

**Study of the Genetic polymorphism and Methylation profile in
the Wnt signalling pathway with risk for occurrence of Lung
cancer in North Indian population**

A

Thesis submitted

in the partial fulfilment of the requirements for the award of degree of

DOCTOR OF PHILOSOPHY

IN

BIOTECHNOLOGY



Submitted by

CHARU BAHL

(Reg. No. 901200003)

Under the supervision of

Dr. SIDDHARTH SHARMA

Associate Professor

Department of Biotechnology

Thapar Institute of Engineering and Technology, Patiala, 147004, Punjab

November 2019

Dedicated to My

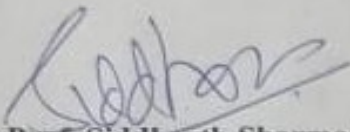
Papa . . .

Late Shri. Ashwani

Kumar Bahl

CERTIFICATE

Certified that the thesis entitled "Study of the Genetic polymorphism and Methylation profile in the Wnt signalling pathway with risk for occurrence of Lung cancer in North Indian population" submitted by Ms. Charu Bahl, Reg. no. 901200003 in the partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy in the Department of Biotechnology, Thapar Institute of Engineering & Technology, Patiala, Punjab is a record of candidate's own independent and original research work carried out by himself under my supervision and guidance. The material embodied in this thesis has not been submitted in part or full to any other University or institute for the award of any degree.



Assoc. Prof. Siddharth Sharma

Supervisor

Department of Biotechnology

Thapar Institute of Engineering &

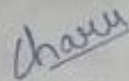
Technology, Patiala, Punjab

CANDIDATE'S DECLARATION

I, hereby declare that the work presented in the thesis entitled "Study of the Genetic polymorphism and Methylation profile in the Wnt signalling pathway with risk for occurrence of Lung cancer in North Indian population" in the partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy at Department of Biotechnology, Thapar Institute of Engineering & Technology, Patiala is an authentic record of my own work during the period from August 2012 to November 2017, under the supervision of Dr. Siddharth Sharma, Assoc. Professor, Department of Biotechnology, Thapar University. This report has not been submitted for the award of any degree or certificate in this or any other university.

Place: Patiala, Punjab

Date:


Charu Bahl

(Reg no. 901200003)

ACKNOWLEDGEMENT

“In the realm of ideas everything depends on enthusiasm...in the real world all rests on perseverance.” - Johann Wolfgang von Goethe

Pursuing for Ph.D was an idea to me planted by the enthusiasm of dreams. Persevering for 5 years and 150 yards to the Department of Biotechnology, Thapar University; has made it possible to accomplish my work and producing this piece of thesis after conducting several trials and finding the new perspective in form the research work. A keen interest in diseases related to lungs and their complexity led the drive in me to push hard for this critical work. Guiding my path was encouragement from my family and constant faith of my mentor Dr. Siddharth Sharma.

It wasn't been without the guidance of my supervisor **Dr. Siddharth Sharma**, Associate Professor who portrayed my dream into the substantial research and paved the pathway in a sequential phased manner, this work would have left to be completed. Apart from the source of vast knowledge, I was blessed to be incubated in an atmosphere of work ethics, boosted morale, high performing work culture and strong values under the motivation of Dr. Siddharth Sharma. Whenever there is a need of inducing a fresh effort, whenever there is a need of fresh perspective, whenever there is a need to overview the research activities, the tall and respectable stature of Dr Siddharth Sharma bails me out. This piece of thesis is encrypted by the colossal efforts of Dr. Sharma. It's very difficult to thank him in words for being the source of my pillar strength and in every other aspect.

I would like to express my gratitude towards **Dr. Moushumi Ghosh**, Professor and Head of Department for she made sure that the infrastructure surely should suffice with our requirements.

This work was supported by grant from the Indian Council of Medical Research, New Delhi, India. (Grant No. 5/13/126/2011/NCD-III).

I would also like to thank **Dr. M.S. Reddy** and **Dr. Dinesh Goyal**, as most of my practical work was completed under their tenure of heading the department. I am extremely overwhelmed for showing a great trust in my conduct and supporting me during critical situations. Their student friendly decisions always helped me grow and overcome all the technical hiccups.

Besides my mentor, **Dr. Handa, Dr. Prakash, Dr. Das, Dr. Saxena** and **Dr. Anita Rajour** played a pivotal role for being a source of inspiration during all the years. They enlightened

me and provided a valuable support while overcoming various challenges. They also gave me precious words of conviction and always believed in my abilities.

I am grateful to **Dr. Digamber Behera**, Head of Department, Department of Pulmonary Medicine, PGIMER who gave this opportunity to have access to the hospital and encourage research work in this field. I would also like to thank **Dr. S.K. Jindal** who was allowed us to collaborate with Department of Pulmonary Medicine, during his tenure of heading the department. He is a great human being and his conduct always inspires me to good great work for humanity.

My sincere gratitude to **Dr. Navneet Singh**, Associate Prof, Department of Pulmonary Medicine, PGIMER. His contribution to this research work is immense as he is the co-supervisor in the ICMR funded project. This fund provided all the necessary consumables required for this work. He enabled us to carry all the sampling work in the hospital. His kind and contemplating nature has always helped me to get comfortable in the environment of the hospital. The collection of clinical data and other information from lung cancer clinic would not have been possible without his presence. Besides, helping in the research activities he has been a great inspiration for me. His passion and grounded nature is a quality I look forward to. His support has been impeccable all through this journey. I have learnt so much being around him and for that I cannot thank him enough.

The other staff members of Department of Pulmonary Medicine, provided me envisions to carry my research work to the ground level. Mr. Santosh and Mr. Parmod were always helpful and help carry out the sampling work. They treated as the part of PGIMER family. Other members of Pulmonary Function Test lab, Mr. Pushpinder, Mr. Chandan, Mrs. Monica, Mr. Babu Ram and Mr. Mandeep help me get trained for sampling work and gave me a friendly environment to work. This experience of working with them has been a lifelong one and it gave confidence to pursue this work.

I would like to mention about my fellow lab mates Amrita Singh, Shweta Lawania, Kinshuk Sharma and Harleen Kaur who helped me complete this endeavor of mine.

I would like to thank PhD scholars **Dr. Raghvendra Amenendi, Dr. Kunal Garg, Dr. Vineet Mesharam, Dr. Nadeem, Dr. Seema, Dr. Sanjog Garyali, Dr. Gurpreet Khaira, Dr. Navdeep Kaur, Dr. Neha Lohia** for guiding me frequently and sharing the latest techniques and development in the scientific world.

I would like to express the sincere gratitude to the fellow research inmates who became my strength during tough times **Prerna** and **Vagish**, for inculcating the research oriented habits. Both of them constantly stood beside me and guided me. They took care of me as family and

always boosted my confidence. They guided a lot during thesis writing process and helped me catch trivial errors.

I like to express my deep joy to work with all the M.Sc and Mtech students who gave a chance to learn more about my field. Their presence kept my enthusiasm high and my passion for this field grew during all these years. Inderjot Madaan, Zeena Anand, Yashila Girdhar, Lavanya Saxena, Prabh Pannu, **Sheetal**, Sheena Bhargava and **Harleen**, with whom I share a good repertoire and a deep mutual understanding.

I am thankful to the lab technicians **Mr. Ram Naval (Babban Ji)**, **Mr. Lallan Mr. Surinder**, **Mrs Lalita** who were always relentless in assisting me for carrying out the research work. Words are not enough to express my sincere thanks to Babban Ji, he is the heart and soul of the department. He has been a constant support and always appreciated my efforts. In my whole journey, he has played a key role. Those conversations and his cute smile always lighted my day. I really respect him for his dedication towards students and providing them with best possible solutions.

No matter how did and what kind of work is completed, the supporting role of the well-wishers and friend played the fulcrum role. To begin with; **Apruva Bhardwaj**, **Piyush**, **Shashank Srivastava**, **Varun Bajoria** and **Nitika Malhotra** who have been a source of huge motivation when and always encouraged carrying out my endeavours. I imbibe the virtue of relaxing in the company of Vipul Roshyan, Ishan Bhateja, Swati Thakur, Chandni Bansal, Amit and Harish. My school friends; Abha Gupta, Priyanka Kathuria, Namrata Deol and Tanvi Jain always gave me the necessary motivation whenever I felt devastated.

Nonetheless, friends and wellwishers know me my thick and thin and there are no ways to express any form of thanking to them. I pay tribute to my father **Late Shri. Ashwani Kumar Bahl**, for whom I am dedicating my research work. It was his dream and this dream along with his love gave me strength to overcome all the hurdles and accomplish this. His absence in my life created emptiness but I never gave up and used this to my strength. I always feel his presence and his blessings are the sole reason of my successful completion of this much awaited milestone of my life.

My mother **Mrs Sangeeta Bahl**, my chachi **Mrs Reepa Bahl** and a special mention of my Chachu, **Mr. Sanjay Bahl** who provide me an encouragement and valuable support to carry out my dreams. His love has been unconditional for all my life. He is my biggest strength and my idol also. I have grown up looking up to him and the way he loves the family. He nurtured every dream of mine and always appreciated my efforts. I am blessed to have him

around. He took care of each and every need of mine. I am so thankful to God who has given me such a lovely family.

Siblings **Vishu Bahl, Sagar Bahl & Diya Bahl** are always a source of motivation for whom I want to set the example. My love for them has been my one of the driving force to perform well and give them happiness. I would like to thank my grandfather, **Late Shri. Banshi Ram**, who has blessed all of us and always taught us the right virtues of living.

Date: 21/01/2020

Charu Bahl

Place: Patiala

LIST OF PUBLICATIONS

Published

1. Bahl, C., Sharma, S., Singh, N., & Behera, D. (2017). Association study between genetic variations in *Axin2* gene and lung cancer risk in North Indian population: A multiple interaction analysis. *Tumor Biology*, 39(4), 1010428317695533. Impact factor: 3.65
2. Bahl, C., Singh, N., Behera, D., & Sharma, S. (2017). Association of polymorphisms in Dickkopf (*DKK*) gene towards modulating risk for lung cancer in North Indians. *Future Oncology*, 13(3), 213-232. Impact factor: 2.13
3. Bahl, C., Sharma, S., Singh, N., & Behera, D. (2017). Single nucleotide variations in the Wnt antagonist *sFRP3* (rs7775 & rs288326) and *sFRP4* (rs1802073 & 1802074) genes and their association with lung cancer risk in North Indians. *Meta Gene*, 13, 159-168.
4. Bahl, C., Singh, N., Behera, D., & Sharma, S. (2018). Genetic variants in the Wnt antagonist genes (*sFRP*, *Dkk* & *Axin2*) predicts the overall survival and prognosis of North Indian lung cancer patients treated with platinum based doublet chemotherapy. *Cancer biotherapy & radiopharmaceuticals*. doi.:org/10.1089/cbr.2018.2491. Impact factor: 1.68
5. Bahl, C., Sharma, S., Singh, N., & Behera, D. (2017). High order gene interactions between the genetic polymorphisms in Wnt and AhR pathway towards lung cancer susceptibility. *Personalized Medicine*, 14(6), 487-502 Impact factor:1.01
6. Budhwar, S., Bahl, C., Sharma, S., Singh, N., & Behera, D. (2017). Role of sequence variations in AhR gene towards modulating smoking induced lung cancer susceptibility and Clinical outcomes in North Indian population: A multiple interaction analysis. *Current Genomics*, 18, doi: 10.2174/1389202918666170915160606.
7. Kumari, A., Bahl, C., Singh, N., Behera, D., & Sharma, S. (2016). Association of *p53* codon 72 polymorphism and survival of North Indian lung cancer patients treated with platinum-based chemotherapy. *Molecular Biology Reports*, 43(12), 1383-1394. Impact factor: 1.82
8. Sodhi, K. K., Bahl, C., Singh, N., Behera, D., & Sharma, S. (2015). Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians. *Future Oncology*, 11(15), 2159-2173. Impact factor: 2.13
9. Pandey, A., Bahl, C., Singh, N., Behera, D., & Sharma, S. (2018). Functional role of CyclinD1 polymorphism (G870A) in modifying susceptibility and overall survival of North Indian lung cancer patients. *Tumori*, 104(3), 179-187. Impact factor: 1.23
10. Bhardwaj, A., Bahl, C., Singh, N., Behera, D., & Sharma, S. (2017). Interactive potential of genetic polymorphism in Xenobiotic metabolising and DNA repair genes in predicting lung cancer predisposition in North Indians". *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 826, 15-24. Impact factor: 1.99

Workshops and Conferences

- Workshop on Biostatistics and Epidemiology for Bio-medical PG students at BETC Training center in the Dept. of Biostatistics & Epidemiology, Institute of Cytology and Preventive oncology, Noida.
- 34th Annual Convention of Indian Association for Cancer Research, IACR Feb, 2015 in Jaipur. Presented a poster on “Genetic Variants in pre micro RNA -*146a* (rs2910164) and *196a2* (rs11614913) genes are associated with risk of Lung Cancer in North Indian Population: A case control study”.
- International conference titled “Global Cancer Summit-2015” with the theme “Frontiers in Oncology: Genetics, Diagnostics and Therapeutics” that was held in Indian Institute of Sciences, Bangalore, India from 18th- 20th Nov, 2015. Presented an oral power point on “Single nucleotide polymorphic variants (rs288326, rs7775, rs1802073 and rs1802074) in Secreted Frizzled related Protein modulate Lung Cancer Susceptibility”

ABSTRACT

Background: Wnt pathway has been implicated in the process of lung carcinogenesis. The genetic differences and methylation profile in the Wnt antagonist genes (*DKK*, *Axin2* and *sFRP*) may lead to chronic activation of this cascade. **Objectives:** To evaluate the role of Single nucleotide polymorphic variants of *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK2* (rs447372, rs419558 and rs17037102) and *DKK4* (rs2073664) genes, *Axin2* (148 C/T, 1365 G/A, 432 T/C, 956+16 A/G, 1712+19 G/T, 1386 C/T and 2062 C/T), *sFRP3* (rs7775 and rs228326) and *sFRP4* (rs1802073 and rs1802074). Also, to detect the methylation status of *sFRP1*, *DKK2*, *DKK3* and *Axin2* gene. **Methodology:** Genotyping of genomic DNA was carried out using PCR-RFLP (polymerase chain reaction-restriction fragment length polymorphism) for each polymorphic site under study. Total number of subjects genotyped in case of *DKK*, *Axin2* and *sFRP* genetic variants were 600 (300 cases and 300 controls), 608 (303 cases and 305 controls) and 610 (300 cases and 310 controls) respectively. Further association analysis was carried out using logistic regression approach to obtain adjusted odds ratio and statistical significance. MDR and CART analysis were applied to evaluate high order interactions between the SNP's of same gene. Also, MDR and CART was also used to analyze the high order interaction between the genetic variants of different genes. Overall Survival was estimated using Kaplan meier survival analysis and adjusted hazards ratio was obtained using cox regression method. DNA isolated from plasma samples of 100 lung cancer patients were screened for the methylation status of promoter region of *sFRP1*, *DKK2*, *DKK3* and *Axin2* gene using methylation specific PCR. Further, the methylation status was correlated with the clinicopathological parameters such as stage, primary tumor extension, lymph node involvement and distant metastasis using logistic regression approach. **Results:** *DKK4* (rs2073664) did not confer any effect on lung cancer susceptibility. Also, the two variants of *DKK3* namely rs3206824 and rs2291599 didn't show any association except rs7396187 which showed a protective effect (OR=0.63, $p=0.01$). However, the subjects with heterozygous genotype of *DKK2* rs17037102 and rs419558 were found to have a higher risk of developing lung cancer. *DKK3* rs3206824 and *DKK2* rs419558 showed a two-fold increased risk of developing lung cancer ($p=0.008$). None of the polymorphic site was found to be associated with the clinic-pathological features. Three out of seven studied polymorphic sites studied for *Axin2* gene showed a strong protective effect in subjects having mutant genotype for *Axin2* 148 C/T and heterozygous genotype for 1365 G/A towards lung cancer risk. The other important finding was the significant association of *Axin2* 148 C/T in SQCC

patients having variant (TT) genotype. Unfavorable genotypes of the three *DKK2* variants collectively (rs447372, rs419558 and rs17037102) exhibited a highly decreased rate of death (HR'=0.37, ***p*=0.03**). ADCC patients carrying the heterozygous (CT) genotype for *DKK4* rs2073664 showed a better OS (log rank ***p*=0.01**). ADCC subjects having TT genotype for *Axin2* 148 showed a better prognosis (HR'=0.48, ***p*=0.003**) and SCLC patients with TT genotype for *Axin2* 1386 showed a poor prognosis (HR'=2.33, ***p*=0.02**). Survival tree analysis depicted *DKK4* rs2073664 as the major contributor. Findings from the present study revealed that the rate of methylation of promoters in plasma of lung cancer patients for *sFRP1*, *DKK3*, *Axin2* and *DKK2* was found to be 24%, 45%, 25% and 10% respectively. **Conclusion:** This preliminary study gave an insight that about the potential of genetic differences in Wnt antagonist genes to modulate lung cancer susceptibility. Genetic variations in Wnt antagonist, *DKK2* (rs419558 and rs17037102) were certainly playing a crucial role in modulating susceptibility of North Indians as compared to *DKK3* and *DKK4*. *Axin2* 148 was also found to play a major. Survival analysis revealed that a cumulative effect of three variants of *DKK2* gene along with *DKK4* rs2073664 can jointly predict the survival of lung cancer patients.

TABLE OF CONTENTS

	Page No.
<i>Acknowledgement</i>	
<i>List of Publications</i>	
<i>Content</i>	
<i>Abstract</i>	
<i>List of figures</i>	
<i>List of tables</i>	
<i>List of abbreviations</i>	
<i>List of symbols</i>	
Chapter	
1. Introduction	1-6
2. Review of Literature	7-55
2.1 Lung cancer	7-25
2.1.1 Epidemiology-World Scenario	11
2.1.2 Epidemiology-Indian Scenario	12
2.1.3 Risk factors for lung cancer	12-15
2.1.4 Histopathological classification of lung cancer	15-18
2.1.5 Molecular pathogenesis of lung cancer	18-23
2.1.6 Diagnosis for lung cancer	23
2.1.7 Treatment for lung cancer	25-26
2.2 Wnt pathway	26-55
2.2.1 History of Wnt pathway	26-29
2.2.2 Canonical Wnt pathway	30-31
2.2.3 Wnt antagonists and agonists	32-34
2.2.3.1 Secreted Frizzled related receptor (sFRP)	33-34
2.2.3.2 Dickkopf proteins (DKK)	34-35
2.2.4 Wnt Receptors and their interactions	35-37
2.2.5 Destruction complex affecting the stability of β -catenin	37-39
2.2.5.1 Axin	39
2.2.6 Role of Wnt signaling in development	39-40
2.2.7 Role of Wnt signalling in stem cell maintenance	40-42
2.2.8 Wnt signalling in cancer	42-47
2.2.8.1 Wnt signaling and lung cancer	46-47
2.2.9 Genetic variations and their relevance in modulating disease susceptibility and clinical outcomes	48-49
2.2.10 Genetic variabilities in the Wnt antagonist genes (<i>sFRP</i> , <i>DKK</i> and <i>Axin2</i>) and their correlation with cancer predisposition	49-51
2.2.11 Association of genetic variations within the Wnt antagonist genes with overall survival and prognosis of cancer patients	51
2.2.12 Epigenetic alterations in the Wnt signalling pathway and their role in carcinogenesis	53-55
3. Gap in Study and Objectives	56
4. Materials and Methods	57-73
4.1 Design of the study and study subjects	57-60

4.1.1	Chemotherapeutic regimen	57
4.1.2	Follow up and response assessment	58-60
4.2	DNA Extraction	60-61
4.2.1	Genomic DNA Extraction	60
4.2.2	DNA Extraction from plasma	60
4.2.3	Qualitative and quantitative estimation of genomic DNA	61
4.3	Genotyping of polymorphic sites of <i>DKK</i> , <i>sFRP</i> and <i>Axin2</i> gene	61-67
4.3.1	Genotyping for <i>DKK</i> variants	61-64
4.3.2	Genotyping for <i>Axin2</i> variants	64-66
4.3.3	Genotyping for <i>sFRP</i> variants	66-67
4.4	Statistical analysis	68-70
4.5	Methylation-Specific Polymerase Chain Reaction (MSP) Analysis for <i>DKK</i> , <i>sFRP</i> and <i>Axin2</i> promoter region	70-73
5. Results		74-22
5.1	Genomic DNA isolation	74
5.2	PCR amplification	74-75
5.3	Genotyping by RFLP	75-80
5.3.1	Genotyping for <i>DKK</i> genetic variants	75-77
5.3.2	Genotyping for <i>Axin2</i> genetic variants	77-79
5.3.3	Genotyping for <i>sFRP</i> genetic variants	80-81
5.4	Epidemiological statistics of North Indian population under study	81-8
5.4.1	Regional Distribution of lung cancer patients	81-82
5.4.2	Distribution of study subjects on the basis of smoking	82-83
5.4.3	Distribution of study subjects on the basis of age	84
5.4.4	Distribution of lung cancer patients on the basis of stage of tumor and TNM	84-85
5.4.5	Distribution of lung cancer patients on the basis of histology of tumor	85-86
5.4.6	Distribution of lung cancer patients on the basis of performance status	86-87
5.4.7	Distribution of lung cancer patients on the basis of doublet platinum based chemotherapeutic regimen	87-88
5.5	Statistical Analysis for polymorphic sites of <i>DKK</i> gene	88-126
5.5.1	Demographic characteristics	88-90
5.5.2	Genotypic and allelic frequencies of <i>DKK</i> polymorphic variants among cases and controls and their role in modifying the susceptibility towards lung cancer	90-93
5.5.3	Genotypic and allelic distribution of the <i>DKK2</i> , <i>DKK4</i> and <i>DKK3</i> gene and risk of lung cancer on basis of histological subtypes	93-95
5.5.4	Effect of smoking status on the association of <i>DKK</i> gene polymorphisms and lung cancer risk	96-99
5.5.5	Association between lung cancer risk and different combinations of <i>DKK4</i> , <i>DKK3</i> and <i>DKK2</i> genotypes	100-101
5.5.6	Linkage Disequilibrium and haplotype analysis in <i>DKK2</i> and <i>DKK3</i> Gene Polymorphisms	102
5.5.7	Relationship of genetic variations in <i>DKK</i> gene with clinico-pathological parameters	102-106
5.5.8	Multiple dimensionality reduction analysis of genetic variants within <i>DKK</i> gene	107-108
5.5.9	Classification and regression tree analysis of genetic variants within <i>DKK</i> gene	108-109
5.5.10	Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of <i>DKK</i> polymorphism	110

5.5.11	Association of <i>DKK</i> genetic variants with overall survival of lung cancer patients	110-116
5.5.12	Association of Wnt antagonist (<i>DKK</i>) polymorphism and overall survival according to histological subtype of lung cancer patients	117-124
5.5.13	Association of Wnt antagonist (<i>DKK</i>) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response	125-126
5.6	Statistical Analysis of polymorphic sites of <i>Axin2</i> gene	126-162
5.6.1	Demographic characteristics	126-128
5.6.2	Genotypic and allelic frequencies of the <i>Axin2</i> variants and their association with overall lung cancer risk	129-130
5.6.3	Genotypic and allelic frequencies of the <i>Axin2</i> variants and their association with overall lung cancer risk and on basis of histological subtypes	130-133
5.6.4	Effects of smoking status and its association between <i>Axin2</i> gene polymorphisms and lung cancer risk	135-138
5.6.5	Combinatorial analysis of the risk associated with different polymorphic sites of <i>Axin2</i> gene	139-141
5.6.6	Relationship of genetic variations in <i>Axin2</i> gene with clinicopathological parameters	142-145
5.6.7	MDR analysis within <i>Axin2</i> gene	146-147
5.6.8	CART analysis within <i>Axin2</i> gene	147-149
5.6.9	Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of <i>Axin2</i> polymorphism	150
5.6.10	Association of <i>Axin2</i> genetic variants with overall survival of lung cancer patients	150-153
5.6.11	Association of Wnt antagonist (<i>Axin2</i>) polymorphism and overall survival according to histological subtype of lung cancer patients	154-160
5.6.12	Association of Wnt antagonist (<i>Axin2</i>) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response	161-162
5.7	Statistical Analysis of polymorphic sites of <i>sFRP</i> gene	163-179
5.7.1	Demographic characteristics	163
5.7.2	Genotypic and allelic distribution of the four genetic variants and their association with risk of lung cancer	163-166
5.7.3	Genotypic and allelic distribution of the four genetic variants and the stratified association analysis based on histological subtypes	167
5.7.4	Genotypic distribution of the four genetic variants based on smoking status and its association with risk of developing lung cancer	168
5.7.5	Clinico-pathological parameters and their association with genotype	169
5.7.6	Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of <i>sFRP</i> polymorphism	170
5.7.7	Association of <i>sFRP</i> genetic variants with overall survival of lung cancer patients	170-172
5.7.8	Association of Wnt antagonist (<i>sFRP</i>) polymorphism and overall survival according to histological subtype of Lung cancer patients	173-177
5.7.9	Association of Wnt antagonist (<i>sFRP</i>) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response	178-179

5.8	Multiple gene interaction analysis (<i>sFRP</i> , <i>DKK</i> and <i>Axin2</i>)	178-196
5.8.1	Demographic characteristics	178-179
5.8.2	Genotypic distribution of double combinations of polymorphisms between Wnt antagonist genes (<i>sFRP</i> , <i>DKK</i> and <i>Axin2</i>) and their association with lung cancer predisposition	180-187
5.8.3	Association of different unfavorable genotypic combinations of <i>DKK2</i> , <i>DKK3</i> , <i>sFRP3</i> and <i>sFRP4</i> gene with overall survival of lung cancer patients	188
5.8.4	High order gene-gene interactions for prognosis of lung cancer patients using survival tree analysis	189-192
5.9	Methylation-Specific Polymerase Chain Reaction (MSP) Analysis for <i>DKK</i> , <i>sFRP</i> and <i>Axin2</i> promoter region	193-194
5.9.1	Relationship of various clinic-pathological features with methylation profile of Wnt antagonists genes	194-196
6. Discussion		197-214
6.1	Wnt antagonist (<i>DKK</i>) polymorphism and lung cancer predisposition	198-200
6.2	Wnt antagonist (<i>Axin2</i>) polymorphism and lung cancer predisposition	201-203
6.3	Wnt antagonist (<i>sFRP</i>) polymorphism and lung cancer predisposition	203-205
6.4	Gene-gene and gene-environment interactions	205-208
6.5	Wnt antagonists and their role in overall survival and objective response of lung cancer patients receiving doublet platinum based chemotherapy	208-212
6.6	Methylation profile of Wnt antagonist genes (<i>DKK2</i> , <i>DKK3</i> , <i>Axin2</i> and <i>sFRP1</i>) and its relationship with various clinico-pathological parameters	212-214
7. Conclusion		215-216
8. Bibliography		217-248
Appendix-I		249-251
Appendix-II		252-254
Appendix-III		255

LIST OF FIGURES

Fig. No.	Figure legend	Page No.
Review of Literature		
2.1	Schematic representation of Lung tissue; a: Anatomical features of lung showing trachea, bronchioles, alveoli and diaphragm; b: Lung tissue harboring a solid neoplastic growth along with lymph node infiltration and mediastinum.	8
2.2	Estimated incidence rate (per 100,000) of lung cancer worldwide which are standardised for age.	11
2.3	Microscopic images of different histological subtypes of lung cancer; a: small cell lung cancer; b: squamous cell carcinoma; c: adenocarcinoma; d: large cell carcinoma	18
2.4	Schematic representation of major molecular abnormalities taking place during lung cancer pathogenesis	23
2.5	Schematic representation of Wnt/ β -catenin signaling; a: showing the signal transduction when Wnt pathway is in off state; b: showing the signal transduction when Wnt pathway is in on state	32
2.6	Schematic representation of Wnt antagonists regulating the pathway at different subcellular levels; a: secretory antagonists of Wnt pathway; b: agonists of Wnt pathway	33
2.7	Overview of structure of frizzled receptors showing its different domains	35
2.8	Protein domains of Dishevelled, Axin, APC, β -catenin and LEF-1 found out either by sequence similarity or by functional assays	36
2.9	Schematic representation of destruction complex for β -catenin degradation regulated by Axin2	37
2.10	Destruction complex showing the phosphorylated sequence of β -catenin at N-terminal carried out by CK1 and GSK-3. The acidic cluster that primes CK1 phosphorylation is marked with gray colour; SCF: Skp1/cullin/F-box complex	38
2.11	Dysregulated canonical Wnt signalling in cancer. Constitutive expression of Wnt occurs due to lack of regulation at different levels of signal cassette in different cancers which has been illustrated	44
2.12	Relation between the tumor cells which are differentiated and cancer stem cells. Cancer stem cells show a comparatively higher levels of Wnt than	46

	differentiated tumor cells. The activity of Wnt depends on both extrinsic and intrinsic factors	
2.13	Epigenetic silencing of negative WNT regulators contributes to the aberrant activation of WNT/ β -catenin signaling in human cancers.	54
Materials and Methods		
4.1	STRAD diagram explains the selection criteria of the patients for overall survival analysis	59
4.2	Schematic representation of thermal cycling conditions used; Stage 1 : Initial Denaturation at 95°C for 5 min; Stage 2 : 30 cycles of Denaturation at 94°C for 30 sec, Annealing at annealing temperature for 45 sec and Extension at 72°C for 30 sec; Stage 3 : Final extension at 72°C for 5 min	62
Results		
5.1	Representative example of agarose gel electrophoresis showing genomic DNA bands	74
5.2	Representative example of agarose gel electrophoresis showing plasmid DNA bands	74
5.3	Representative example of agarose gel electrophoresis showing PCR amplification for <i>Axin2</i> 1365 G/A	75
5.4	Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for <i>DKK</i> variants ; A: <i>DKK2</i> rs17037102 G/A, B: <i>DKK2</i> rs419558 C/T, C: <i>DKK4</i> rs2073664 C/T, D: <i>DKK3</i> rs3206824 G/A, E: <i>DKK3</i> rs2295199 C/T, F: <i>DKK2</i> rs447372 A/G	76
5.5	Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for <i>Axin2</i> variants; A: <i>Axin2</i> 2062 C/T, B: <i>Axin2</i> 1386 C/T, C: <i>Axin2</i> 148 C/T, D: <i>Axin2</i> 956+16 A/G, E: <i>Axin2</i> 1365 G/A, F: <i>Axin2</i> 432 T/C, G: <i>Axin2</i> 1712+19 G/T	78-79
5.6	Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for <i>sFRP</i> variants; A: <i>sFRP3</i> rs7775 C/G, B: <i>sFRP3</i> rs288326 C/T, C: <i>sFRP4</i> rs 1802073 C/A, D: <i>sFRP4</i> rs1802074 G/A	80
5.7	Pie chart illustrating regional distribution of North Indian Lung cancer patients	81
5.8	Pie chart illustrating regional distribution of North Indian Lung cancer patients a: regional distribution of males; b: regional distribution of females	82
5.9	Graph showing the distribution of study objects on the basis of smoking status.	83

	a: Among cases, b: Among controls	
5.10	Graph showing the distribution of male study subjects on the basis of smoking status	83
5.11	Graph showing the age wise distribution of lung cancer on the basis of gender	84
5.12	Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of clinical stage	85
5.13	Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of histological subtypes	86
5.14	Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of karnofsky performance status (KPS)	87
5.15	Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of Eastern Cooperative Oncology Group-performance status (ECOG-PS)	87
5.16	Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of Doublet platinum based chemotherapeutic regimen	88
5.17	MDR interactions with <i>DKK</i> variants.	108
5.18	CART analysis of <i>DKK</i> variants; W: wild type; M: variant type.	109
5.19	Kaplan meir survival curves for <i>DKK</i> variants; A: <i>DKK4</i> rs2063664 C/T (CC: reference), B: <i>DKK3</i> rs3206824 G/A (GG: reference), C: <i>DKK2</i> rs447372 A/G (AA: reference) for lung cancer patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time	115
5.20	Kaplan meir survival curves for <i>DKK</i> variants in lung cancer patients stratified according to their histology; A: <i>DKK4</i> rs2063664 C/T (CC: reference) in ADCC patients, B: <i>DKK3</i> rs3206824 G/A (GG: reference) in ADCC patients, C: <i>DKK3</i> rs3206824 G/A (GG: reference) in SQCC for lung cancer patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time, ADCC: adenocarcinoma, SQCC: squamous cell carcinoma	123-124
5.21	Multifactor dimensionality reduction (MDR) interaction dendrogram.	147
5.22	CART analysis of <i>Axin2</i> variants; W: wild type; M: variant type.	149
5.23	Kaplan meir survival curves for <i>Axin2</i> variants; A: <i>Axin2</i> 148 C/T (CC: reference), B: <i>Axin2</i> 432 T/C (TT: reference), C: <i>Axin2</i> 1712+19 G/T (TT: reference) for lung cancer patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time	153
5.24	Kaplan meir survival curves for <i>Axin2</i> variants in lung cancer patients	159

	stratified according to their histology; A: <i>Axin2</i> 148 C/T (CC: reference) in ADCC patients, B: <i>Axin2</i> 432 T/C (TT: reference) in ADCC patients, C: <i>Axin2</i> 1712+19 G/T (TT: reference) in SQCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time, ADCC: adenocarcinoma, SQCC: squamous cell carcinoma	
5.25	Kaplan meir survival curves for <i>Axin2</i> variants in lung cancer patients stratified according to their histology; D: <i>Axin2</i> 432 T/C (TT: reference) in SQCC patients, E: <i>Axin2</i> 1712+19 G/T (TT: reference) in ADCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time, ADCC: adenocarcinoma, SQCC: squamous cell carcinoma	160
5.25	Kaplan meir survival curves for <i>sFRP3</i> rs7775 in ADCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time	163
5.26	Kaplan meir survival curves for <i>sFRP</i> variants in lung cancer patients stratified according to their histology; A: <i>sFRP3</i> rs7775 C/G (CC: reference) in SQCC patients, B: <i>sFRP3</i> rs7775 C/G (CC: reference) in SCLC patients, C: <i>sFRP4</i> rs 1802073 C/A (CC: reference) in ADCC patients, D: <i>sFRP4</i> rs 1802074 G/A (GG: reference) in SCLC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time, ADCC: adenocarcinoma, SQCC: squamous cell carcinoma, SCLC: Small cell lung cancer	171
5.27	Survival Tree analysis of <i>sFRP</i> , <i>DKK</i> , and <i>Axin2</i> variants; MST: median survival time in months; HR: Hazards ratio with 95% C.I. (univariate analysis)	191
5.28	Overall survival according to different nodes identified by survival tree analysis	192
5.29	Representative example of methylation specific PCR; a: MS-PCR of <i>sFRP1</i> gene, b: MS-PCR of <i>DKK3</i> gene, c: MS-PCR of <i>Axin2</i> gene, d: MS-PCR of <i>DKK2</i> gene, 1: marker DNA, UM: unmethylated primer, M: methylated primer, NC: negative control	193

LIST OF TABLES

Table no.	Table title	Page no.
Review of Literature		
2.1	TNM stage groupings	9
2.2	Tumor, node, and metastasis (TNM) definitions	10
2.3	List of occupational hazards	14
2.4	Oncogenic Driver mutations commonly found in Adenocarcinoma	24
2.5	Milestones of Wnt pathway in cancer biology	28-30
2.6	Genetic variations of the negative regulators of Wnt signalling cascade in different human cancers	52
2.7	Dysregulated epigenetic activation of the genes of Wnt signalling cascade in different human cancers	55
Materials and Methods		
4.1	List of primers and restriction enzymes used for <i>DKK</i> genetic variants	63
4.2	List of primers and restriction enzymes used for <i>Axin2</i> genetic variants	66
4.3	List of primers and restriction enzymes used for <i>sFRP</i> genetic variants	67
4.4	List of primers used for Methylation specific PCR	73
Results		
5.1	Demographic characteristics among cases and controls	89-90
5.2	Genotypic and allelic distribution of the <i>DKK</i> genetic variants and their association with risk of lung cancer	92-93
5.3	Genotypic distribution of the <i>DKK</i> genetic variants and their association with risk of lung cancer based on histology	95
5.4 (a)	Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer (smokers)	98
5.4 (b)	Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer (non-smokers)	99
5.5	Genotypic distribution based on different genotypic combinations and their association with lung cancer risk	101
5.6 (a)	Haplotype frequencies and lung cancer risk associated with the <i>DKK3</i> haplotypes	102

5.6 (b)	Haplotype frequencies and lung cancer risk associated with the <i>DKK2</i> haplotypes	102
5.7 (a)	Relationship of different genotypes of <i>DKK</i> gene with the clinico-pathological parameters (stage and primary tumor extension)	105
5.7 (b)	Relationship of different genotypes of <i>DKK</i> gene with the clinico-pathological Parameters (lymph node invasion and metastasis)	106
5.8	Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions of <i>DKK</i> variants with lung cancer	107
5.9	Risk estimates based on CART analysis for <i>DKK</i> gene	109
5.10	Distribution of demographic and clinical parameters	111
5.11	Univariate and Multivariate analysis for Wnt antagonist (<i>DKK</i>)	116
5.12 (a)	Association of <i>DKK</i> SNPs on overall survival according to tumor histology (ADCC)	120
5.13 (b)	Association of <i>DKK</i> SNPs on overall survival according to tumor histology (SQCC)	121
5.13 (c)	Association of <i>DKK</i> SNPs on overall survival according to tumor histology (SCLC)	122
5.13	Relationship of different <i>DKK</i> genotypes with the chemotherapy response	126
5.14	Demographic characteristics among cases and controls	128
5.15	Genotypic distribution of the <i>Axin2</i> genetic variants and their association with risk of lung cancer	132
5.16	Genotypic distribution of the <i>Axin2</i> genetic variants and their association with risk of lung cancer based on histology	134
5.17 (a)	Genotypic distribution of the <i>Axin2</i> variants based on smoking status and its association with risk of lung cancer (smokers)	137
5.17 (b)	Genotypic distribution of the <i>Axin2</i> variants based on smoking status and its association with risk of lung cancer (non-smokers)	138
5.18	Genotypic distribution based on different genotypic combinations of <i>Axin2</i> variants and their association with lung cancer risk	140-141
5.19 (a)	Relationship of different genotypes of <i>Axin2</i> variants with the clinico-pathological parameters (stage and primary tumor extension)	144
5.19 (b)	Relationship of different genotypes of <i>Axin2</i> variants with the clinico-pathological parameters (lymph node invasion and metastasis)	145

5.20	Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions of <i>Axin2</i> variants with lung cancer	146
5.21	Risk estimates based on CART analysis for <i>Axin2</i> variants	148
5.22	Univariate and Multivariate analysis for Wnt antagonist (<i>Axin2</i>)	152
5.23 (a)	Association of <i>Axin2</i> SNPs on overall survival according to tumor histology (ADCC)	156
5.23 (b)	Association of <i>Axin2</i> SNPs on overall survival according to tumor histology (SQCC)	157
5.23 (c)	Association of <i>Axin2</i> SNPs on overall survival according to tumor histology (SCLC)	158
5.24	Relationship of different <i>Axin2</i> genotypes with the chemotherapy response	162
5.25	Demographic characteristics among cases and controls	164
5.26	Genotypic distribution of the <i>sFRP</i> genetic variants and their association with risk of lung cancer	166
5.27	Genotypic distribution of the <i>sFRP</i> genetic variants and their association with risk of lung cancer based on histology	167
5.28 (a)	Genotypic distribution of the <i>sFRP</i> variants based on smoking status and its association with risk of lung cancer (smokers)	168
5.28 (b)	Genotypic distribution of the <i>sFRP</i> variants based on smoking status and its association with risk of lung cancer (non-smokers)	168
5.29	Relationship of clinico-pathological features and <i>sFRP3</i> & <i>sFRP4</i> genotypes in lung cancer patients	169
5.30	Univariate and Multivariate analysis for Wnt antagonists (<i>sFRP</i>)	172
5.31 (a)	Association of <i>sFRP</i> SNPs on overall survival according to tumor histology (ADCC)	174
5.31 (b)	Association of <i>sFRP</i> SNPs on overall survival according to tumor histology (SQCC)	175
5.31 (c)	Association of <i>sFRP</i> SNPs on overall survival according to tumor histology (SCLC)	176
5.32	Relationship of different <i>sFRP</i> genotypes with the chemotherapy response	179
5.33	Demographic characteristics among cases and controls	179
5.34	Genotypic distribution based on different genotypic combinations and their	183-

	association with lung cancer risk	187
5.35	Association of combination of unfavorable genotypes with overall Survival	188
5.36	High order gene-gene interactions using survival tree analysis	190
5.37	Distribution of demographic and clinical parameters of lung cancer patients analyzed for their methylation profile	194
5.38 (a)	Relationship of various clinic-pathological features with methylation profile of Wnt antagonists genes (stage and primary tumor extension)	196
5.38 (b)	Relationship of various clinic-pathological features with methylation profile of Wnt antagonists genes (lymph node invasion and metastasis)	196

List of Abbreviations

Abbreviated Form

VEGF
 LPR5/6
 GSK-3 β
 Fz
 Dsh
 APC
 TCF/ LEF
 sFRP
 CRD
 NTR
 DKK
 NSCLC
 SNP
 MDR
 CART
 DNA
 IASLC
 FNAC
 PET
 CT
 TNM
 IARC
 NCRP
 ICMR
 PBCR
 HBCR
 NNK
 PAH
 TCDD
 EPA
 SCLC
 LCC
 ADCC
 SQCC
 TTF-1
 CEA
 CAM
 EMA
 Ck
 Rb
 EGFR
 PIK3CA
 mTOR
 MAPK
 ATP
 TK1
 GTP
 ERBB
 BPDE
 EML-4
 ALK

Full Form

Vascular Endothelial Growth Factor
 Lipoprotein Receptor Related Protein 5/6
 Glycogen Synthase Kinase-3 Beta
 Frizzled receptor
 Dishevelled
 Adenomatous Polyposis Coli
 T-Cell Factor/ Lymphoid Enhancer Factor
 Secreted Frizzled Related Protein
 Cysteine Rich Domain
 Netrin Like Domain
 Dickkopf
 Non-Small Cell Lung Cancer
 Single Nucleotide Polymorphism
 Multifactor Dimensionality Reduction
 Classification And Regression Tree Analysis
 Deoxyribonucleic Acid
 International Association for the Study of Lung Cancer
 Fine Needle Aspiration Cytology
 Positron Emission Tomography
 Computed Tomography
 Tumour Node And Metastasis
 International Agency For Research On Cancer
 National Cancer Registry Program
 Indian Council Of Medical Research
 Population Based Cancer Registry
 Hospital Based Cancer Registry
 4-(Methylnitrosamino)-1- (3-Pyridyl)-1-Butanone
 Polycyclic Aromatic Hydrocarbon
 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin
 Environmental Protection Agency
 Small Cell Lung Cancer
 Large cell carcinoma
 Adenocarcinoma
 Squamous Cell Carcinoma
 Thyroid Transcription Factor 1
 Carcinoembryonic Antigen
 Calcium/calmodulin-dependent protein
 Epithelial Membrane Antigen
 Cytokeratin
 Retinoblastoma
 Epidermal Growth Factor Receptor
 Phosphatidylinositol 3-Kinase, Catalytic, Alpha
 Mammalian target of rapamycin
 Mitogen-Activated Protein Kinases
 Adenosine Triphosphate
 Tyrosine kinase inhibitor
 Guanosine Triphosphate
 Erb-b2 receptor tyrosine kinase
 BenzopyreneDiethyloxide
 Echinoderm Microtubule-Associated Protein Like-4
 Anaplastic Lymphoma Kinase

MDM2	Mouse Double Minute 2 Homolog
PTEN	Phosphatase Tensin
TGF- β	Transforming Growth Factor
E2F	Elongation factor
BCL-2	B-Cell Lymphoma-2
PIP3	Phosphatidylinositol (3,4,5)-Trisphosphate
LOH	Loss Of Heterozygosity
FHIT	Fragile histidine triad protein
CD	Cluster of differentiation
Ret	Ret proto-oncogene
KIF5B	Kinesin-1 heavy chain
NCOA4	Nuclear receptor coactivator 4
B-RAF	B-Raf proto-oncogene
MET	Met proto-oncogene
FDA	Food And Drug Administration
MMTV	Mouse Mammary Tumor Virus
m-RNA	Messenger Ribonucleic Acid
Int1	Integration 1
β -TrCp	Beta transducin containing protein
PP2A	Protein phosphatase 2A
WIF1	Wnt Inhibitory Factor 1
SOS1	Son of sevenless homolog 1
LGR5	Leucine-rich repeat-containing G-protein coupled receptor 5
ZNRF3	Zinc and ring finger 3
RNF3	Ring finger protein 3
PCP	Planar cell polarity
CREB	cyclic AMP responsive element binding protein
STAT3	Signal Transducer And Activator Of Transcription 3
AMP	Adenosine monophosphate
RGS	Regulator of G-protein signalling
SCF	Stem cell factor
BMP	Bone morphogenetic protein
Porcn	Porcupine
ESC's	Embryonic stem cells
IPSC's	Induced pluripotent stem cells
Sox-2	Sex Determining Region Y-Box 2
Oct-4	Octamer-Binding Protein 4
Cox-2	Cyclooxygenase-2
MSC's	Mesenchymal Stem Cell
HSC's	Hematopoietic Stem Cell
FAP	Fibroblast activation protein
EMT	Epithelial To Mesenchymal Transition
CRBP	Cellular retinol-binding protein
CR	Complete Response
PR	Partial Response
SD	Stable Disease
PD	Progressive Disease
EDTA	EthyleneDiaminetetraAcetic Acid
TBE	Tris/Borate/Edta
RBC	Red Blood Cells
HR'	Adjusted hazards ratio
ZEB1	Zinc finger E-box-binding homeobox 1
ZEB2	Zinc finger E-box-binding homeobox 2
Hox-B9	Homeobox protein 9
CSC's	Cancer stem cells

PGIMER	Postgraduate Institute of Medical Education and Research
HGF	Hepatocyte growth factor
SDF-1	Stromal cell derived factor-1
PCR	Polymerase Chain Reaction
OD	Optical density
RFLP	Restriction Fragment Length Polymorphism
dNTP	Deoxyribose Nucleoside Triphosphates
BSA	Bovine Serum Albumin
TE	Tris EDTA
NEB	New England biolab
PAGE	Polyacrylamide Gel Electrophoresis
APS	Ammonium Persulfate
TEMED	N,N,N',N' Tetramethylethylene-1,2-Diamine
HWE	Hardy-Weinberg equilibrium
LD	Linkage disequilibrium
FDR	False discovery rate
CVC	Cross validation consistency
MST	Median survival time
OS	Overall Survival
MSP	Methylation specific PCR
DAAM2	DeoxyadenosineMethylase 2
Smad7	Sma Mothers Against Decapentaplegic 7
CTNNB1	Catenin-B1
WHO	World Health Organization
HR	Hazard Ratio
OR	Odds Ratio
W	Wild
H	Heterozygote
M	Mutant
Fzd	Frizzled Domain
Fc	Fragment Crystallizable
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
Cys	Cysteine
Glu	Glutamate
Gln	Glutamine
Gly	Glycine
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
MAF	Minor Allele Frequency
EML4-ALK	Echinoderm microtubule-associated protein-like 4-alk fusion protein

List of Symbols

Symbol	
%	Percentage
<	Less than
≥	Equals to or greater than
°	Degree
°C	Degree celsius
μg	Microgram
μl	Microliter
μm	Micrometer
cm	Centimetre
g	Gram
h	Hour
rpm	Revolution per minute
mg/m ²	Milligram per square meter
l	Litre
M	Molar
χ ²	Chi-square
mg	Milligram
min	Minute
ml	Milliliter
mm	Millimeter
mM	Millimolar
mg	Magnesium
nm	Nanometer
rpm	Revolutions per minute
df	Degree of freedom
s	Second
U	Unit (activity)
v	Volume
Cl	Chloride
D'	Coefficient of linkage disequilibrium
r ²	Extent of linkage disequilibrium
μg /ml	Microgram per millilitre
μM	Micromolar
AgNO ₃	Silver nitrate
HCl	Hydrochloride
ng/μl	Nanograms per microliter
bp	Base pair
α	Alpha
β	Beta
γ	Gamma
±	Plus minus
NaCl	Sodium chloride
ε	Epsilon
A	Adenine
R	Arginine
N	Asparagine
D	Aspartic acid
B	Asparagine or aspartic acid
C	Cytosine
E	Glutamic acid
Q	Glutamine
Z	Glutamine or glutamic acid
G	Guanine
H	Histidine
I	Isoleucine
L	Leucine

K	Lysine
M	Methionine
F	Phenylalanine
P	Proline
S	Serine
T	Thymidine
W	Tryptophan
Y	Tyrosine
V	Valine

CHAPTER 1

INTRODUCTION

Globally, one may have witnessed a massive reduction in overall consumption of cigarette; despite this decreased trend lung cancer continues to remain the reason for major deaths in countries such as United States and also the developing nations such as India. WHO has projected an estimate that by year 2030, lung cancer will account for 12% deaths in India. This unexplained cause of the disease has forced the researchers to understand the genetics behind the same. Glancing at the national lung cancer numbers, as per the Globocan 2012, India was estimated to have close to 70,275 lung cancer cases (inclusive of all ages and sexes) [Globocan, 2012]. The incidence rate for the disease was found to be 5.6 cases per 100,000. Of all the forms of cancer, lung cancer has been ranked as the fourth most common form. With volumes of literature, research and scientific innovations around lung cancer, heterogeneity of lung cancer outlines and demarcates the role of not just one but many parameters governing the incidence of the disease. Lung cancer is steered by the several factors, each of them individually contributing by its percentage, broadly including genetic mutations and dynamic epigenetic alterations. The unmining of these factors, have helped to not only understand the disease, its cause, and prognosis better but has also played an instrumental role in the designing the novel treatment better.

Lung cancer is characterized by complex molecular perturbations which are a result of dysregulated expression of genes involved in the process of cell proliferation, differentiation and apoptosis [Winn *et al.*, 2002]. Beside smoking and other environmental exposures, the genetic, epigenetic alterations and variations in genes encoding proteins involved in signaling pathways also lead to molecular abnormalities which cater to the process of lung carcinogenesis [Yokota *et al.*, 2010]. Thus, proper understanding of multiple signalling pathways which play an important role in the molecular pathogenesis of lung cancer is highly imperative for the development of diagnostic/prognostic markers and also for treatment strategies that can target molecular aberrations and their downstream activated pathways [Surana *et al.*, 2014]. It has been established that tobacco intake is the prime initiator of lung cancer in 80% of the total cases, [Wogan *et al.*, 2004] but recently evidences from familial clustering and segregation studies have revealed that genetic susceptibility might confer an important role in occurrence of lung cancer. A distinct amalgam of polymorphic traits in each individual affects his/her susceptibility towards lung cancer. [Yokota *et al.*, 2010]. Therefore, it is important to understand the role of various cancer related genes including the dominant oncogenes, recessive tumor suppressor genes and their respective polymorphic forms in the process of lung carcinogenesis, so as to find out the populations who would be at a higher risk of acquiring the disease. Additionally, this understanding on the same shall prove to be a tool for the identification of both predictive and prognostic markers for lung carcinogenesis [Amos *et al.*, 1999].

Certain signaling cascades have shown some promising evidences which would help to understand the disease pathology to a better extent. One of these cascades is the Wnt/ β -catenin signalling pathway that has been reported to play an important role in the process of oncogenesis and particularly lung cancer development [Mazieres *et al.*, 2005]. Wnt signalling pathway is involved in various processes such as cell proliferation, survival, polarity, adhesion and development of various tissues by altering the transcription of downstream targets such as *cyclin D1*, *VEGF-A*, *c-myc* etc [Yang, 2012].

Categorically, Wnt pathway is broadly classified into canonical and non-canonical pathway. Focusing, on the canonical pathway which is centrally mediated by β -catenin levels in the cytoplasm, the activation of this pathway is augmented by binding of the Wnt ligands to the frizzled (Fz) receptors and low-density-lipoprotein-related protein5/6 (LRP5/6) complex, therefore these interactions lead to inhibition of phosphorylation of β -catenin by glycogen synthase kinase- β (GSK-3 β) within a large cytoplasmic complex including Dishevelled (Dsh), adenomatous polyposis coli (APC) and Axin [Akiyama, 2000]. This leads to the destruction of this degradation complex and hence the ubiquitination of the β -catenin is prevented and it is translocated to the nucleus where it interacts with the transcription factors T-cell factor/lymphoid enhancer-binding factor (TCF/LEF) which activate target genes such as *matrix metalloproteins*, *cox-2*, *cyclin D1*, *VEGF*, *c-myc* etc. Aberrant activation of Wnt signalling pathway has been related to human cancers such as colorectal, melanoma, gastric, hepatocellular, prostate, ovarian and breast carcinoma [Giles *et al.*, 2003]. Overexpression of Wnt target genes have been confirmed in various tumor tissues [Polakis *et al.*, 2000]. The internalization of the β -catenin to the nucleus is prohibited by the regulation of the Wnt signalling by the Wnt antagonists, which block the signalling cascade by allowing the Axin2 protein to promote the ubiquitination of β -catenin [Luo *et al.*, 2007].

Secreted Frizzled related proteins (sFRP) are extracellular proteins which exhibit a structural homology with the Frizzled- related receptors which is responsible for negatively regulating signal cascade by competing with the Frizzled related receptors for binding to the Wnt ligands which activate this pathway [Shi *et al.*, 2007]. Till date five *sFRPs* (*sFRP 1-5*) have been discovered with 300 amino acid long signal sequence. All five members of sFRP protein family show structural similarity with cysteine rich domain (CRD) ligand binding domain of frizzled family. CRD is one of the most important characteristic of sFRP molecular structure which is characterized by the presence of ten cysteine residues at conserved positions. Presence of sFRP in Wnt signaling pathway may cause morphological alterations by sequestering Wnt through netrin (NTR) and CRD domain. This has been believed to increase cytoplasmic level of β -catenin and inhibit activation of various expression genes. Both CRD and NTR are reported in activating Wnt/ Ca^{2+} signaling pathway by increasing intracellular Ca^{2+} levels This gene is considered to have tumor suppression potential by decreasing cell cycle progression [Longman *et al.*, 2012].

Dickkopf (DKK) proteins are extracellularly secreted glycoproteins and antagonize the Wnt signaling by binding to co-receptors LRP5/6 thereby preventing the formation of active Wnt-Frizzled-LRP5/6 complexes [Zorn, 2001]. DKK family includes four proteins DKK1-DKK4 that are identified by two cysteine rich domains separated by a non-conservative linker region where in the location of 10 cysteine residues is highly conserved within the DKK family and it also possess a N-terminal signal peptide [Krupnik *et al.*, 1999]. The *DKK* genes are located on different chromosomes, namely *DKK2* located at 4q25 and contains four exons whereas *DKK3* is located at 11p15 and comprises of seven exonic regions. *DKK4* is located at 8p11 and has seven exons [Niehrs, 2006]. Recent studies have demonstrated that *DKK3* is downregulated in case of breast, cervical, NSCLC, prostate and renal cancer [Veeck *et al.*, 2012]

Axin2, being a scaffold protein it forms the important part of the destruction complex along with Dsh and APC which mediate the phosphorylation of β -catenin by GSK-3 β required for its ubiquitination which inhibits its translocation into nucleus to bind to transcription factors TCF/LEF regulating downstream target gene expression [Kikuchi, 1999]. Lung cancer being multifactorial in nature, causes cellular abnormality which arises because of certain heterogeneous and unknown genetic and molecular anomalies [Polakis, 2000; Winn *et al.*, 2002] and recently some reports have established the role of this developmental pathway in the oncogenesis of lung tissue [Winn *et al.*, 2002; Mazieres *et al.*, 2005]. Several types of cancers have evidently reported the modifications in Axin2 protein [Huang and He, 2008]. The expression of *Axin2* gene was found to be dysregulated in various human cancers such as colorectal, ovarian, breast and hepatocellular *etc.* [Gong *et al.*, 2015].

Germline variations in the genes of this pathway may lead to its constitutive expression. Chronic expression of this intermeshing transduction cascade has been reported in various human cancers such as colorectal, head and neck cancer, melanoma, leukemia and lung cancer [Polakis, 2000]. Evidently many studies have evaluated the role of genetic differences in affecting disease susceptibility for Wnt antagonists genes namely *sFRP3* [Shanmugam *et al.*, 2007; Berndt *et al.*, 2009; Paez *et al.*, 2014], *sFRP4* [Hirata *et al.*, 2009; Frank *et al.*, 2010; Alanazi *et al.*, 2013; Yilmaz *et al.*, 2015], *DKK2*, [Hirata *et al.*, 2009; Alanazi *et al.*, 2013; Paez *et al.*, 2014; Yilmaz *et al.*, 2015], *DKK3* [Hirata *et al.*, 2009; Alanazi *et al.*, 2013; Paez *et al.*, 2014; Kim *et al.*, 2015; Yilmaz *et al.*, 2015], *DKK4* [Hirata *et al.*, 2009; Alanazi *et al.*, 2013] and *Axin2* [Kanzaki *et al.*, 2006; Gunes *et al.*, 2009; Gunes *et al.*, 2010; Mostowska *et al.*, 2014; Liu *et al.*, 2014; Ma *et al.*, 2014; Wu *et al.*, 2015]. Also, the epigenetic hyper methylation induced silencing of *sFRP1* gene promoter has been established in many cancers such as hepatic [Shih *et al.*, 2006], ovarian [Takada *et al.*, 2004], breast cancer [Veeck *et al.*, 2006]. Also this epigenetic dysregulation of *sFRP1* is reported in NSCLC cancer cell lines [Fukui *et al.*, 2005; Shi *et al.*, 2007; Licchesi *et al.*, 2008; Taguchi *et al.*, 2016] and NSCLC tissue samples [Zhang *et al.*, 2009; Zhang *et al.*, 2011].

Furthermore, the epigenetic hypermethylation of the *DKK3* promoter has been observed in high percentage of lung tumors [Yue *et al.*, 2008]. The expression of *DKK2* gene is reported to decrease in melanoma and gastric cancer [Sato *et al.*, 2007; Kuphal *et al.*, 2006]. There are evidences supporting the oncogenic role of *DKK3* gene in the process of lung cancer progression which is frequently inactivated in lung cancer [Nozaki *et al.*, 2001]. Certain epigenetic alterations have also been seen in renal carcinoma in context of *DKK2* gene [Hirata *et al.*, 2009]. The epigenetic silencing of the genes in this pathway has been observed to alter the level of expression of target genes in various tumors [Tennis *et al.*, 2007]. Methylation of Wnt antagonists, such as *sFRP* has been elucidated in breast [Veeck *et al.*, 2006], hepatic [Shih *et al.* 2006], lung [Tang *et al.*, 2006], and ovarian cancer [Takada *et al.*, 2004]. There are evidences supporting the oncogenic role of *DKK3* gene in process of lung cancer progression which is frequently inactivated in lung cancer [Nozaki *et al.*, 2001].

Interplay of numerous genetic variabilities within the signalling pathways at multiple loci and environmental factors synergistically modulate the propensity of occurrence of a disease. Genetic susceptibility towards a disease has been recently accounted as an important contributing factor in the process of carcinogenesis [Mucci *et al.*, 2001]. Recently, candidate gene based approach has emerged as the new concept hypothesis relating to disease susceptibility. Predisposition towards a disease such as cancer cannot be dramatically effected by a single predictor that might not show promising results when analyzed alone [Neumann *et al.*, 2005]. Several individual genetic variants have given inconsistent outcomes in the various studies pertaining to the contradiction about their involvement in development of the disease. Various factors which bring these differences include smaller sample size, variation in statistical models and lack of evaluation of interaction among the genes involved in same physiological pathway. Therefore, it is plausible that a single genetic polymorphism would not individually determine the risk for cancer [Horne *et al.*, 2005]. Etiology of lung cancer has been elucidated by distinct and complex signalling pathways and their interaction with environmental factors. Being a multifactorial disease, various candidate gene based studies have never given a clear picturesque about its predisposition. Multiple pathways and their interaction with other environmental parameters might unfold the cumulative effect of low penetrance alleles in modulating the susceptibility [Qian *et al.*, 2016]. The complexity of the Wnt signalling cascade drives to analyze the interactive effect of the genetic variation within genes controlling the pathway which is intrinsic and needs to be explored.

Despite of numerous efforts, the average five year survival of lung cancer patients has not improved from 15% [Chen *et al.*, 2014]. The clinical outcomes in case of lung cancer have been very poor as compared to the other cancers such as breast, prostate and colorectal, this difference is attributed to variations in the biology of lung tumors [Borczuk *et al.*, 2003]. Another important aspect of the β -catenin mediated Wnt

signalling is that, it is found to nurture and dictate the differentiation of cancer stem cells that have great tumorigenic capacity. Hyperactivity of β -catenin mediated pathway when controlled by the optimum activity of the Wnt antagonists such as sFRP, DKK protein and Axin2 protein has been found responsible for decreasing the dividing capacity of this stem cell population [Pinto *et al.*, 2003, Korinck *et al.*, 1998]. Exploration of Wnt signalling pathway as a prognostic and predictive marker has been carried out in various malignancies such as hematological [Ashihara *et al.*, 2015], breast [Kazi *et al.*, 2016], ovarian [Bodnar *et al.*, 2014], pancreatic ductal adenocarcinoma [Nakamoto *et al.*, 2014] and esophageal squamous cell carcinoma [Deng *et al.*, 2015]. The involvement of the genetic variations within this developmental pathway has been evaluated in NSCLC in Caucasians [Cosico *et al.*, 2014] and Korean ethnicities [Yoo *et al.*, 2016]. Another study has been done in renal cancer patients where the polymorphic sites of negative regulators of Wnt pathway was examined to find out their predictive role in clinical outcomes [Hirata *et al.*, 2009].

The primary drive behind evaluating the genetic variants of Wnt antagonist genes (*DKK*, *Axin2* and *sFRP*) and their cumulative role lies in the fact that there are no or a very few reports available so far for these SNP's with regard to lung cancer and the above mentioned findings state the they might be important in explaining the aberrant activation of Wnt cascade in lung cancer pathogenesis which is majorly regulated by Wnt antagonists such as *DKK*, *Axin2*, *sFRP3* and *sFRP4*. The current literature remains scanty in terms of evidences supporting the role of these variations in lung carcinogenesis, we thought it would be interesting to screen some of these genetic variability within the Wnt pathway for any evidence to show whether the polymorphic variants within the *DKK*, *Axin2*, *sFRP3* and *sFRP4* genes plays a role in altering the predisposition of North Indians towards lung cancer. As far our knowledge is concerned, this is the first attempt to investigate the role of polymorphic sites of *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK2* (rs447372, rs419558 and rs17037102) and *DKK4* (rs2073664) genes, *Axin2* (148 C/T, 1365 G/A, 432 T/C, 956+16 A/G, 1712+19 G/T, 1386 C/T and 2062 C/T), *sFRP3* (rs7775 and rs228326) and *sFRP4* (rs1802073 and rs1802074). In this study, attempt was made to find the susceptible subgroups within the sample population by stratifying the population on basis of their histology, clinicopathological features and smoking habits. Furthermore, we also tested whether haplotypes of these genes contribute to host susceptibility to lung cancer. The present study was also designed to explore the high order gene-gene and gene-environment interactions that may cater to the difference in lung cancer susceptibility in north Indians. Multifactor dimensionality reduction (MDR) and Classification and regression tree analysis (CART) was used as major tools to find out cumulative effects of all these polymorphic sites and their interaction with environmental factors such as smoking. These single nucleotide polymorphisms (SNP's) have been examined for their association with overall survival of north Indian lung cancer patients. Also, these SNPs were screened to find their ability to act as

predictive and prognostic markers. The epigenetic status of the promoter region of the genes such as *sFRP1*, *DKK2*, *DKK3* and *Axin2* was also investigated in the plasma DNA samples of the north Indian lung cancer patients. The relationship of the methylation status with various clinic-pathological factors has also been evaluated.

CHAPTER 2

REVIEW OF LITERATURE

2.1 Lung cancer

Lung cancer is characterized by a condition of aberrant cell behaviour which leads to the development of expansive masses within the lung tissue. These abnormal neoplastic cells have the peculiar property of intruding into nearby tissues and consequently undergoing the process of metastasis and hence spreading to the other organs of the body [Lung cancer-Genetics Home Reference, 2017]. The human body is equipped with a respiratory system which facilitates the exchange of gases. Lungs which are located in the thoracic cavity, are a complex pair of organs which are interconnected to the nasal passage through a tracheal system. The tracheal system further branches into bronchi and consequently into bronchioles leading to the alveolar sacs where the actual exchange of oxygen and carbon dioxide takes place as illustrated in fig 2.1(a) [Nagaishi and Nagasawa, 1972]. The region between the two lungs is termed as mediastinum. There is specific feature which enhances the immunological potential of the lungs, which is the presence of the lymph nodes in these organs. Also, the lungs are placed in the pleural cavity surrounded by the pleural fluid which is enclosed by the pleural membrane. The abnormal growth in the different regions of the lung tissue leads to the formation of a solid tumor which derives all the blood supply and inherently impairs the proper functioning of the lung tissue by invading the adjacent cells and consequently other organs through lymph node infiltration as illustrated in fig 2.1(b). The lung tissue subsequently suffers from mucous formation, fibrosis and necrosis due to malignant properties of the tumor [Hyde and Hyde, 1974]. The major clinical implications which are evident in the condition of lung carcinoma include:

- Coughing and pain in the chest, shoulder, or back unrelated to pain from coughing
- Shortness of breath
- Changes in the voice or being hoarse
- Harsh sounds with each breath (stridor)
- Conditions such as bronchitis or pneumonia
- Coughing up phlegm or mucus, especially if it is tinged with blood
- Loss of appetite
- Weakness
- Hemoptysis
- Dysphonia and asthenia [Yoder, 2006]

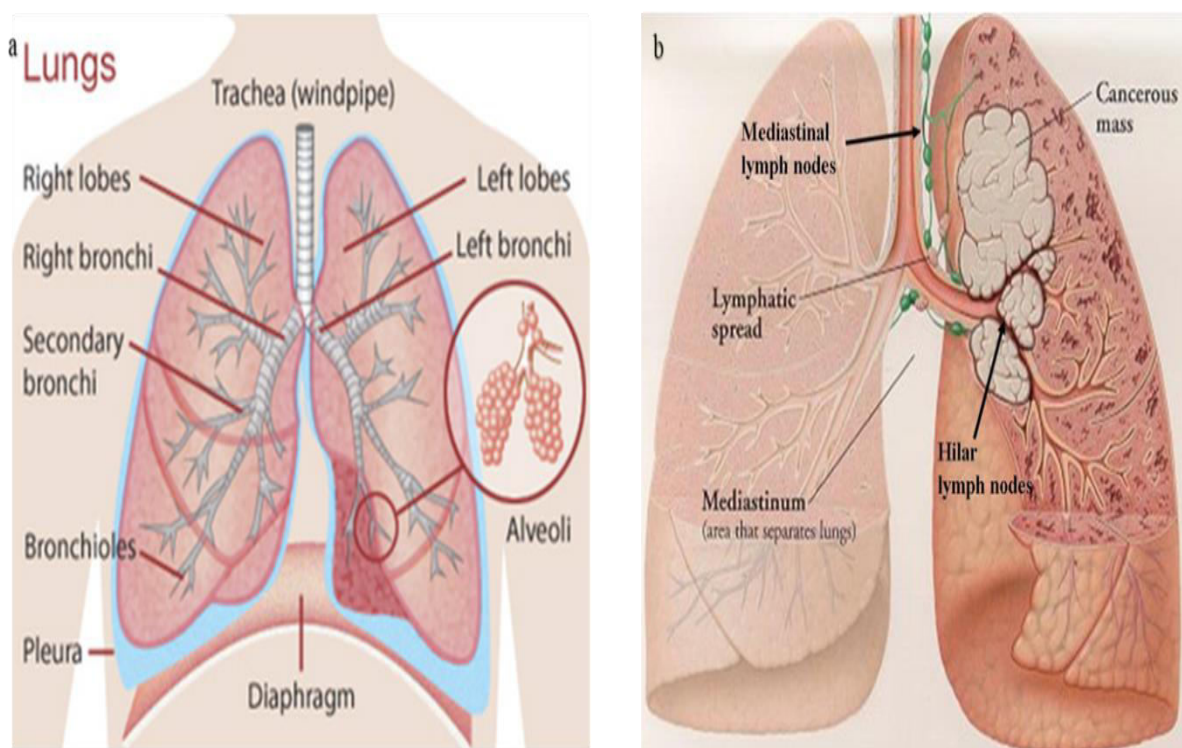


Fig 2.1: Schematic representation of Lung tissue; a: Anatomical features of lung showing trachea, bronchioles, alveoli and diaphragm; b: Lung tissue harboring a solid neoplastic growth along with lymph node infiltration and mediastinum (Source: Gadgeel, 2010).

Staging of lung tumors, these days is done in accordance to the 7th edition of TNM classification of lung cancer tumors given by International Association for study of lung cancer (IASLC). The classification of the lung tumors is done primarily on the basis of tumor size and location, status of the lymph node involvement and metastatic condition. Major proportion of the lung tumors are diagnosed at advanced stages only because of the asymptomatic behavior of early stage lung tumors. Staging is usually done after confirmed diagnosis through fibre optic bronchoscopy, Computed Tomography (CT) guided fine-needle aspiration cytology (FNAC), microscopic identification of tissue by immunohistochemistry and Positron emission tomography (PET) scan. This procedure basically demarcates the candidates who are eligible for surgery, those having loco-regional status, and finally the one's characterized with metastatic disease. The lower stage of the tumor pertains to a better prognosis for the patient [Rami-Porta *et al.*, 2009]. Table 2.1 summarizes the criteria generally used by clinicians for classifying the stages of lung tumors. TNM classification and the parameters help in reckoning the stage of the tumor with the details of the same have been tabulated in table 2.2.

Table 2.1: TNM stage groupings (Source: Weissferdt and Moran, 2012)

Occult carcinoma	TX	N0	M0
Stage 0	Tis	N0	M0
Stage IA	T1a	N0	M0
	T1b	N0	M0
Stage IB	T2a	N0	M0
Stage IIA	T2b	N0	M0
	T1a	N1	M0
	T1b	N1	M0
Stage IIB	T2a	N1	M0
	T2b	N1	M0
	T3	N0	M0
Stage IIIA	T1a	N2	M0
	T1b	N2	M0
	T2a	N2	M0
	T2b	N2	M0
	T3	N1	M0
	T3	N2	M0
	T4	N0	M0
Stage IIIB	T4	N1	M0
	T1a	N3	M0
	T1b	N3	M0
	T2a	N3	M0
	T2b	N3	M0
	T3	N3	M0
	T4	N2	M0
Stage IV	T4	N3	M0
	Any T	Any N	M1a
	Any T	Any N	M1b

Table 2.2: Tumor, node, and metastasis (TNM) definitions (Source: Weissferdt and Moran, 2012)

Primary tumor (T)	
TX	Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy
T0	No evidence of primary tumor
Tis	Carcinoma in situ
T1	Tumor 3 cm or less in greatest dimension, surrounded by lung or visceral pleura, without bronchoscopic evidence of invasion more proximal than the lobar bronchus (i.e., not in the main bronchus) ^a
T1a	Tumor 2 cm or less in greatest dimension
T1b	Tumor more than 2 cm but 3 cm or less in greatest dimension
T2	Tumor more than 3 cm but 7 cm or less or tumor with any of the following features (T2 tumors with these features are classified T2a, if 5 cm or less): involves main bronchus, 2 cm or more distal to the carina; invades visceral pleura (PL1 or PL2); associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the entire lung
T2a	Tumor more than 3 cm but 5 cm or less in greatest dimension
T2b	Tumor more than 5 cm but 7 cm or less in greatest dimension
T3	Tumor more than 7 cm or one that directly invades any of the following: parietal pleural (PL3), chest wall (including superior sulcus tumors), diaphragm, phrenic nerve, mediastinal pleura, parietal pericardium; or tumor in the main bronchus (less than 2 cm distal to the carina ^a but without involvement of the carina); or associated atelectasis or obstructive pneumonitis of the entire lung or separate tumor nodule(s) in the same lobe
T4	Tumor of any size that invades any of the following: mediastinum, heart, great vessels, trachea, recurrent laryngeal nerve, esophagus, vertebral body, carina, separate tumor nodule(s) in a different ipsilateral lobe
Regional lymph nodes (N)	
NX	Regional lymph nodes cannot be assessed
N0	No regional node metastases
N1	Metastases to ipsilateral peri-bronchial and/or ipsilateral hilar node(s) and intrapulmonary nodes, including involvement by direct extension
N2	Metastases to ipsilateral mediastinal and/or subcarinal lymph node(s)
N3	Metastases in contralateral mediastinal, contralateral hilar, ipsilateral or contralateral scalene, or supraclavicular node(s)
Distant metastasis (M)	
M0	No distant metastasis
M1	Distant metastasis
M1a	Separate tumor nodule(s) in a contralateral lobe tumor with pleural nodules or malignant pleural (or pericardial) effusion
M1b	Distant metastasis (in extrathoracic organs)

2.1.1 Epidemiology-World Scenario

In the past two decades, the incidence of lung cancer has increased immensely all around the globe. Annually, 1.61 million new cases of lung cancer were reported worldwide with 1.38 million deaths in 2008 and 1.69 million deaths in 2015 [Major cancer types by deaths worldwide, 2015]. A report by the American cancer society has published certain important facts about the incidence and mortality occurring due to lung cancer. According to the present world scenario, lung cancer is the second most common solid tumor among men and women after prostate cancer in men and breast cancer in women. The number of deaths due to lung cancer estimated by American cancer society in 2017 in United States may reach to 155, 870 cases. Also, the estimated incidence in 2017 in US will rise to 222,500 cases [Key Statistics for Lung Cancer, 2017]. Occurrence of lung cancer shows a distinct geographical pattern among males and females as shown in fig 2.2. Out of the developed and developing countries, mortality rate is higher in developed countries (22%) as compared to 14% in developing countries. Geographically, the maximum proportions of males having lung cancer are reported in Europe, North America and South America [Müller-Hermelink *et al.*, 2004]. Other countries showing a moderate rate of lung cancer incidence include China, Japan and South East Asia. Southern Asia (Pakistan and India) and sub-Saharan Africa have shown the lowest rate of lung cancer incidence among men [Parkin *et al.*, 2002]. This pattern among females mainly showcases the historical scheme of tobacco smoking around the globe. As the rate of lung cancer occurrence can be easily correlated with the changing pattern of smoking habits in females among the developed and developing nations. North America and North West Europe exhibit the highest rate of lung cancer incidence followed by Australia, New Zealand and China [Parkin *et al.*, 2002].

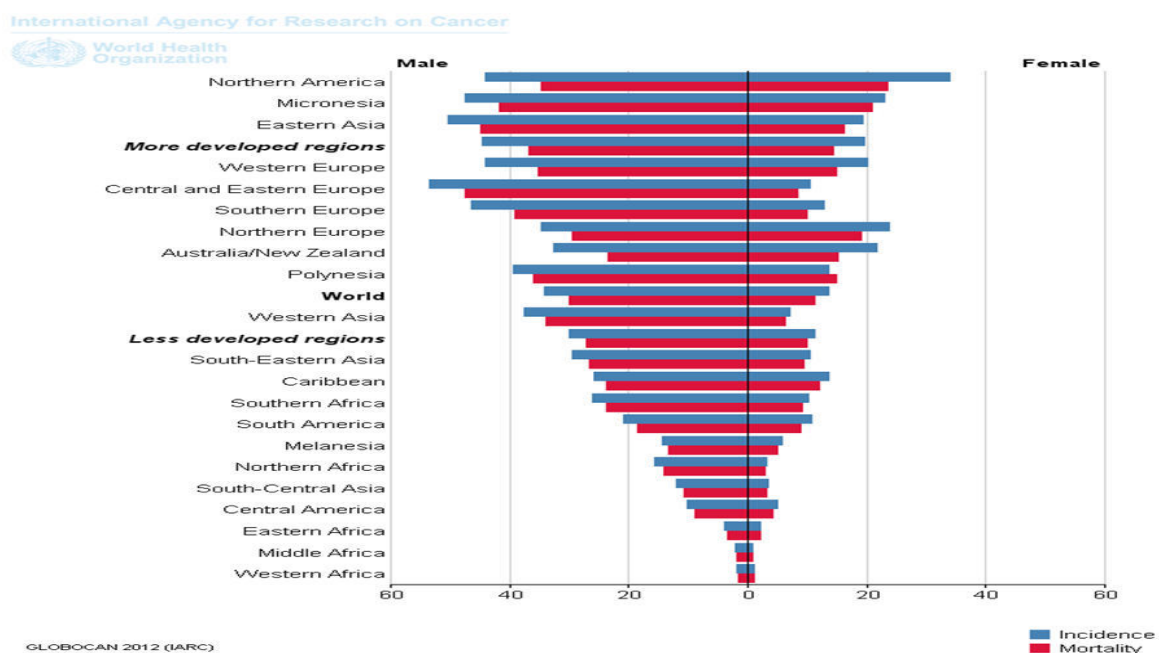


Fig 2.2: Estimated incidence rate (per 100,000) of lung cancer worldwide which are standardised for age. (Source: Globocan, 2012 (IARC), Section of Cancer Surveillance (17/5/2017))

2.1.2 Epidemiology-Indian Scenario

During the early 19th century, the incidence of lung cancer in India was found to be extremely rare. In early 1950's 27.4 per million cases were reported and in 1959 it increased to 78.6 per million [Behera and Balamugesh, 2004]. National Cancer Registry Programme (NCRP) initiated by Indian Council of Medical Research (ICMR) has a major objective of collecting the data from all urban and sub-urban parts of the country. NCRP has made a network of cancer registries all around the country. There are three population based cancer registries (PBCR) in Chennai, Bangalore and Mumbai and three hospital based cancer registries (HBCR) at Chandigarh, Dibrugarh and Thiruvananthapuram which commenced from 1 January 1982. According to the PBCR's fact sheet (2012-2014), the highest age adjusted incidence rate of lung cancer cases in males was found in three areas of Mizoram namely, Aizawl District (37.9), Mizoram State (32.6) and Mizoram excluding Aizawl District (29.3). Similarly, the highest age adjusted incidence rates in females were observed in the three areas of Mizoram namely, Aizawl District (40.8), Mizoram State (29.3) and Mizoram excluding Aizawl District (21.1). Also, lung cancer came out as the most prevalent type of carcinoma in males in 11 registries, the contribution of lung cancer varied from 21.5% in Sikkim PBCR to 48.2% in Manipur state. Similarly, in case of females, the contribution of lung cancer was found to be maximum in 8 registries and the prevalence varied from 24.1% in Chennai PBCR to 72.2% in Manipur PBCR. According to HBCR's annual report, lung cancer contributes up to 8% out of all the tobacco related cancers in all the hospital based cancer registries among males. However, the incidence rate of lung cancer in females is much lower approximately 3% [Three-Year Report of Population Based Cancer Registries 2012-2014, 2016]. According to Globocan 2012 cancer fact sheet, around 53728 cases of lung cancer are estimated to be reported in India per 100, 1000 individuals. WHO has projected an estimate that by year 2030, lung cancer will account for 12% deaths in India. Currently, lung cancer is the fourth largest cause of cancer related mortality in India [Globocan, 2012].

2.1.3 Risk factors for lung cancer

Lung cancer is one of the most frequently occurring solid tumor amongst the tobacco-related malignancies. **Tobacco smoking** is crowned as the major causative agent for neoplasm of the lung tissue [Gandini *et al.*, 2007]. 80% of the lung cancer deaths are attributed to smoking [Ankathil, 2010; Alberg *et al.*, 2013]. Smoke produced by various tobacco products is an amalgam of numerous potentially carcinogenic compounds such as nitrogen oxides, isoprene, buta-diene, benzene, styrene, formaldehyde, acetaldehyde, acrolein, and furan in vapour phase. Certain components of cigarette smoke have been categorised as potent carcinogens for pulmonary tract which include aza-arenes, tobacco-specific nitrosamines, *e.g.* 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK), 1,3-butadiene, ethyl carbonate, ethylene oxide, nickel, chromium, cadmium, polonium-210, arsenic, and hydrazine [Hoffman *et al.*, 2001]. Particulate matter of the smoke has many carcinogenic compounds such as nicotine, polycyclic aromatic hydrocarbons (PAH), benzo (a) pyrene, N-nitrosamines,

aromatic amines, and metals [Pfeifer *et al.*, 2002]. Commonly used form of tobacco products in India are cigarettes, hookah and bidis. Cigarettes generally possess a significantly lower amount (13.3-16.5 mg/g) of nicotine as compared to bidi (21.2 mg/g) [Malson *et al.*, 2001]. Individuals consuming bidi tend to inhale to a greater depth which ultimately leads to an enhanced exposure to carbon monoxide and other damaging constituents. The initiation and progression of bronchial and oesophageal cancers is evidently driven more by bidi smoking. Also, it has been proposed that bidi smokers are more susceptible for developing lung cancer [Kumar *et al.*, 2010]. 85% of lung cancers in males and 47% of lung cancer in females have been attributed solely as result of tobacco consumption worldwide. Major sections of Indian population who are diagnosed with lung malignancy are found to have strong history of bidi smoking. Rural population (31%) of India represent more number of smokers as compared to urban people (21.5%) [Mishra *et al.*, 2016]. Smoking status of any individual is quantified by calculating the “pack years” or the Brinkman index. Pack years are defined as the number of packs of cigarette/bidi a person smoked multiplied by number of years he/she smoked. It gives the cumulative estimate of the smoking dosage [Asami *et al.*, 1997]. Other than cigarette and bidi, tobacco is consumed in the form of hukkah, chillum and chutta in India [Jindal and Behera, 1990]. **Second hand smoke or passive smoking** has been pointed out as one of the risk factors for lung cancer in never smokers. 10 to 15% cases of lung cancer are reported in never smokers. Passive smoking along with other genetic and environmental factors synergistically acts as an important contributor towards lung cancer development in never smokers. There have been various reports addressing the increased susceptibility for lung cancer in individuals who are indulged in involuntary smoking [Samet *et al.*, 2009]. Individuals inhaling second hand smoke have been detected with the presence of 4-amino -byphenyl adduct of haemoglobin formed by the nitrosamines such as NNK present in the involuntary tobacco smoke. This evidence clearly depicts the carcinogenetic potential of passive smoking. Relative risk associated with light smokers was in par with risk associate with individuals who are exposed to second hand smoke of tobacco related products [IARC, 2004].

In addition to the second hand smoke, several **occupational hazards** such as exposure to asbestos, crystalline silica, radon, mixtures of polycyclic aromatic hydrocarbons and heavy metals have been found to contribute in augmenting the carcinoma of lung tissue. Several reports since 1950's have confirmed the relevance of exposure towards these agents in etiology of lung cancer [Boffetta and Trichopoulos, 2002]. A list of such occupational hazards has been enlisted by International Agency for Research on Cancer (IARC) as shown in table 2.3.

Radioactive contaminant such as radon was the first one to be established as an environmental carcinogen for respiratory tract. In United States, the ionizing radiations from radon are the second largest contributor in mortality due to lung cancer [National Research Council, 1988]. According to Environmental Protection Agency (EPA), radon is estimated to cause 21,000 lung cancer deaths annually and 2,900 of them occur in never smokers [Beir, 1999]. This inert gas is produced from

radium during the natural decay series of uranium.

Table 2.3: List of occupational hazards (Source: IARC, 2004)

Agents, mixture, circumstances	Main industry, use
• Arsenic and arsenic compounds • Asbestos • Beryllium and beryllium compounds • Bis(chloromethyl)ether and • Chloromethyl methyl ether • Cadmium and cadmium compounds • Chromium(VI) compounds • Dioxin(TCDD) • Nickel compounds • Plutonium-239 • Radon-222 and its decay products • Silica, crystalline • Talc containing asbestiform fibers • X- and gamma- radiation • Coal-tar pitches • Coal-tars • Soots • Exposure circumstances • Aluminum production • Coal gasification • Coke production • Haematite mining (underground) with exposure to radon • Iron and steel founding • Painter (occupational exposure)	• Glass, metals, pesticides • Insulation, filters, textiles • Aerospace • Chemical intermediate • Dye/pigment • Metal plating, dye/pigment • Chemical industry • Metallurgy, alloy, catalyst • Nuclear • Mining • Stone cutting, minning, glass, paper • Paper, paints • Medical, nuclear • Construction, electrodes • Fuel • Pigments

Alpha-particles which are characterized as high energy and high mass particles are emitted from the inhaled radon progeny that induces the uncontrolled division in lung tissue. Mechanisms underlying the effect of these particles on lung tissue are introduction of DNA base mutations by forming DNA adducts and also breaking the chromosome to affect genomic integrity [Samet, 1989]. Several epidemiological studies have correlated the radon exposure in mine workers with lung cancer development. Meta-analysis study conducted on 2,798 never smoking miners from 11 different cohorts has estimated a three-fold increased risk in never smokers as compared to smokers. Radon is also a common containment of indoor air so, it is solely responsible for lung cancer in never smokers [Samet *et al.*, 2009].

Asbestos is another occupational agent which is known to elevate the relative risk for lung cancer especially in non-smoker workers who are exposed to asbestos in the industry in comparison to unexposed non-smokers. Although, there exists a great variation in the risk associated among the various workers depending on the level of exposure and type of fibres. The contribution of asbestos exposure remains questionable in individuals indulged in tobacco smoking, however the independent carcinogenetic potential of asbestos is evident in never smokers [IARC, 2004]. Certain **silica** based

industries have been found to be hazardous in nature towards risk for lung cancer. Workers suffering from silicosis have shown a higher relative risk for lung cancer in numerous association studies. The risk was found to be positively related to exposure duration. Evidently, silicosis generally indicates an increased risk of lung cancer in non-smokers [Samet *et al.*, 2009]. **Arsenic**, a heavy metal is also a candidate whose exposure modulates the risk for lung cancer in workers employed in copper smelting industry. Increased risk of lung cancer has been observed in non-smoker smelter workers exposed to $\geq 15 \mu\text{g}/\text{m}^3/\text{year}$ of arsenic. However, certain conflicting reports have also been there [Järup and Pershagen, 1991].

Indoor and outdoor air pollutants such as combustion of solid fuel, biomass, cooking fumes, coal and arsenic in drinking water *etc.* have a cumulative role along with tobacco and environmental agents in increasing the susceptibility for this disease [Samet *et al.*, 2009]. Exposure to particulate matters produced by burning solid fuels such as benzo[a]-pyrene, formaldehyde, and benzene enhances the potential of this indoor smoke to interact with the bronchial tract as they have an inherent capacity to form DNA adducts. The profusely increasing incidence of lung cancer in never smokers and non-smokers females can be attributed to all the above mentioned factors [Loprieno, 1975].

Along with these external factors, other epidemiological factors affecting lung cancer incidence are **age, gender, ethnicity, familial history and geographical variability** [Samet *et al.*, 2009]. Lung cancer incidence is greatly reflected by the age of the population, rate of lung cancer development is very rare in subjects under 40 years of age. This pattern has been similar in case of smokers and never smokers among different ethnic groups such as Caucasians, Japanese and African Americans [Thun *et al.*, 2006]. Gender related biasness in lung cancer occurrence is evident all around the world. Females above the age of 60 years who never smoke are at much higher risk of developing the disease as compared to males who don't smoke [Thun *et al.*, 2006]. However, cohort studies of prospective type have shown an increased mortality rate for men as compared to women [Samet *et al.*, 2009]. Ethnicity has a major impact on the lung cancer risk in never smokers; it is highest in case of African Americans [Thun *et al.*, 2006] and Asians followed by Caucasians [Subramanian and Govindan, 2007].

2.1.4 Histopathological classification of lung cancer

Lung cancer, being a highly heterogeneous disease is classified on the basis of the histopathological characteristics of lung tumor cells. Broadly, there exist two different types of lung cancer namely Small cell lung carcinoma (SCLC) and Non-small cell lung carcinoma (NSCLC). Further, NSCLC is subdivided into many different types majorly comprising of Adenocarcinoma (ADCC), Squamous cell carcinoma (SQCC) and large cell carcinoma (LCC). All the subtypes differ typically in regard of the clinical and biological features. Approximately, 80% of lung cancer patients are diagnosed with NSCLC and only the remaining 20% are diagnosed with SCLC. Both the types of tumors can be

easily identified on the basis of their distinct microscopic morphological characteristics [Lung cancer, 2017].

Small cell lung carcinoma, the most frequent pulmonary neuroendocrine neoplasm is defined as the tumor of small epithelial lung cells having scanty cytoplasm, irregular cellular borders, granular chromatin and lack of conspicuous nucleoli. Cells have peculiar round and spindle shape showing extensive necrosis and are rapidly dividing. It is the most aggressive and robust type of malignancy which has the highest proliferation rate. SCLC specimen samples are identified by certain cytological features such as neoplastic cells that are arranged in loose irregular clusters, with high nuclear/cytoplasmic ratio as illustrated in fig 2.3(a). Immuno-histochemical determinants of SCLC cells include neuroendocrine markers such as CD56, chromogranin and synaptophysin along with TTF-1 in 90% of the cases [Nicholson *et al.*, 2002]. The invasiveness of this solid tumor is impeccable as 60-70% cases are diagnosed with metastasis and prognosis is very poor. Also, there is lack of any prevalent prognostic and predictive biomarker. 95% of the total cases of SCLC are attributed to tobacco consumption. Generally, the patients with age of more than 60 years are detected with SCLC. The chances of relapse in case of SCLC patients are maximum. Clinically, the general symptoms observed at the time of diagnosis include weight loss, fatigue, lack of breath (anorexia), cough, chest pain, dyspnea and hemoptysis. Paralysis of vocal cords is another signature clinical implication of SCLC. In case of advanced and metastatic stage of SCLC, certain specific indications are seen in extra pulmonary tissues such as jaundice in the liver, seizures in the brain, pruritus and pain in bones. During brain metastasis, a change in mental status and ataxia are also manifested [Carter *et al.*, 2014].

Squamous cell carcinoma (SQCC) is a type of NSCLC occurring in epithelial cells which is characterized by keratinization and intercellular connections formed between bronchial epithelial cells. 90% of the SQCC cases are attributed to cigarette smokers. Another important occupational agent which is considered as the second most important causative agent for SQCC is arsenic [Spiro and Porter, 2002]. The prominent cytological features include large size tumor cells bearing hyperchromatic nuclei and excess of cytoplasm. The cells are either spindle or tadpole shaped. They also exhibit cohesive aggregates and the color of the tumor cells vary from white to grey, reflecting the fibrosis of the lung tissue [Funai *et al.*, 2003]. SQCC cells show the presence of keratin (34BE12), cyto-keratins 5/6, and carcinoembryonic antigen (CEA) when analyzed by immune-histochemical properties as shown in fig 2.3(b) [Chu and Weiss, 2002]. This type of tumor is generally located in the central part of bronchial system from where it invades the wall of bronchus to infiltrate into nearby tissues leading to the cessation of release of secretions in the bronchial tract and consequently condition of atelectasis, bronchial dilatation, obstructive lipoid pneumonia and infective bronchopneumonia may arise [Funai *et al.*, 2003]. SQCC staging is done in accordance to the TNM

system [Rami-Porta *et al.*, 2009]. As compared to other histological subtypes, SQCC exhibits lesser frequency of metastasis [Quint *et al.*, 1996].

Adenocarcinoma (ADCC) is a type of NSCLC where in the epithelial cells undergo glandular differentiation and produce mucin to form different growth patterns comprising acinar, papillary and bronchioalveolar. It is the most frequently occurring subtype among individuals who never smoke or have minimum exposure towards smoke [Khuder, 2001]. This type of tumor is generally located around the peripheral nodules rather than the central hilar or perihilar nodes. 15% of the cases show the involvement of pleural membrane and chest wall which is highest among all the other histological subtypes [Kondo *et al.*, 2002]. Cytological identification of the adenocarcinoma cells depends on both the morphology of the cells and their architecture of cell clusters. The cells are arranged in sheets or a three dimensional morulae, acini and papillae along with fibrovascular cores. Cytoplasm of the cells is generally varying in amount but usually it is in abundance and is relatively more translucent than squamous cell carcinoma. Another typical feature is that the cellular borders of these cells are delineated. Single, eccentric and oval shaped nuclei are visible in these cells. Prominent, single macronucleoli which are round or irregular are evident as illustrated in fig 2.3(c). The invasiveness of this tumor is visualized by fibrosis, along with necrosis, cavitation and hemorrhagic area [Walter and Pryce, 1955]. Typical immuno-histochemical characteristics of these cells are elucidated by the presence of epithelial markers such as AE1/AE3, CAM 5.2, epithelial membrane antigen (EMA) and carcinoembryonic antigen (CEA) [Colby, 1995]. Differentiation of primary ADCC from metastatic ADCC can be done on the basis of immuno-phenotype for cytokeratin (CK) 7 and CK20. Characteristic immuno-phenotype of pulmonary adenocarcinoma tumor is that it stains positive for CK7 and negative for CK20 [Rubin *et al.*, 2001]. Adenocarcinoma is distinguished into different grades on the basis of extent of variation of the architectural frame of tumor cells from the normal lung tissue. The three grades include grade 1 as well differentiated, grade 2 as moderately and grade 3 as poorly differentiated [Travis and Harris, 2004].

Large cell carcinoma is an undifferentiated type of NSCLC which does not exhibits the cytological and architectural features of small cell carcinoma and glandular or squamous carcinoma. Large cell carcinoma accounts for almost 9% of the total cases in most of the studies [Travis and Harris, 2004]. This type of tumor is not characterized by any specific cytological features. The borders of the cells are not defined and hence the no specific pattern is observed in the syncytium formation. The shape of the nucleus is also not defined as it varies from being round to irregular. Chromatin distribution is also haphazard however, a prominent nucleolus is evident. The cytoplasm is not in abundance, is basophilic in nature along with an increased nuclear to cytoplasm (N/C) ratio as shown in fig 2.3(d). Some of these features are identical to neuroendocrine tumors except the presence of defined nucleoli. These tumors are identified in chest radiographs as large, peripheral masses which might invade some segment or large bronchi. Invasiveness of this type of tumor is evident in viscera, pleural wall and

nearby areas [Silverman and Geisinger, 1996]. Significant symptoms of this tumor include necrosis, haemorrhage and cavitation. Certain immune histochemical markers which are used to differentiate this type of tumor from others include chromogranin, synaptophysin and N-CAM (CD56) [Lantz *et al.*, 1989]. Molecular and genetic perturbations are almost similar to those which are redundant in other NSCLC types as it is a poorly differentiated form of that tumor. These alterations include *K-ras* mutations, *p53* mutations and *Rb* signalling dysregulation and overexpression of *cyclin D1* [Travis and Harris, 2004]. The prognostic and predictive response for one type of large cell carcinoma which is large neuro-endocrine carcinoma almost similar to SCLC [Takei *et al.*, 2002].

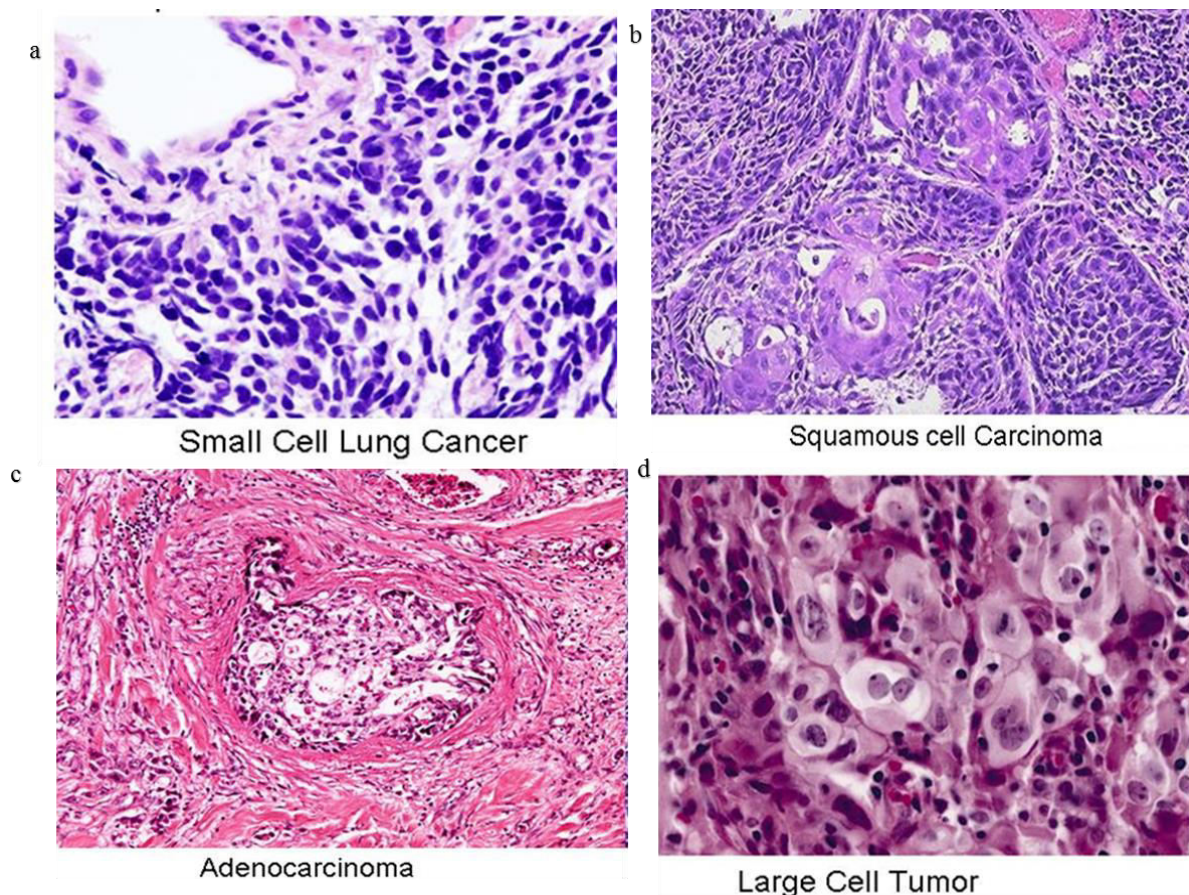


Fig 2.3: Microscopic images of different histological subtypes of lung cancer; a: small cell lung cancer; b: squamous cell carcinoma; c: adenocarcinoma; d: large cell carcinoma (Source: Weiss 2010).

2.1.5 Molecular pathogenesis of lung cancer

Accumulation of diverse alterations at both genetic and molecular level has been ensued to contribute in the pathogenicity of lung carcinoma [Travis and Harris, 2004]. The most important property acquired by the cells undergoing the path of dysplasia, hyperplasia and malignant neoplasia is the capability to proliferate without responding to any regulatory factors, evade the cellular apoptotic pathway, overexpress the factors responsible for vascular angiogenesis, undergo unregulated

replication by hyperactivity of telomerase and finally the ability to invade the adjacent tissues and distant sites [Hanahan and Weinberg, 2000]. All these functional disparities are attributed to differential expression of the genes and proteins controlling multiple processes of homeostasis. Another important characteristic is that these heterogeneous lung tissue cells arise from the precursor cancer stem cells from different lineages which dedifferentiate to form these aggressively dividing cells. This attributes to differences in pathogenesis of different histological subtypes of lung cancer [Nuciforo and Fraggetta, 2004]. Recently, the epigenetic changes have been elucidated to play an important role in dysregulating the expression of the genes controlling cellular proliferation and differentiation. Promoter methylation can directly or indirectly keep a check on the genome integrity and hence methylation profile of all these target genes can be of great relevance in disclosing the pathogenicity of the lung tissue [Zöchbauer-Müller *et al.*, 2002]. Different types of molecular changes that take place in the cells during the event of carcinogenesis include dysregulation of tumor suppressor genes, activation of proto-oncogenes under the effect of any environmental carcinogen, loss of heterozygosity and genomic instability arising due to loss or gain of whole chromosomes or fragments [Massion and Carbone, 2003]. Several studies have investigated the specific genetic alterations taking place in the course of lung cancer development and progression. There have been a few distinctions among the molecular aberrations that are identified among SCLC and NSCLC [Fong *et al.*, 2003].

Some of the common genetic determinants which are evident in generally all the histological subtypes of lung cancer includes allelic loss of several chromosomal loci containing proto-oncogenes such as 3p14-23, 8q21-23, 9p21, 13q, 17q, 18q and 22p [Panov, 2005]. The following genetic aberrations are commonly seen in different lung cancer tumors are explained in detail:

Activation of proto-oncogenes: The uncontrolled proliferative capacity of the lung tumor cells can be attributed to this class of growth stimulating genes. One such example is **protein tyrosine kinases**, which form the major regulatory force for cellular pathways that controls the process of cell division and differentiation. They are also influential in governing cell to cell communication. Hence, by doing so they provide a conducive microenvironment by forming paracrine and autocrine cycles of growth stimulating factors for cancer cells to divide [Fong *et al.*, 2003]. Enhanced expression of cyclin D1 [Keum *et al.*, 1999], cyclin B [Fukuse *et al.*, 2000] and cyclin E [Soria *et al.*, 2000] has been seen in several NSCLC tissues. Another similar type of molecular modulator called **EGFR (Epidermal growth factor receptor)** is a type of tyrosine kinase which augments a ligand activated cascade of auto-phosphorylation of mediators inside the cells to finally activate the effector molecules responsible for important events such as cell division, apoptosis, motility and neovascularization. *EGFR* is located at chromosome 7p11.2, covering around 200kb and has 28 exons. EGFR protein acts through two different intermediate pathways, PIK3CA/AKT1/mTOR pathway and the RAS/RAF1/MAP2K1/MAPK1 pathway. The downstream targets of the EGFR pathway recruited

after hetero or homo dimerization includes p21-Ras and mitogen-activated protein kinases (MAPK's) [Danesi *et al.*, 2003]. Genetic aberrations in *EGFR* gene have been reported in 40-50% of advanced NSCLC cases particularly ADCC. Moreover, these mutations are more frequently reported among non-smokers [Kitamura *et al.*, 2008]. Commonly identified mutations in ADCC cases include three major types located in exon 18-21 of *EGFR* gene. The first are the in-frame deletions found in exon19, second are insertional mutations found around exon 20 and last are missense substitutions in the region between exon 18-21. Most of these genetic changes alter the ATP binding cleft of the EGFR protein, and hence cause the uncontrolled proliferative effect. EGFR inhibitors, called the tyrosine kinase inhibitors (TKI's) have gained immense clinical importance as they competitively fight with ATP to bind to the active site and in doing so, they block the active cleft and the cascade is blocked [Pao *et al.*, 2005]. These genetic alterations are cohesively associated with the prognosis and resistance to platinum based chemotherapy. The frequently employed tyrosine kinase inhibitors include erlotinib, gefitinib and afatinib which show 70% objective response [Wu *et al.*, 2012].

Other important proto- oncogenes which have been implicated in lung cancer development include **Her-2/neu and MYC proto-oncogene**. Her-2/neu is a transmembrane glycoprotein of 185kDa which shares a homology with EGFR and belongs to the ERBB family of proteins. Several studies have revealed the overexpression of this protein in around 1-4% of the NSCLC cases primarily adenocarcinoma [Fong *et al.*, 2003]. MYC proto-oncogenes belong to family of transcription factors which regulate the expression of genes controlling cellular proliferation and apoptosis. These phosphoproteins are located in the nucleus. They work on the cascade activation by forming heterodimers with the other proteins controlling cellular differentiation. It also interacts with the histone deacetylase which directly governs the chromatin conformation responsible for affecting gene transcription [Forgacs *et al.*, 2001]. Commonly seen aberration in the c-myc gene is overexpression and gene amplification. This is evident in around 80-90% of the SCLC cases [Rom *et al.*, 2000].

Another important proto-oncogene which is found to be dysregulated during lung carcinogenesis is **RAS** gene. RAS family of proteins has various members namely, *H-Ras* gene (homologous to the oncogene of the Harvey rat sarcoma virus), the *K-Ras2* gene (homologous to the oncogene of the Kirsten rat sarcoma virus) and the *N-Ras* gene (initially cloned from human neuroblastoma cells). These proteins are found anchored on the inner side of the cell membrane and they counteract by activating the GTPase enzyme to cleave a phosphate moiety from GTP to reduce it to GDP which activates a cascade in which Ras- dependent kinases acts on MAPK's and it moves to the nucleus to recruit the transcription factors. Constitutive expression of these proteins leads to the uninterrupted growth of cells [Danesi *et al.*, 2003]. *K-Ras* is one of the oncogenic driver mutation which is certainly reported in 25% of the smokers exhibiting the invasive mucinous adenocarcinoma or large cell carcinoma [Calles *et al.*, 2015]. The most frequent type of genetic alteration observed in this gene is the missense substitution in the coding region mainly codon 12, 13 and 61. These substitutions are

generally categorized as G to T trans-versions which are result of DNA damage induced by the benzo[a]pyrene diethyloxide (BPDE), nitrosamines *etc.* present in the tobacco smoke [Graziano *et al.*, 1999]. One of the commonly reported fusion mutation comprising the 5' end of the Echinoderm microtubule-associated protein like-4 (EML 4) and 3' end of the anaplastic lymphoma kinase (ALK) is highly significant chromosomal rearrangement. This EML4-ALK chimeric protein is found to play an oncogenic role in as its N-terminal has been derived from EML4 and the C-terminal contains the whole tyrosine kinase domain of *ALK* gene. This fusion mutation is more likely seen in non-smokers. This is observed in nearly 5% of the NSCLC cases [Soda *et al.*, 2007].

Neoplastic SCLC cells have a distinct feature of secreting certain important peptides such as angiotensin, bombesin, insulin-like growth factor 1, vasopressin and serotonin providing a property of stimulating themselves to aggressively proliferate [Sethi *et al.*, 1992].

Inactivation of tumor suppressor genes: Tumor suppressor genes regulate cellular proliferation, DNA damage response and further repair mechanism. The loss of activity of anti-proliferative genes is frequently evident in lung tumorigenesis. Dysfunctional tumor suppressor genes are attributed to loss of allele at certain chromosomal loci, genetic alteration or epigenetic changes including the hyper methylation of the promoter regions of these genes [Wistuba *et al.*, 2000]. The most redundant mutation found in almost all the histological subtypes is the alterations in *p53* gene, which plays the most important role in maintaining the integrity of the human genome. It has multidimensional functions to perform for which it has several specific features such as DNA binding proficiency and certain specific sequences which control the expression of genes which are active growth arrest phase of the cells under genotoxic condition that might arise due to hypoxia, DNA damage and switching on of oncogenes. It interacts with the MDM2 protein which is largely responsible for regulating cell division [Oren, 2003]. The common mutation which leads to the malfunctioning of *p53* protein include missense mutations (mainly G→T trans-versions) clustered in the middle of the gene at codons 157, 245, 248, and 273. TP53 is found to be mutated in almost 50% of NSCLC cases and 75% of SCLC. The genetic alterations which lead to loss of tumor suppressor activity are generally related to cigarette smoking [Sholl, 2016].

The next tumor suppressor gene which has been found to be implicated in lung cancer pathogenesis is **Retinoblastoma (Rb) protein**. The phosphorylation status of the RB protein and its interaction with transcription factor E2F is most important for the regulation of G0/G1 cell cycle transition. Alterations of the *Rb* gene in lung cancer cover deletions, nonsense mutations, pathogenic splicing variations and chromosomal deletions [Forgacs *et al.*, 2001]. Almost 90% of the SCLC and 15-30% of the NSCLC neoplasms have dysregulated or no *Rb* expression [Rom *et al.*, 2000].

Other tumor suppressor genes, which are accounted to play a part in augmenting lung cancer initiation and progression include ***Pten*** (Phosphatase Tensin Homolog Deleted on Chromosome Ten) gene is located at chromosome 10q23 and codes for a lipid phosphatase which removes the phosphate moiety

from PIP3 and carries a anti proliferative effect [Forgacs *et al.*, 2001] and **Transforming growth factor** (TGF- β) which has multiple functions as a protein where in, it stops the proliferation of many epithelial cells through binding with a set of cell receptors. Lower expression of this protein has been evident in NSCLC tissues when analyzed by immune histochemical staining [Panov, 2005].

Escaping the programmed cell death: Apoptosis or programmed cell death has been the most important metabolic tool for controlling the repercussions that arise from the exposure towards carcinogens, DNA damage, and necrotic cells. The inherent mechanism of the cells to undergo the apoptotic pathway maintains the homeostasis of the human body. These cellular pathways are dysregulated whenever a tumor formation is initiated [Panov, 2005]. *Bcl-2 (B-cell lymphoma-2)* gene was the first oncogene found to function through the production of an inhibitor of apoptosis. The *bcl-2* gene family consists of more than 15 members, which either promote or inhibit the apoptosis [Reed, 1997]. Bcl-2 overexpression, detected by immunohistochemistry was found in 75%-95% of SCLC tumors, 25%-30% of the squamous cell carcinomas and in 10% of adenocarcinomas [Kaiser *et al.*, 1996].

The pathogenesis of lung tissue can be demonstrated through fig 2.4 which is assembled on the basis of earlier reports. Removal of 3p21.3 allele is triggered in many epithelial cells of respiratory system at distinct sites due to the exposure of tobacco and other pollutants. Further, chromosomal loss of tumor suppressor gene located at 3p21.3 takes place which makes it haplo-insufficient. The next set of event is characterized by inactivation of growth regulating genes such as *Rb*, *p53* or *p16* due to mutation or epigenetic silencing. This subsequently allows the clonal proliferation of cells which were initially transformed. It is proposed that during the pathogenesis of the SCLC neoplastic cells arise directly either from normal or hyperplastic epithelial cells without passing through characteristic preneoplastic intermediate pathological stages (parallel theory of lung cancer pathogenesis). On the contrary, the NSCLC pathogenesis is accompanied with sequential morphological changes (sequential theory) [Panov, 2005].

Some of the oncogenic driver mutations have gained clinical importance and are recently been employed in therapeutic regimen of the ADCC patients along with the ongoing chemotherapy which largely comprises of the platinum based drugs. These potential targets have been of great importance in increasing the average survival time of lung cancer patients, particularly ADCC patients [Sholl, 2016]. Some of these driver mutations along with the inhibitors have been tabulated in table 2.4.

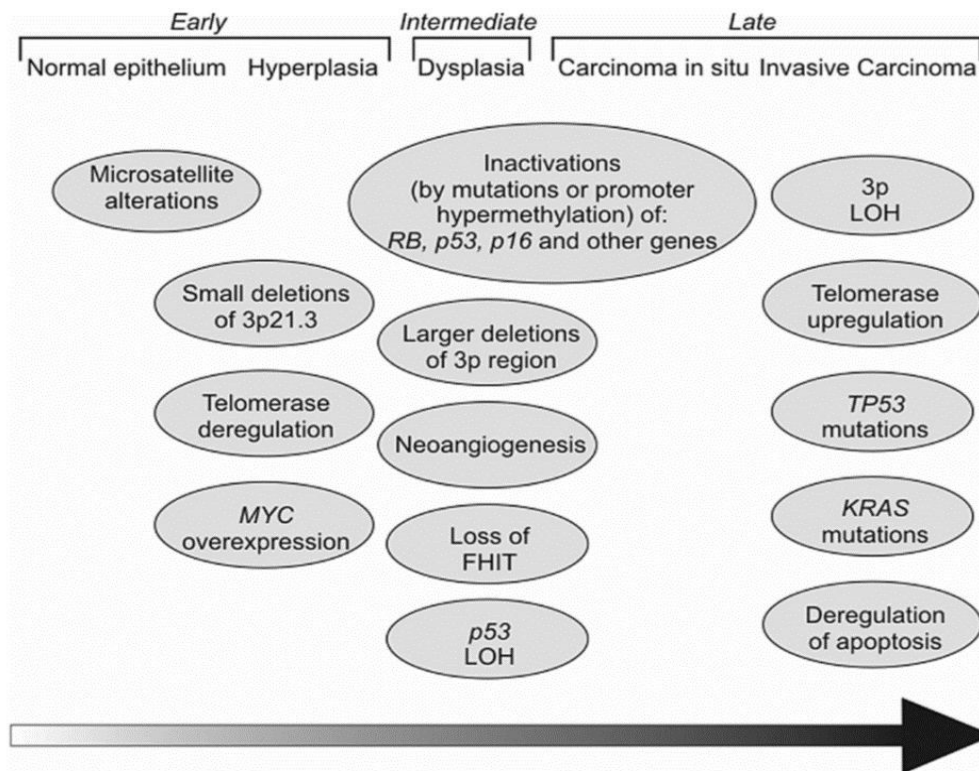


Fig 2.4: Schematic representation of major molecular abnormalities taking place during lung cancer pathogenesis. (Source: Panov, 2005)

2.1.6 Diagnosis for lung cancer

Lung cancer is generally symptomatic at the time of the initial evaluation. Primary diagnostic procedure includes assessment of the history of the patients along with the preliminary physical examination. These physical tests includes the complete blood count, spirometry, estimation of alkaline phosphatase and hepatic transaminase, calcium levels and determination of physiological parameters such as electrolytes, creatinine, and urea concentration in blood and urine and radiography of chest. After the initial screening, in order to confirm the presence of the tumor which might be lying beneath the mediastinum or at any other place within the lungs a contrast enhanced computed tomography (CT) followed by positron emission tomography (PET) to find out the stage of the tumor on the basis of the 7th revised edition of IASLC and exact status of the spread of the tumor around the lung tissue and other organs. The next step involves the confirmation of histology of the lung tumor using molecular identification techniques such as immuno-histochemical characterization. In order to carry out this molecular testing, a section of the lung tumor tissue is required. The approach used to obtain the tissue largely depends on the location and size of the tumor and its metastatic accessibilities. For the centrally located lesions, fibre optic bronchoscopy is used, whereas for lesions located near the periphery CT-guided transthoracic needle aspiration is a typical method. Endobronchial ultrasound and electromagnetic navigation are the techniques of the new diagnostics [Latimer and Mott, 2015].

Table 2.4: Oncogenic Driver mutations commonly found in Adenocarcinoma (Source: Sholl, 2016)

S.No.	Gene	Type of Mutation	Frequency of occurring	Target population	Targeted Therapy
1.	<i>KRAS</i>	Missense mutations at codons 12, 13, 61	25%	Smokers as compared to non-smokers	Chemotherapy, MEK inhibitors under clinical trail
2.	<i>EGFR</i>	Missense and insertion-deletion mutations at hotspots in exons 18–21; common ones include Ex19del, L858R Mutations	15%	Non-smokers as compared to smokers; Women are effected more than men ; widely found in Asian ethnicty	Tyrosine kinase inhibitors: Erlotinib, Gefitinib, Afatanib
3.	<i>ALK</i>	Rearrangements such as <i>EML4-ALK</i> are the most common	5%	Non-smokers as compared to smokers; Women are effected more than men ; younger patients	Multikinase Tyrosine kinase inhibitors: Crizotinib, Ceritinib, Alectinib
4.	<i>ROS1</i>	Rearrangements: CD74-ROS1, SCL34A2-ROS1, EZR-ROS1, FIG1-ROS1	1%–2%	Non-smokers as compared to smokers; Women are effected more than men ; younger patients	Multikinase Tyrosine kinase inhibitors: Crizotinib
5.	<i>RET</i>	Rearrangements: KIF5B-RET, CCDC6-RET, NCOA4-RET	1%–2%	Non-smokers as compared to smokers; Women are effected more than men ; younger patients	Multikinase Tyrosine kinase inhibitors: Sunitinib, Cabozantinib, Vandetanib
6.	<i>BRAF</i>	V600E and other missense mutations at hotspots in exons 11 and 15	4%	Smokers as compared to non-smokers	RAF-MEK inhibitor trials Dabrafenib, trametinib
7.	<i>ERBB2</i>	Small exon 20 insertions Rarely may have concomitant Amplification	2%	Non-smokers as compared to smokers;	HER inhibitor trials: Trastuzumab, Neratinib, Afatinib, Lapatinib
8.	<i>MET</i>	Exon 14 skipping mutations Amplification	3%	Smokers as compared to non-smokers; older patients	Multikinase Tyrosine kinase inhibitors: Crizotinib

2.1.7 Treatment for lung cancer

Cancers that are limited to a small area and have not spread very far are best treated with a local treatment to remove or kill the entire tumor. Surgery and radiation are forms of local treatment. Surgery generally is more effective than radiation in removing the tumor but not all the patients can safely tolerate surgery. Sometimes, chemotherapy is given before or after local treatments depending upon the spread of the cancer cells to other organs. Chemotherapy after surgery is called adjuvant therapy. It is given to help kill any remaining cancer cells not removed by surgery. When chemotherapy is given before surgery it is called neoadjuvant therapy. Chemotherapy can also be given as the only treatment. In non-small cell lung cancer, chemotherapy commonly consists of a platinum drug (either cisplatin or carboplatin) along with a second drug (paclitaxel, docetaxel, gemcitabine or vinorelbine). After first-line chemotherapy, a few drugs are also approved for second-line therapy of non-small cell lung cancer such as pemetrexed. For small cell lung cancer patients, most commonly used regimen is cisplatin/irinotecan or carboplatin/irinotecan. Radiation therapy to the chest is routinely used in patient with limited-stage disease, *i.e.*, tumor confined to the chest. Radiation therapy to the brain is commonly used in limited-stage patients, in an effort to prevent brain metastasis [Behera, 2007].

Targeted therapy refers to medicines, unlike traditional chemotherapy, that target a specific mechanism or pathway in the cancer cells to slow or stop the growth of the cancer cells. Currently, targeted therapies are only available for a few types of advanced stage NSCLC. The first approved drug was gefitinib, now largely supplanted by erlotinib. They work by inhibiting the epidermal growth factor tyrosine kinase receptor (*EGFR* inhibitors) [Cabanero *et al.*, 2017]. Cetuximab is a monoclonal antibody directed against *EGFR*, currently in trials in lung cancer. *VEGF* (Vascular Endothelial Growth Factor) inhibitors affect the new blood vessel formation in tumors. Bevacizumab is FDA-approved for the treatment of advanced colon cancer, and is being evaluated in a variety of other cancer settings. It was just approved for the treatment of advanced non-small cell lung cancer, in conjunction with carboplatin/paclitaxel, showing a two month increase in average survival over chemotherapy alone [Wu *et al.*, 2012].

Another upcoming therapy is immunotherapy where immune checkpoint inhibitors such as Nivolumab and Pembrolizumab. Immunotherapy is now being approved as a first line therapy. Checkpoint immunotherapy uses highly selective humanized monoclonal antibodies against checkpoint signals such as programmed cell death receptor and programmed cell death ligand. By blocking these receptors and signals, the immune system can be reactivated to fight the tumor. Immunotherapy for advanced NSCLC has resulted in a new paradigm of treatment options resulting in improved survival and response rates and has a less severe yet unique toxicity profile when

compared to chemotherapy. Programmed cell death ligand inhibitors, nivolumab and pembrolizumab, and programmed death ligand receptor inhibitor, atezolizumab, are currently approved by regulatory authorities for the treatment of advanced NSCLC [Shine and Sehgal, 2018]. There are several biomarkers which are predictive markers for estimating the likelihood of response of the anti PD-L1 inhibitors in NSCLC patients. Checkpoint inhibition (i.e., antibodies targeting against pathways involving adaptive immune suppression such as PD-L1, PD-1, and CTLA-4) is a uniquely encouraging anti-tumor strategy that appears to be clinically pertinent across a number of tumour types. Furthermore, the extended clinical response observed through the generation of memory T cell responses can provide long-term immunity. In addition, on the basis of the enhanced knowledge of the biological mechanisms fundamental for cancer immune interactions, a wave of novel immunotherapeutic procedures that target activating and inhibitory T cell receptors, adoptive cell therapies (ACT) comprising tissue-infiltrating lymphocytes (TILs), chimeric antigen receptors (CARs), T cell receptor (TCR)-modified T cells, and specific antibodies are now clinically assessed, and two chimeric antigen receptor T cell (CAR-T) products have received recent approval in the United States for treating patients with refractory hematologic malignancies (specifically acute lymphoblastic leukemia and diffuse large B cell lymphoma). These biomarkers include the PD-L1 expression, Microsatellite instability and tumor mutation burden. For Pembrolizumab indication, PD-L1 expression (Tumor Proportion Score i.e TPS >1% is a criteria in ADCC NSCLC patients. Similarly, another milestone is inclusion of Tumor mutation burden in guidelines as an emerging marker due to Checkmate 227 trial which now indicates that 10 mutation per megabase can show better response to nivolumab or ipilimumab and nivolumab [Prelaj *et al.*, 2019].

2.2 Wnt Pathway

2.2.1 History of Wnt pathway

Metazoan body formation has evolved during the course of evolution where in a high degree of cellular differentiation and diversity along with a typical spatial positioning is the key in obtaining the required cellular symmetry and polarization [Loh *et al.*, 2016]. The major force behind the orchestration of the various processes taking place during embryonic development is an evolutionary conserved signalling cascade known as the Wnt pathway. This pathway is a myriad of all the different biological processes including cellular migration, cellular polarity, cell division, patterning of the neural network and morphogenesis which are witnessed during the development of an organism. Wnt cascade has been found responsible for maintaining the homeostasis of all the animals. Various interactions taking place between membrane receptors and Wnt proteins govern various intracellular responses. Wnt signalling is the epitome of embryonic development, tissue differentiation and regeneration in adults. It also controls numerous physiological processes due to its complex and intermeshing property [Yamaguchi, 2001].

For the first time, the Wnt pathway and its implication in oncogenesis came under the eye of the researchers in the year 1982 when Roel Nusse and Harvold Varmus stated that in MMTV (mouse mammary tumor virus) induced mammary tumors of mice, a gene named as *Int-1* (*Integration1*) is activated due to the integration of proviral DNA. This gene is now called as *Wnt1*. *Wnt1* is classified as a proto-oncogene which is responsible for encoding a cysteine rich secretory protein [Nusse and Varmus, 1982]. This drew the interest of scientific community in finding the potential of this proto-oncogene in the field of carcinogenesis. Then, in the year 1987 another important advancement took place where in a homologue of *Wnt1* gene was identified in *Drosophila melanogaster*, the gene called *Wingless* (*Wg*) which is responsible for segment polarity [Rijsewijk *et al.*, 1987]. Subsequently, various studies were done to define the phenotypic effect of Wnt signalling genes by doing epistasis and mutation experiments in *Drosophila* and this work done on the mutants of *Drosophila* was awarded with Nobel prize in the field of physiology in 1995 [Klaus and Birchmeier, 2008]. Several proteins contributing in delineating the segmental polarity of this organism was identified which included porcupine, dishevelled, armadillo (β -catenin) and GSK3 [Noordermeer *et al.*, 1994]. The term “Wnt” used these days was coined by combining *Wingless* (*Wg*) and *Int-1* (*Integration1*). At first instance, the functional significance of *Wnt1* gene in vertebrates was identified in *Xenopus* by injecting mRNA of *Wnt1* in frog embryos which led to the formation of duplicate body axis in them [McMahon and Moon, 1989]. Various outcomes from the studies done in *Drosophila* and *Xenopus* led to the disclosure of this evolutionary conserved pathway termed as canonical Wnt pathway. After this, to fill the gaps in deciphering the canonical Wnt pathway, TCF/LEF transcription factors were identified which are the nuclear effector molecules of this pathway. Other molecules which acts as receptors to the Wnt ligands called Frizzled receptors and co-receptors LRP5/6 were also identified [Clevers and Nusse, 2012]. Furthermore, a detailed list of all the important milestones of Wnt field in cancer biology has been summarized in table 2.5.

Table 2.5: Milestones of Wnt pathway in cancer biology

S.No	Year	Important advancements in the field of Wnt pathway and cancer	Reference
1.	1982	MMTV-Induced mammary gland tumor was related to a mouse proto-oncogene or <i>Int1</i>	Nusse and Varmus 1982
2.	1987	Identification of a homologous gene <i>Wingless</i> in <i>Drosophilla</i> which was found similar to <i>Int1</i>	Rijsewijk <i>et al.</i> , 1987 Cabrera <i>et al.</i> , 1987
		Deletions at chromosome 5q21-22 were linked with familial adenomatous polyposis, an heritable human colon cancer	Bodmer <i>et al.</i> , 1987 Leppert <i>et al.</i> , 1987
3.	1989	Patients suffering from familial adenomatous polyposis and sporadic cancer were found to screened for alterations in <i>APC</i> gene and it came out positive	Ashton <i>et al.</i> , 1989 Grodén <i>et al.</i> , 1989
		<i>Drosophilla</i> was studied to find other molecules of the Wnt cascade by screening for various other mutations	Perrimon <i>et al.</i> , 1989
		Effect of <i>Wnt1</i> gene was analyzed by injecting its mRNA in the embryo of <i>Xenopus laevis</i> and it led to the formation of secondary axis of the body	Mcmohan and Moon, 1989
4.	1990	Multiple neoplasia of intestine in mice was attributed to the mutation which was found to be dominant called <i>Min</i>	Moser <i>et al.</i> , 1990
		<i>Wingless</i> gene was reported to regulate the accumulation of Armadillo in <i>Drosophilla</i>	Riggelman <i>et al.</i> , 1990
5.	1991	Molecular cloning of β -catenin was carried out	McCrea <i>et al.</i> , 1991
		Cloning of TCF and LEF transcription factors was also done	Waterman <i>et al.</i> , 1991
6.	1992	<i>Min</i> mutation which was found in mice was identified as a truncation of the APC protein	Su <i>et al.</i> , 1992
		A new molecule Shaggy was elucidated to mediate Wnt signalling	Siegfried <i>et al.</i> , 1992
7.	1993	An interaction between APC and β -catenin was observed	Rubinfeld <i>et al.</i> , 1993
8.	1994	Antagonising effects of Shaggy and Wingless were observed in stabilizing and phosphorylating Armadillo	Peifer <i>et al.</i> , 1994
		Another important component of Wnt pathway was elucidated, named as Dishevelled	Noordermeer <i>et al.</i> , 1994
9.	1995	Stability of β -catenin was found to be associated with APC protein	Munemitsu <i>et al.</i> , 1995
		Genetic influences on the developmental pathway of <i>Drosophilla</i> were deciphered by Weischaus and Lewis. Also, a Noble prize in the field of physiology was awarded to Nusslein Volhard.	
10.	1996	In colorectal cancer, a high concentration of β -catenin was observed in the nucleus.	Inomata <i>et al.</i> , 1996
		Wnt ligands were processed through a transmembrane protein called Porcupine.	Kadowaki <i>et al.</i> , 1996
		Body axis formation in <i>Xenopus</i> was found to regulated by the location and symmetry of β -catenin in the nuclear region	Schneider <i>et al.</i> , 1996

		An important receptor at the surface of the cell responsible for interacting with Wnt ligands was identified and termed as Frizzled (FZD) transmembrane receptor	Bhanot <i>et al.</i> , 1996
		A strong affinity of β -catenin with the transcription factors TCF and LEF was reported which was responsible for translocating them to the nucleus and hence regulating the gene expression	Huber <i>et al.</i> , 1996
11.	1997	Mutations in β -catenin were evident in melanoma and colon cancer cell lines	Rubinfeld <i>et al.</i> , 1997
		Determination of three dimensional structure of β -catenin	Huber <i>et al.</i> , 1997
		The process of proteolytic ubiquitination of β -catenin is mediated by its phosphorylation carried by F3 ligase and its interaction with β -TrCP	Salomon <i>et al.</i> , 1997
		Homeotic gene named as <i>Ubx</i> was identified as the target gene of the Wnt pathway	Riese <i>et al.</i> , 1997
		Evidences reported the soluble extracellular negative regulators of the Wnt pathway called secreted frizzled related receptors (sFRP2 and FRZB)	Wang <i>et al.</i> , 1997
12.	1998	An internally located inhibitor of Wnt pathway called <i>Axin1</i> was identified	Zeng <i>et al.</i> , 1997
		Another target gene of Wnt pathway which was identified is the proto-oncogene <i>MYC</i>	He <i>et al.</i> , 1998
		The small intestine of TCF 4 null mice lacked the stem cell compartment	Korinek <i>et al.</i> , 1998
		A new interaction of β -catenin, GSK3 β and APC with Axin1 and Axin2 was reported. This interaction was found to promote the GSK3 β mediated phosphorylation which leads to proteolytic degradation of β -catenin	Behrens <i>et al.</i> , 1998
		Another secreted Wnt antagonist called Dickkopf (DKK) protein was elucidated	Glinka <i>et al.</i> , 1998
		Co-repressor of transcription factors LEF and TCF were identified and termed as Grucos	Levanon <i>et al.</i> , 1998
13.	1999	Intestinal polyposis was observed in mice having conditional gain of function alteration in β -catenin	Harada <i>et al.</i> , 1999
		Another molecule regulating the β -catenin functioning called Casein kinase 1 (CK1)	Peters <i>et al.</i> , 1999
		Another protein called Protein phosphatase 2A (PP2A) was identified and it was found to interact with β -catenin degradation complex and influences the function of GSK3 β	Seeling <i>et al.</i> , 1999
		Another Wnt antagonists was identified, WIF1 which interacts with Wnt ligands	Hsieh <i>et al.</i> , 1999
14.	2000	Genetic alterations in <i>Axin1</i> and <i>Axin2</i> gene were evident in case of colon cancer patients	Liu <i>et al.</i> , 2000
		Few co-receptors of Frizzled related receptor were identified namely Arrow, LRP5 and LRP6	Pinson <i>et al.</i> , 2000
		Determination of three dimensional structure of complex formed by β -catenin and TCF	Graham <i>et al.</i> , 2000
		Genetic silencing in <i>Axin1</i> gene have been implicated in hepatocellular carcinoma	Satoh <i>et al.</i> , 2000
15.	2001	Implication of epigenetic silencing of <i>sFRP</i> and enhanced expression of FRZD was elucidated in the process of carcinogenesis	Katoh <i>et al.</i> , 2001
		Another dimension of Wnt signalling was deciphered where in LRP5 was found to augment the Wnt cascade by	Mao <i>et al.</i> , 2001

		translocating axin towards the cell membrane	
16.	2002	Wnt antagonist Kremen was recognised as the DKK receptor which negatively regulates Wnt signalling by relocating the LRP receptor from the membrane into cytoplasm	Mao <i>et al.</i> , 2002
17.	2003	Overexpression of Dishevelled protein was observed during carcinogenesis	Okino <i>et al.</i> , 2003
		Lack of expression of extracellular Wnt inhibitory factor (WIF1) has been revealed in different cancers	Wissmann <i>et al.</i> , 2003
18.	2004	Evidences support the presence of negative regulatory molecules that control the tumorigenic effect of β -catenin -TCF complex	Lepourcelet <i>et al.</i> , 2004
		Activating factors of the Wnt signalling were reported, they were identified as Norrin and R-spondins.	Xu <i>et al.</i> , 2004
19.	2005	Sclerostin (SOST) was classified as an inhibitor of Wnt signalling. It blocks the signal cascade by binding to LRP5 and LRP6	Li <i>et al.</i> , 2005
20.	2006	Tumors of the human sebaceous gland showed the presence of mutations in LEF1. This depicts that Wnt signalling is dysregulated in sebaceous glands tumor.	Takeda <i>et al.</i> , 2006
		Another transcription factor called Hyrax was found to be associated with β -catenin	Mosimann <i>et al.</i> , 2006
21.	2007	LRP mutations enhancing the Wnt signalling were recognised in thyroid tumors	Bjorklund <i>et al.</i> , 2007
		Polymerization of Dishevelled protein near the cell membrane and localization of axin proteins to activate the Wnt signalling	Schwarz-Romond <i>et al.</i> , 2007
		Expression of two microRNAs is also influenced by Wnt and Nodal signalling cascade	Martello <i>et al.</i> , 2007
		A new intestinal stem cell marker LGR5 has been identifies as the target gene of Wnt pathway	Barker <i>et al.</i> , 2007
22.	2012	ZNRF3 and RNF3 possess E3 ubiquitin ligase activity which acts on the Frizzled related receptor leading to the down regulation of Wnt cascade. These two transmembrane proteins are found to be mutated in various cancers.	Koo <i>et al.</i> , 2012
23.	2016	Biallelic truncated mutations have been reported in <i>sFRP4</i> gene in the patients suffering from Pyle's disease	Simsek Kiper <i>et al.</i> , 2016

2.2.2 Canonical Wnt pathway

Canonical pathway is the primary pathway which is essential for the formation of primitive streak during the development of the embryos. Various secretory glycoproteins of Wnt family along with the different lipoprotein receptors form the centrally important transduction pathway which takes into account all the major cell regulatory functions in adults and intracellular signals during embryogenesis [Lindsley *et al.*, 2006]. In mammals, 19 Wnt genes have been identified which form the part of 12 conserved Wnt subfamilies and encode for cysteine rich protein having 350-400 amino acids. They also contain a signature signal peptide at the N-terminal [Kusserow *et al.*, 2005]. The first Wnt ligand which was purified and characterized biochemically was murine Wnt-3A [Willert *et al.*, 2003]. When the interaction of the wnt ligand takes place with the receptor, three different cascades are augmented in the cell, namely β -catenin dependent canonical pathway, planar cell polarity (PCP) Non-canonical pathway and Wnt/ Ca^{2+} cascade [Veeman *et al.*, 2003]. Canonical pathway is mediated by a central protein molecule, β -catenin which transduces the signal into nucleus. The level of β -catenin in the cytosol determines the fate of the signal cascade. Nuclear accumulation of β -catenin is essential to transverse the effect of Wnt ligand to the downstream target genes of the signalling pathway. In the absence of a Wnt ligand binding to its receptor complex, the cytoplasmic β -catenin is degraded by the “destruction complex”. In this complex, Axin, along with APC, GSK-3 β and casein kinase 1 α (CK1- α) bind to facilitate the sequential phosphorylation of β -catenin. Accordingly, phosphorylated β -catenin is recognized by β -transducin-repeat-containing protein (β -TrCP) and constantly degraded by the ubiquitin-proteasome pathway [Gordon and Nusse, 2006] as shown in fig 2.5 (a).

Wnt signalling is activated *via* ligation of Wnt molecules to their respective cell surface receptors which are basically composed of the seven trans-membrane frizzled (Fz) proteins and LRP5/6. Upon docking to their receptors, the cytoplasmic protein Dvl is recruited, phosphorylated and activated. Activation of Dvl induces the dissociation of GSK-3 β from Axin2 and leads to the inhibition of GSK-3 β . Hence, the phosphorylation and degradation of β -catenin is inhibited as a result of the inactivation of the “destruction complex”. Subsequently, stabilized β -catenin translocate into the nucleus, where it forms an active transcription complex with members of the T-cell transcription factor family (LEF1, TCF1, TCF3, TCF4) and the transcription initiator p300 [Angers and Moon, 2009]. Successful assembly of the transcription complex leads to target gene activation as demonstrated in fig 2.5 (b). Target genes of the canonical β -catenin pathway include matrix metallo-proteinases (*MMP2*, *MMP3*, *MMP7*, and *MMP9*) [Tamamura *et al.*, 2005], *cyclin D1* [Shtutman *et al.*, 1999], *cox-2* [Mann *et al.*, 1999], *c-myc* [He *et al.*, 1998], *c-jun* [Mann *et al.*, 1999], *Fra-1* [Longo *et al.*, 2002] and *VEGF* [Zhang *et al.*, 2001]. These lead to changes in different target gene expressions that regulate cell proliferation, differentiation, survival, cell polarity and angiogenesis.

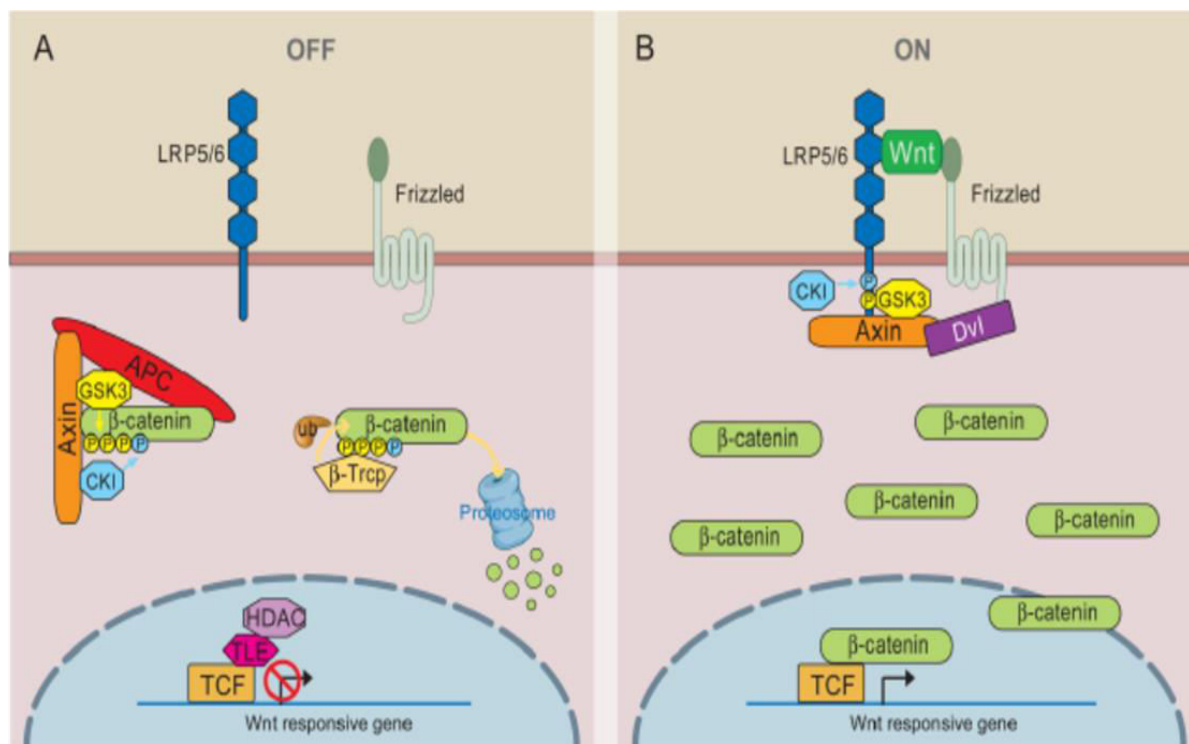


Fig 2.5: Schematic representation of Wnt/β-catenin signaling; a: showing the signal transduction when Wnt pathway is in off state; b: showing the signal transduction when Wnt pathway is in on state. (Source: McDonald *et al.*, 2009)

2.2.3 Wnt Antagonists and agonists

Regulation of this centrally important pathway called as Wnt signalling cascade is crucial and is checked at multiple cellular stages. The tight regulatory machinery of the Wnt pathway has several major antagonists such as sFRP, WIF-1, Xenopus cerberus, Wise, and DKK families [Clevers and Nusse, 2012] as illustrated in fig 2.6 (a).

sFRP and WIF-1 proteins are extracellular protein which exhibit a structural homology with the Frizzled- related receptors and are responsible for negatively regulating the signal cascade by competing with the Frizzled related receptors for binding to the Wnt ligands which activate this pathway. DKK proteins are secreted glycoproteins involved in the process of antagonising of Wnt pathway at the next cellular stage where it binds with the LRP co-receptor and due to this interaction there is disruption of Wnt ligand/Frizzled receptor/LRP complex [Clevers and Nusse, 2012]. There are two families of agonists also as shown in fig 2.6 (b), which come into play during the process of embryogenesis. They include Norrin and R-spondin proteins. Norrin is a specific ligand for FZD4 and acts through FZD4 and LRP5/6 during retinal vascularization. R-spondin proteins exhibit synergy with Wnt, Fz, and LRP6, and show genetic interaction with LRP6 during embryogenesis, but their mechanism of action is still unclear [Clevers and Nusse, 2012].

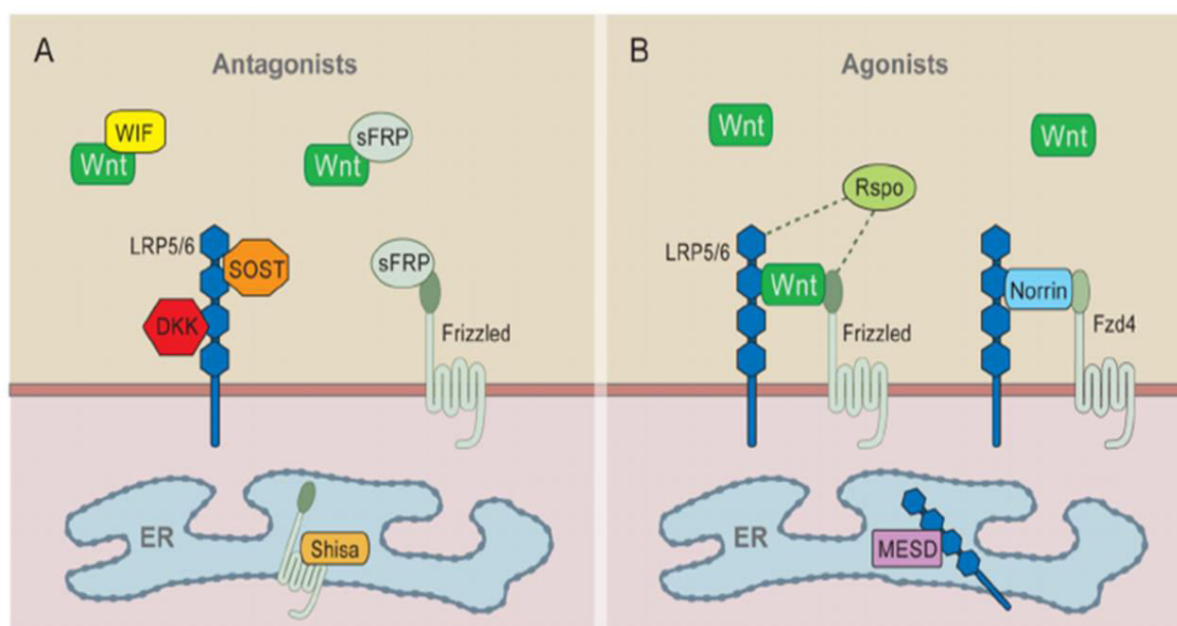


Fig 2.6: Schematic representation of Wnt antagonists regulating the pathway at different subcellular levels; a: secretory antagonists of Wnt pathway; b: agonists of Wnt pathway (Source: McDonald *et al.*, 2009)

2.2.3.1 Secreted Frizzled related receptor (sFRP)

Secreted frizzled related proteins belong to the family of extracellular glycoproteins which include five different members, *sFRP1*- *sFRP5*. Out of the five members, *sFRP1*, *sFRP2* and *sFRP5* are closely related and are coded by three exons located on chromosome 8p12-p11.1, 4q31.3 and 10q24.1 respectively. On the other hand, *sFRP3* and *sFRP4* form another group and both are encoded by six exons located on chromosome 2q31-q33 and 7p14-p13, respectively. Phylogenetically, *sFRP1*, *sFRP2* and *sFRP5* appear most closely related, whereas *sFRP3* and *sFRP4* form a separate subgroup. All the five members have been confirmed to antagonise the Wnt pathway [Jones and Jomary, 2002]. Structurally, sFRPs comprise a C-terminal domain and an N-terminus domain, each operating independently of the other. The N-terminus is comprised of a 120 amino acid long CRD made up of a conserved region of ten cysteine residues. It also contains a 20–30 amino acid signal peptide [Shi *et al.*, 2007]. Reports have shown the CRD is important in the inhibition of angiogenesis [Zerlin *et al.*, 2008]. Within the C-terminal domain is the netrin-like domain (NLD), which precedes a conserved region of six cysteine residues. Recent studies have demonstrated that the NLD can augment the process of apoptosis [Longman *et al.*, 2012]. The CRD and the NLD have both been shown to increase intracellular calcium levels, leading to the activation of the Wnt/Ca²⁺ signalling pathway. Characterization of *sFRP4* gene revealed the presence of identical motifs for potential regulatory elements such as the STAT3 (signal transducer and activator of transcription-3) and CREB (cyclic AMP responsive element binding protein) and these may play roles in regulating tissue-specific expression patterns [Pohl *et al.*, 2015].

This pathway is regulated by secreted-frizzled related proteins which basically shares a structural homology at the N-terminal with the Frizzled- related receptors. This structural homology aids them in antagonising the complex signal cascade by competing with the Frizzled related receptors for binding to the Wnt ligands which transduce the signal cascade. Alongside, recent studies support that an altered expression of these proteins have significantly contributed to various cancers such as colorectal, ovarian, prostate cancer *etc.* [Surana *et al.*, 2014]. Allelic loss of the genomic region where *sFRP3* is located has been observed in various cancers such as colorectal, prostate, gastric, papillary bladder, head and neck cancer, and non-small cell lung carcinomas [Shanmugam *et al.*, 2007]. Malignant mesothelium and osteogenic sarcoma when experimentally examined reported a lower expression of sFRP3 protein. *sFRP3* gene has a distinct mechanism of regulation which is still unclear as it lacks the CpG-island in the promoter region unlike the other members of the sFRP family [Shi *et al.*, 2007]. Other disease circumstances where sFRPs have shown increased expression are those associated with increased cell death, particularly those where apoptosis is occurring such as heart failure and degenerative human retinal disease *etc* [Lee *et al.*, 2006].

2.2.3.2 Dickkopf proteins (DKK)

Wnt signalling cascade is regulated by the secreted glycoproteins DKKs which antagonise the signal cascade by binding the LRP co-receptor, in doing so it disrupts the interactive complex of Wnt ligand/Frizzled receptor/LRP. DKK family includes four proteins DKK1-DKK4 which are identified by two cysteine rich domains (cys1-cys2) separated by a non-conservative linker region where in the location of 10 cysteine residues is highly conserved within the DKK family and it also possess a N-terminal signal peptide. N-terminal CRD (cys-1) is unique to the DKK family and the C-terminal CRD (cys-2) shares structural similarity with a colipase fold found in a wide range of proteins. Cys-2 appears to be responsible for the Wnt inhibitory function of DKK1, 2, and 4. The *DKK* genes are located on different chromosomes, *DKK2* located at 4q25 contains four exons whereas *DKK3* is located at 11p15 and is comprised of seven exonic regions. *DKK4* is located at 8p11 and has seven exons. It appears that *DKK3* is a divergent member of the DKK family by both structural and functional criteria. Although *DKK4* has displayed Wnt antagonist activity indistinguishable from *DKK1* and *DKK3* do not seem to function in Wnt signalling [Niehrs, 2006]. *DKK2* shows both agonistic and antagonistic function of Wnt [Choi *et al.*, 2012].

Recent studies have demonstrated that *DKK3* is downregulated in case of breast, cervical, NSCLC, prostate and renal cancer [Veeck and Dahl, 2012]. Also, the epigenetic hypermethylation of the *DKK3* promoter has been observed in high percentage of lung tumors [Yue *et al.*, 2008]. The expression of *DKK2* gene is reported to be downregulated in melanoma and gastric cancer [Sato *et al.*, 2007]. There are evidences supporting the oncogenic role of *DKK3* gene in process of lung cancer progression which is frequently inactivated in lung cancer [Kuphal *et al.*, 2006]. Certain epigenetic alterations have also been seen in renal carcinoma in context of *DKK2* gene. Recent evidences have shown the

role of DKK3 protein as a key factor in lung cancer pathogenesis [Nozaki *et al.*, 2001]. This can be clearly understood by a study which states that DKK3 protein is antagonising β -catenin driven Wnt pathway in lung cancer cells, which in turn leads to the lowering of expression of Wnt target genes such as *c-myc* and *cyclin D1* specifically in NSCLC cell lines [Yue *et al.*, 2008].

2.2.4 Wnt Receptors and their interactions

Frizzled receptors of the Wnt pathway are classified as the obligate receptors, without them the signal induced by Wnt ligand cannot be transduced to the effector molecules of the signalling cassette. A total of ten frizzled receptors have been identified in mammals which provide the platform for Wnt ligands to bind their extracellular domain which is cysteine rich region called CRD at the N-terminal [Wang *et al.*, 1996]. Certain structural features provide a distinct characteristic to all the Frizzled receptors as shown in fig 2.7. Some of them are listed as follows:

- These receptors certainly possess a signal peptide.
- Also, the presence of the 120 amino acid long sequence having 10 highly conserved cysteine residues is a characteristic feature.
- Another important feature is the flexible linker region which is composed of diverse 40-100 amino acids.
- Structurally, Fz receptors are composed of seven transmembrane domains connected by an extracellular cytoplasmic loop along with the cytoplasmic tail.

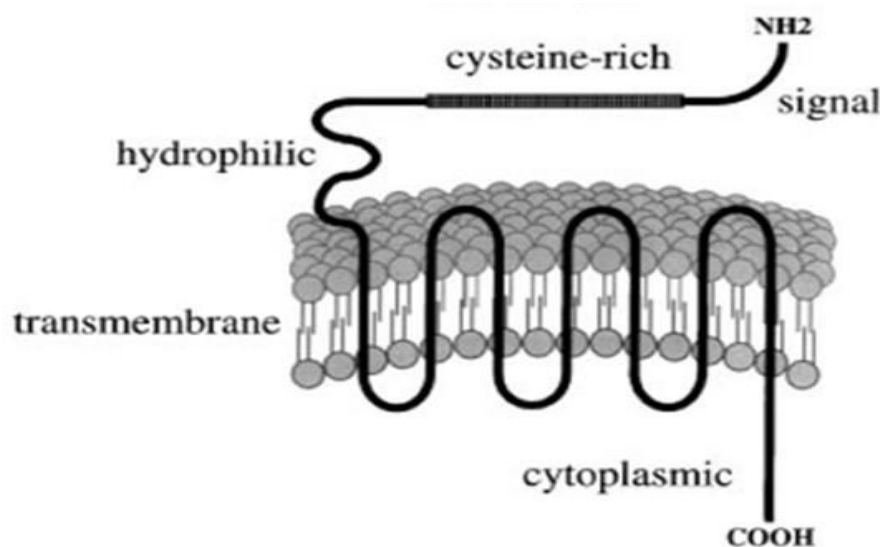


Fig 2.7: Overview of structure of frizzled receptors showing its different domains (Source: Jone *et al.*, 1996).

CRD forms the major ligand binding site for the Fz receptor which is hydrophobic in nature to enable its interaction with the lipid region of the Wnt ligand. Expression of cysteine rich region attached to the cell membrane *via* a glycosyl phosphatidylinositol molecule forming the anchor for cysteine rich

domain and thus it provides enough binding sites for Wnt ligands to interact. The signal induced due to the binding of Wnt ligand travels through the signal cassette *via* two major proteins which form the immediate interaction with the Fz receptor inside the cell. These proteins are Dishevelled (Dsh) and Axin which have the proficient domains that enable them to do so as shown [Wodarz and Nusse, 1998] in fig 2.8. Another important property which is not found in all the Fz receptors, is the presence of specific sequence S/TXV which is the putative site for PDZ domains which is responsible for anchoring the membrane receptor proteins to the cytoskeletal elements. The most potential candidate to bind to the C-terminus of Fz receptor is Dsh, this may be because of the genetic similarity it confers with the receptors containing a PDZ domain. Another nature of Fz-Wnt ligand interaction is its promiscuous nature where in it a single receptor can bind to various Wnt ligands and vice versa. This fact has been deciphered by analyzing the structure of the Wnt-Fz complex [Clevers and Nusse, 2012].

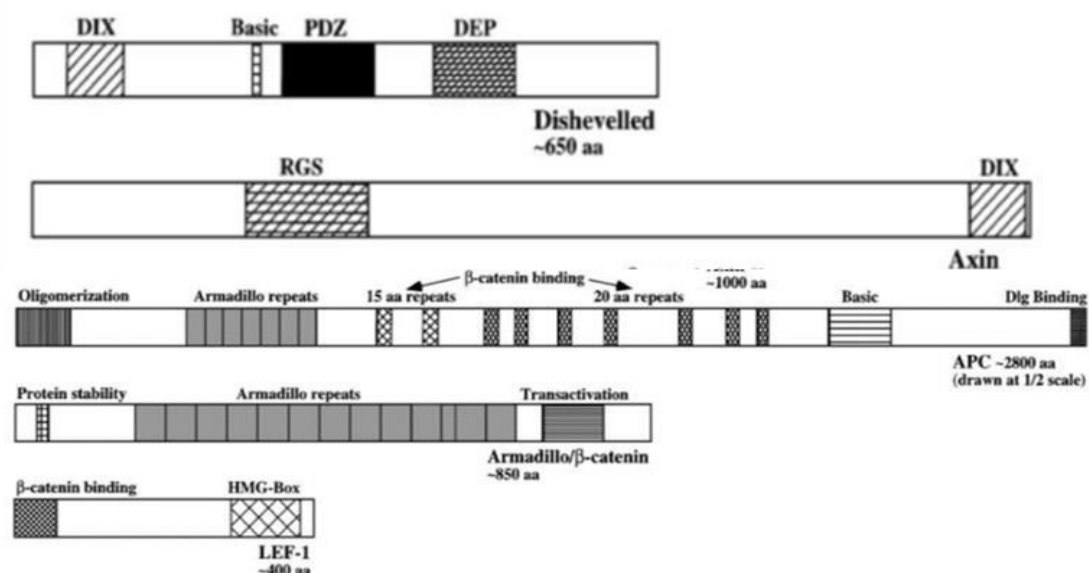


Fig 2.8: Protein domains of Dishevelled, Axin, APC, β-catenin and LEF-1 found out either by sequence similarity or by functional assays (Source: Wodarz and Nusse, 1998)

Another cognate receptor which cooperates and coordinates with Fz receptor is LRP family of proteins. LRP is characterized as a single pass transmembrane protein and forms the part of the Wnt receptor complex. It possesses distinct sites for interacting with different Wnt proteins. These dimeric Wnt receptors undergo a conformational change when a ligand binds and further it transduces the signal by initializing the phosphorylation of the target proteins. LRP is also responsible for another pivotal interaction with the Axin protein which binds to its cytoplasmic tail. This binding depends on the phosphorylation status of the LRP6 tail which is carried out by two kinases namely GSK3β and CK1γ. PPPSP (proline rich) motif of all the Wnt proteins such as Axin, APC, β-catenin and potentially LRP's is phosphorylated at the serine residue by GSK3β [Mao *et al.*, 2001]. CK1γ is anchored in the membrane by its palmitoylated C terminus and it phosphorylates the same proteins on

residues adjacent to PPPSP motif. This kinase is induced as a direct response to the signal [Davidson *et al.*, 2005]. The interaction which facilitates the binding of tail region of LRP with Axin protein is the coordination that takes place between cytoplasmic region of Frizzled receptor and Dsh. The DIX domain present on dishevelled protein can also communicate with the similar region located on the Axin protein. These two domains can bind directly to each other. The structure of this domain is not elucidated. The basic purpose of this domain is to aid in the oligomerization both hetro and homodimers of the proteins [Fiedler *et al.*, 2011].

A unique characteristic of this signalling pathway is the way this cascade works which is entirely different as compared to other pathways. Generally, the protein phosphorylation in other pathways leads to the amplification of the signal where multiple substrates are phosphorylated by a single kinase. On the contrary, in this pathway LRP6 phosphorylation which is induced by the binding of Wnt ligand titrates away the antagonist Axin elucidating a stoichiometric way of transduction as compared to catalyst mediated way of signal transduction. It has also been stated that level of GSK3 β is also regulated by Wnt signalling in a stoichiometric manner by sequestering the enzyme inside multi-vesicular endosomes [Clevers and Nusse, 2012].

2.2.5 Destruction complex affecting the stability of β -catenin

Cytoplasmic levels of β -catenin are controlled by the degradation complex which is a large assembly of multimeric proteins responsible for mediating the Wnt cascade. This large multimeric protein assembly comprises of a scaffold protein called Axin which interacts with both β -catenin, APC, WTX and the two serine threonine kinases which are expressed constitutively [Macdonald *et al.*, 2009] as shown in fig 2.9.

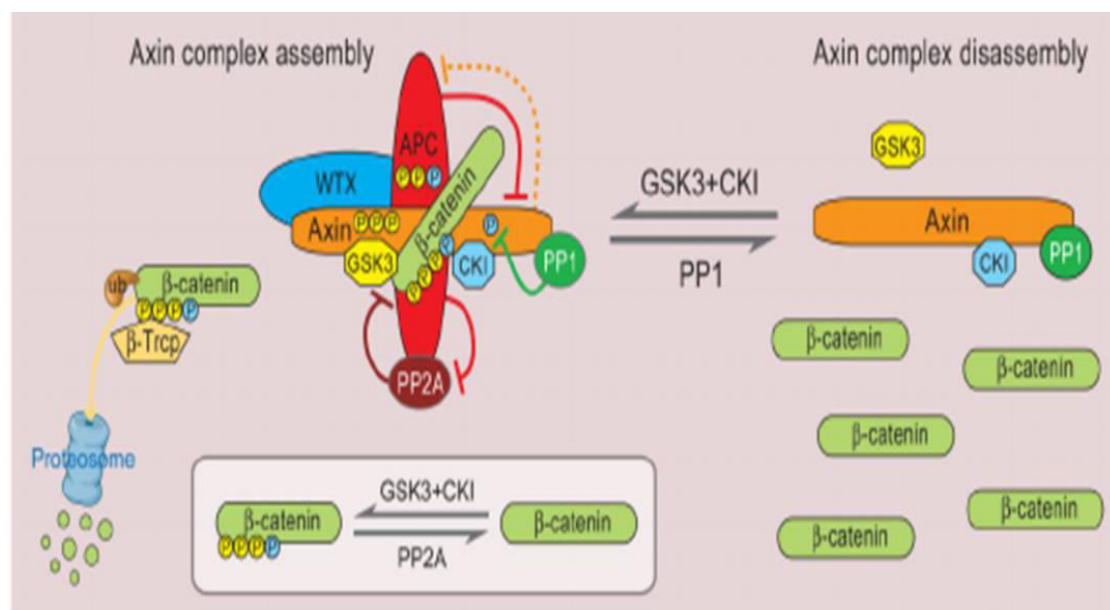


Fig 2.9: Schematic representation of destruction complex for β -catenin degradation regulated by Axin2 (Source: Mcdonald *et al.*, 2009)

APC, being a large protein has interaction sites both for β -catenin and three sites for Axin which are spread between a span of 15 and 20 amino acids repeats which is a binding site for β -catenin as illustrated in fig 2.8. The indispensable presence of APC in the destruction complex has been already established by studies done on colorectal cancer, however its exact functionality at the molecular level still remains unclear. The sequence of events which takes place when there is no ligand and Wnt receptor complex is not involved, the two kinases namely CK1 and GSK3 start the phosphorylation of β -catenin attached to Axin at the Ser/thr residues located near the N-terminal as shown in fig 2.10 [Stamos and Weis, 2013]. The next step is the recognition of the motif which is phosphorylated by the F-box repeat protein β -TrCP which is the constituent of E3 ubiquitin ligase complex. Consequently, ubiquitination of β -catenin takes place and proteasome carries the targeted destruction further blocking the activation of Wnt target genes in the nucleus [Hart *et al.*, 1999]. In the other case, when the receptor complex is engaged with the ligand the first sequential step that takes place is the recruitment of Axin to the tail of LRP5/6 which is already phosphorylated. Due to the displacement of Axin protein from destruction complex, the β -catenin fails to undergo ubiquitination as the degradation complex reaches to the saturated levels of the phosphorylated β -catenin. This in turn increases the cytoplasmic content of β -catenin and it gets accumulated in the nucleus [Macdonald *et al.*, 2009].

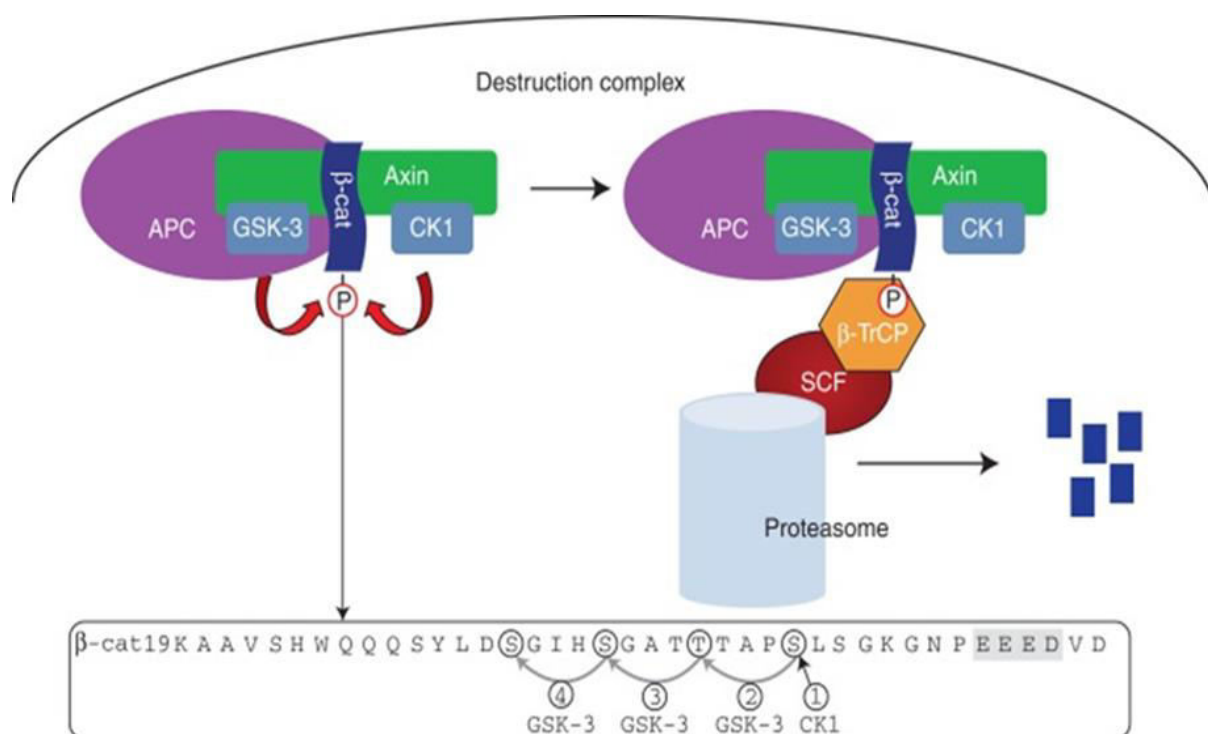


Fig 2.10: Destruction complex showing the phosphorylated sequence of β -catenin at N-terminal carried out by CK1 and GSK-3. The acidic cluster that primes CK1 phosphorylation is marked with grey colour; SCF: Skp1/cullin/F-box complex (Source: Stamos and Weis, 2013)

2.2.5.1 Axin

Discovery of Axin took place when a disruption at the mouse fused (Fused) locus lead to the duplication of body axis in embryo [Zeng *et al.*, 1997]. The vetrebrates and nematodes express two axin genes, *Axin1* and *Axin2/conductin* whereas *Drosophilla* bears a single axin gene. In case of mouse, the isoforms of *Axin* gene are interchanged at mechnasitic level and they differ only in their transcription pattern. The more widely expressed gene is *Axin1* whereas *Axin2* exhibits a differential pattern which is dictated by Wnt cascade [Chia and Costantini, 2005]. *Axin2* is an architectural protein which forms the scaffold of this destruction complex. In a way, *Axin2* negatively controls the Wnt signalling pathway and it has been established as a tumor suppressor gene. *Axin2* is located on chromosome 17q24 in humans and codes for 843 amino acid long protein. *Axin2* gene comprises of 10 exons and forms a major component of multiprotein complex comprising Dsh, APC, and GSK-3 β which regulate the level of β -catenin in the cytoplasm. β -catenin has the ability to interact with the transcription factors such as TCF/LEF and augment the process of transcription of downstream targets of Wnt pathway such as *matrix metalloproteins*, *cox-2*, *cyclin D1*, *VEGF*, *c-myc* etc [Dong *et al.*, 2001]. Axin is reported to be a flexible and mainly unstructured protein which has binding sites for CK1, β -catenin and GSK3 β [Rubinfeld *et al.*, 2001]. The carboxy terminal bears the β -catenin binidng site and GSK3 β binding site interacts with catalytic domain of PP2A and phosphatase PP1. There is special pattern of binding mechanism in which a simultaneous binding of Axin occurs with kinases and β -catenin which enhances the effective concentration of substrate and enzyme [Yamamoto *et al.*, 2001].

In colon cancer cell lines, where there is nuclear accumulation of β -catenin due to genetic alterations in APC, an enhanced expression of *Axin2* can rescue the colon cancer cells from this genetic defect [Sakanaka *et al.*, 1998]. Reports have shown that highly frequent allelic loss and rearrangement in the human chromosome 17q23-24 region, on which the *Axin2* gene is located, in numerous cancers provide evidence that *Axin2* protein plays a role as a tumor suppressor [Phelan *et al.*, 1998]. The variations in the *Axin2* gene may cause certain modulations in the protein structure and as it forms the scaffold of the destruction complex, this may affect its binding with APC and alter the process of addition of phosphate moiety to β -catenin which is the primary step for its ubiquitination [Davoodi *et al.*, 2015]. *Axin2* gene has been found to be dysregulated in various tumors such as breast, colorectal, ovarian, hepatocellular *etc.* [Gong *et al.*, 2015].

2.2.6 Role of Wnt signalling in development

Beginning of embryonic development is marked with the formation of the primitive streak and followed by patterning of the embryo by assigning the anterior posterior axis that begins with the nascent primitive streak. Primitive streak is composed of subsets of cells present in the posterior epiblast lying near to the extra embryonic ectoderm which is characterized by a peculiar gene

expression. This primitive streak migrates to the suitable locations in the embryo [Rossant and Tam, 2004]. Formation of this primitive streak is regulated by different pathways such as Notch, Bone morphogenetic protein (BMP) and Wnt signalling. In order to induce spatially proficient cellular actions in the embryonic and extra embryonic tissues, these factors and their antagonists function in a synchronized manner to achieve the same. The activity of Notch/BMP/Wnt proteins together creates a gradient which determines the localized cellular fates in the epiblast. However, the accurate involvement of Wnt pathway is still unclear in controlling the precise gene expression schemes [Tam *et al.*, 2003]. Wnt pathway generally functions through β -catenin molecule and activates the target genes which have TCF as their transcription factor. Also, it is stated that Wnt pathway can directly induce gene expression more prominently and it might contribute in modulating the effect of other factors [Eastman and Grosschedl, 1999]. Mesoderm formation from the pluripotent epiblast cells is also largely influenced by β -catenin mediated Wnt signalling. Canonical Wnt pathway is primarily essential for the generation of primitive streak. The same has been elucidated by an experiment where mice lacking the Wnt-3A ligand, co-receptors LRP5/6 and β -catenin, the mediator of Wnt signalling fail to develop any primitive streak and also showed the absence of mesoderm. Hence, Wnt signalling plays its role in the gastrulation of embryos. However, not much is known about its contribution in the extra-embryonic tissues and preimplantation of mammalian embryos [Kelly *et al.*, 2004; Huelsken *et al.*, 2000; Lindsley *et al.*, 2006]. A secretory factor called Porcn (Porcupine gene homolog in mice) is essential for the attaching the acetyl group and then releasing the 19 Wnt ligands. The developmental functions have been revealed by experiments in mouse where deletions were carried out in Porcn in the zygotic oocyte and visceral endoderm specific regions. This at first negated the fact that Porcn mediated Wnt signalling is indispensable for preimplantation process or for implantation. It was also confirmed that Wnt signalling initiates the process of gastrulation and it is the primary event that takes place in the embryo due to Wnt cascade. Also, the initial footstep which is induced by Wnt signalling in the extra embryonic tissue is the fusion of chorion and allantois sac. Altogether, these findings elucidate the significance of Porcn dependent Wnt signalling in development of embryo and placenta and it also gives an insight about its dwindling role in lineage specification in blastocyst or changes required during preimplantation [Logan and Nusse, 2004].

2.2.7 Role of Wnt signalling in stem cell maintenance

Determination of cellular fates in terms of proliferation and differentiation is majorly influenced by canonical Wnt signalling pathway along with the specification of the axis in metazoans. These processes work in complete synchronization to enable a single cell to undergo morphogenesis to give rise to complex structures [Willert and Jones, 2006]. The other important dimension of the Wnt signalling pathway is its ability to maintain the adult stem cells located in the various organs such as skin, intestine, blood and brain. Canonical Wnt signalling has an indispensable role in nurturing the self-renewal ability of stem cells and differentiation of cells into specific lineages in various tissues

and cell types. Wnt cascade has a peculiar feature of affecting the microenvironment as the signals are short ranged which is apt for stem cells maintenance and the fate of the progenitor cells [Kim and Kahn, 2014]. Embryonic stem cells (ESC's) are characterized by pluripotent ability and their capacity to form the whole organisms from a single cell which are obtained from the inner layer of cells of blastocyst. Wnt signalling has been found to play a pivotal role in establishing and providing the self-renewal ability to ESC's. Wnt ligand: Wnt-3A is an inhibitor of GSK3 β and was found to increase the generation of colonies similar to ESC's. The mechanism that underlies in promoting the self-renewability of these cells is elucidated by the ability of Wnt signalling pathway to repress the activity of TCF3 which is an endogenous transcription factor governing the process of differentiation. Genetic mutations in proteins such as APC and GSK3 β lead to teratoma differentiation due to aberrant Wnt signalling [Merrill, 2012]. Similarly, somatic cell reprogramming and induced pluripotent stem cells (iPSC's) generation is also found to be enhanced by Wnt cascade. However, this process of somatic cell reprogramming depends on the stage and temporal activity of various factors such as LEF1/TCF1 and TCF3/4. As, it is known that the pluripotent ability of the cell can be maintained by the expression of several transcription factors such as Sox2, Oct4 and Nannog. Evidences have shown the interaction of β -catenin with these factors inducing the expression of genes responsible for pluripotency [Lien and Fuchs, 2014].

Mesenchymal stem cells are the stem cells generated from mesenchymal tissues such as adipose or placental tissue and bone marrow. They have the inherent property to differentiate into diverse cell types. Canonical Wnt signalling can differentially regulate the mesenchymal cells in a biphasic way which mainly depends on the intensity of signal. Promotion of self-renewal ability of MSC's takes place under lower levels of Wnt signalling whereas the differentiation of mesenchymal cells for osteogenesis takes place when the Wnt cascade is expressed in high levels [Yang *et al.*, 2016].

Wnt signalling has also been found to play an important role in nurturing the hematopoietic stem cells (HSC's), vascular endothelial cells formation and transformation of the HSC's during the haematopoiesis. Activation of Wnt signalling is essential when the generation of red blood cells is initiated [Clevers *et al.*, 2014]. Wnt cascade along with other pathways such as BMP aids in maintaining the lineage of progenitor blood cells. The expression of genes involved in the differentiation of the erythrocytes and transcription factors such as Smad1 and TCF which are required during the repair process of hematopoietic system is regulated by Wnt and BMP pathway [Clements *et al.*, 2011]. Wnt signalling cascade also directs the division and maintenance of cells located in the intestine. It nurtures the stem cells of intestinal cryptic regions. Both constitutive and less proficient activity of Wnt signalling has certain repercussions on the fate of intestinal cells and stem cells also. Insufficient Wnt activity leads to depletion of stem cells located in the intestinal crypts and also results in the lack of diving epithelial cells required for intestinal lining. Inversely, the chronic activation of this pathway induces the uncontrolled division of epithelial cells lining the

intestinal wall which is correlated with the initiation and progression of colorectal cancer [Gehart and Clevers, 2015; Yang *et al.*, 2016].

2.2.8 Wnt signalling in cancer

Several genetic abnormalities have been found to play a role in altering the transcriptional regulation of metabolic switches that are generally responsible for maintaining the cell number by regulating the cell proliferation rate and their survival. Advanced techniques have aided in deciphering specific profile of gene expression which can be used as an identification denominator for a particular type of cancer. Impaired activity of the cell signalling cassettes which control the genetic expression of several cell cycle regulating genes, tissue differentiation genes and apoptotic genes have been implicated in cancer. Wnt pathway has already been established as a hallmark in the field of signalling cascades influencing the process of oncogenesis [Barker and Clevers, 2006]. The oncogenic potential of Wnt pathway majorly rides on the stability and cytoplasmic content of the central mediator of this pathway that is β -catenin. However, there is plethora of proteins including several receptors, ligands, kinases and antagonists which need to function in a synchronized manner to avoid the subverted state of canonical Wnt cascade [Klaus and Birchmeier, 2008]. The first investigation which elucidated the evolving role of this signalling pathway in cancer biology was the discovery where in APC that regulates the β -catenin levels was found to be associated with familial adenomatous polyposis (FAP), which is an inheritable syndrome leading to the formation of hundreds of polyps and subsequently leading to colorectal cancer at later stages of life [Preisinger *et al.*, 1991]. Further support about the tumorigenic effect of this pathway was provided by the understanding of the gene regulatory transcription factors such as TCF/LEF and their interaction with β -catenin. Wnt pathway is regulated at various subcellular levels where both agonists and antagonists have been found to dictate the nuclear accumulation of β -catenin and its rate of phosphorylation which governs its proteolytic degradation. There have been findings stating the loss of function mutations in the molecules suppressing the transduction of signal induced [Giles *et al.*, 2003]. Losses of function mutation in APC have also been reported in majority of sporadic cancers. The stability of β -catenin in cytoplasm is largely governed by this protein and functionality of this protein depends on both the alleles encoding it. Certain truncating mutations in APC have also been found to affect its functionality where in its binding sites for the major scaffold protein of degradation complex, called Axin are lost [Polakis, 2007].

Axin is responsible for recruiting the enzymes required for phosphorylation of β -catenin namely GSK3 β and CKI which facilitates its destruction by E3 ubiquitin ligase β -TrCP complex. Mutations in the *Axin1* and *Axin2* gene have also been reported in colorectal and hepatocellular carcinoma [Lammi *et al.*, 2004]. Another important channel that leads to chronic activation of Wnt pathway is the presence of genetic defects in the central mediator, β -catenin itself where in certain mutations can help in altering its structural properties such that it can inhibit its proteasome mediated degradation.

Such mutations have been found frequently in hepatocellular cancer and medulloblastoma [Polakis, 2012]. On the contrast, the mutations in GSK3 β were found to depict no correlation with human cancers despite of the role of GSK3 β in maintaining the stability of β -catenin [Abrahamsson *et al.*, 2009]. The cognizant receptor LRP has been found associated with parathyroid and breast cancer, aberrant splicing process leads to the truncation of this protein and hence the region which interacts with secreted Wnt ligands is cleaved and this protein becomes functionally challenged [Bjorklund *et al.*, 2009]. The frame shift mutations in TCF4 in the 3' end of the gene where β -catenin binds have been commonly reported in the many cases of colorectal cancer along with instability of microsatellites [Cuilliere-Dartigues *et al.*, 2006]. These kinds of mutations ultimately lead to the constitutive expression of the Wnt cascade as they pose inhibition in the binding of several molecules such as CtBP which act as transcriptional repressors [Sjoblom *et al.*, 2006;]. Certain diseases such as acute lymphoblastic leukaemia and B-cell lymphoma have been identified with genetic alterations in CREB (cAMP response element-binding protein) encoding gene [Mullighan *et al.*, 2011]. Important interrelation of CREB protein functioning and Wnt signalling transduction is the carboxyl end of β -catenin is directly associated with CREB protein. This interactive signalling enforces a major regulation on the maintenance and division of stem cell population. As a result, CREB is an important therapeutic target in case of aberrant Wnt signalling driven tumours [Teo and Kahn, 2010].

Chronic activation of Wnt cascade in breast cancer cases is not frequently associated with the genetic abnormalities in this pathway rather the overproduction of the Wnt ligands leads to its constitutive expression. In a subset of cancer cell lines such as ovarian and breast cancer, an autocrine pattern of Wnt ligand production is observed which consequently leads to the enhanced stability of β -catenin [Bafico *et al.*, 2004]. This overexpression of the Wnt cascade can be checked by the enhancing the expression of negative regulators such as DKK which are receptor-ligand antagonist along with sFRP which is an extracellular inhibitor of Wnt cascade. Subsequently, the activity of the Wnt cascade can be lowered down leading to reduction in proliferative capacity of cells [Ai *et al.*, 2006]. Another type of tumor called medulloblastoma, is also characterized as a heterogeneous tumor having different molecular subtypes imposing varied clinical outcomes. Out of all, not more than 10% of the tumors are identified as Wnt driven tumors as β -catenin activating genetic alterations are seen at a higher frequency in these tumors. However, these tumors manifest a comparatively better outcome as compared to Myc driven disease. *Myc* is one of the target genes of canonical Wnt pathway, therefore other than aberrant Wnt signalling there lie some other additional factors which contribute to the heterogeneity of the clinical manifestations of these tumor subgroups [Kool *et al.*, 2008]. A comprehensive illustration of the dysregulation of Wnt signalling during cancer is shown in fig 2.11.

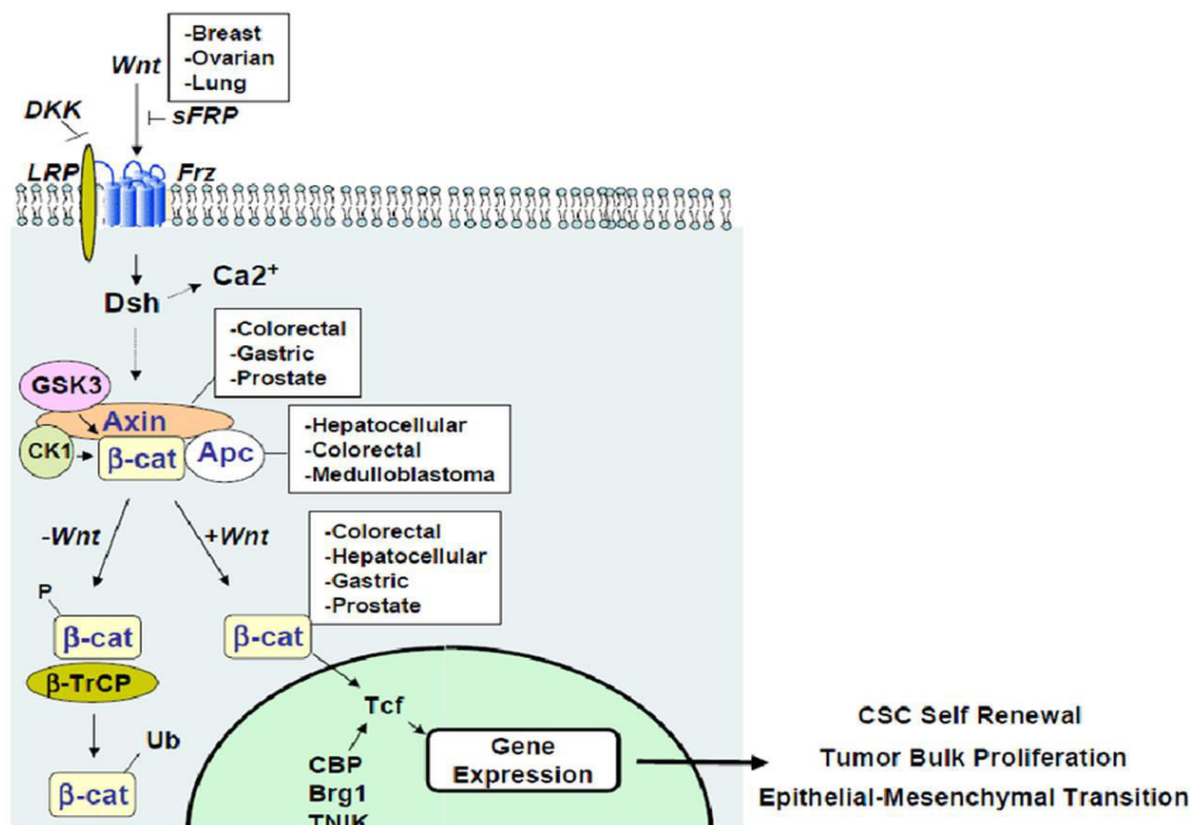


Fig 2.11: Dysregulated canonical Wnt signalling in cancer. Constitutive expression of Wnt occurs due to lack of regulation at different levels of signal cassette in different cancers which has been illustrated (Source: Agnihotri *et al.*, 2010)

The next important feature in any malignancy is its ability to metastasize and invade the adjoining tissues by surpassing through the extracellular cellular matrix and getting beyond the basement membrane of the prime tumor site [Coghlin and Murray, 2010]. Metastasis can be easily correlated with several Wnt target genes such as *E-cadherin*, *VEGF*, *MMP-7* and *CD44* which facilitate the process of tissue invasion. In a normal epithelial cell, the complex of E-cadherin and β -catenin is found to be responsible for maintaining the tight junctions and cellular polarity [Heuberger and Birchmeier, 2010]. The transition of epithelial into mesenchymal (EMT) origin is a one of the most important feature adopted by epithelial cells during the process of metastasis. Activated Wnt signalling leads to increased accumulation of β -catenin in the nucleus and hence, it disturbs the E-cadherin/ β -catenin complex facilitating the epithelial to mesenchymal transition (EMT) process [Vermeulen *et al.*, 2010; Gonzalez and Medici, 2014]. Wnt signalling has been found to enhance the expression of genes promoting EMT including *Snail1* and *Snail2*, *KCF8*, *E47*, *ZEB1* and *ZEB2* [Lamouille *et al.*, 2014]. Constitutive expression of canonical Wnt signalling cascade that place through the effector molecules regulating the transcription of other genes such as *LEF* and *HOXB9* has been characterized as an important identification marker for brain and bone metastasis that occurs during the adenocarcinoma of lung tissue [Nguyen *et al.*, 2009]. This signalling has also found its implication in the process of angiogenesis, as it leads to the overexpression of *cyclooxygenase-2* (*cox-*

2) which is responsible for the formation of new blood vessels and hence eventually promoting the process of metastasis [Previdi *et al.*, 2010].

Recently, a new concept of cancer stem cells (CSC's) is an upcoming dimension which contributes to the explanation of the carcinogenesis process. Apparently, role of Wnt signalling has been elucidated in the maintenance of these stem like tumor cells which cater to the process of tumorigenesis. There is a population of cells residing in special niches in each of the adult epithelial tissues responsible for restoring the self-renewal capacity of the same and also help in maintaining the homeostasis of the tissue. These cells possess some peculiar features of dividing indefinitely which makes them a hotspot for activating mutations that may lead to initiation of tumor formation. These cells are endowed with self-renewal capability and also have potential to differentiate into multiple types of cells and so, they are named as cancer stem cells [de Sousa e Melo and Vermeulen, 2016]. The maintenance of the stem-cell like population intimately depends on the level of Wnt activity taking place in them as shown in fig 2.12. The cells which apprehend a higher activity of Wnt signalling are posed to exhibit cancer driving properties. More than 80% of the colorectal cases are identified with APC inactivation mutations which ultimately increase Wnt signalling [Fearon, 2011]. Despite of this global increase in activity of Wnt cascade in all cancer cells, regulation may take place at additional levels and these cells with a differential activity are associated with clear biological significance. This regulation occurs at several levels and exhibit similarity to the Wnt cascade regulation taking place in normal stem cells. There are niches in the normal tissues which support the stem-cell like population by producing the Wnt ligands for example: paneth cells of the intestine are found responsible for the secretion of Wnt ligands in order to support the crypts base columnar stem cells. Similarly, in the hematopoietic stem cell population Wnt signals play a pivotal role in the process of haematopoiesis and functions of these stem cells [Malhotra and Kincade, 2009]. Various malignancies have elucidated the importance of the Wnt ligands in nurturing the CSC's. Wnt-3A has been found to enhance the sphere producing ability in estrogen (ER) negative breast cancer [Lamb *et al.*, 2013]. Also, Wnt-1 and 3A cater to the stem-like cells in glioblastomas [Kaur *et al.*, 2013]. Hyperactivity of β -catenin mediated pathway when controlled by the optimum activity of the Wnt antagonists such *sFRP*, *DKK* and *Axin2* has been found responsible for decreasing the dividing capacity of this stem cell population [Pinto *et al.*, 2003].

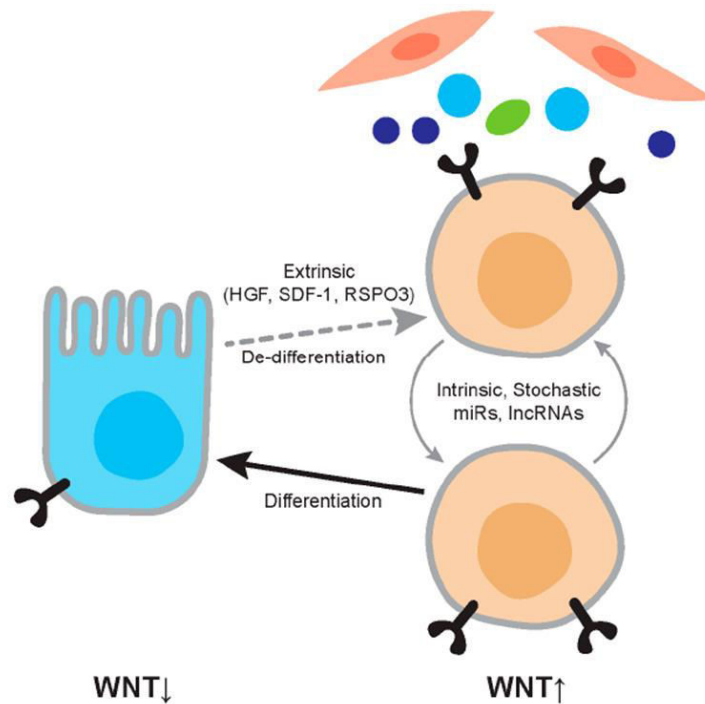


Fig 2.12: Relation between the tumor cells which are differentiated and cancer stem cells. Cancer stem cells show a comparatively higher levels of Wnt than differentiated tumor cells. The activity of Wnt depends on both extrinsic and intrinsic factors. (Source: de Sousa e Melo and Vermeulen, 2016)

2.2.8.1 Wnt signalling and lung cancer

Heterogeneity and complexity of lung carcinoma has always been a major milestone in disclosing the pathological features of this disease. Several genetic and molecular perturbations consequently lead to the formation of primary tumor of lung and then its metastasis. Mechanisms underlying the aggressive and robust capacity of lung tumors to develop and spread are largely unknown [Nguyen *et al.*, 2009]. Uncovering the genes and pathways involved in abrupt initiation and rapid spread of this disease can help in introducing the early diagnostic methods, treatment modulations and also targeting new molecules which have hidden therapeutic potential. The ontogeny of lung cancer cells has been recently related to expression of those genes which recapitulate the events taking place during embryonic development of lung tissue [Chen *et al.*, 2014]. Considering, the role of the developmental pathways in lung carcinogenesis and progression, there is a need of exploring the signal transduction cascades such as canonical Wnt pathway which has elucidated roles in maintenance of embryo, adult cells and also other variety of processes such as cell polarity, differentiation, apoptosis, and survival [Pongraez and Stockley, 2006]. Several evidences in the murine models have exhibited the effect of increased Wnt signalling in enhancing the initiation of tumor formation and particularly in case NSCLC [Vaughan *et al.*, 2011]. Another report unveiled that the respiratory epithelium which was cultured in a microenvironment having cigarette smoke showed the upregulation of Hedgehog and

Wnt pathway [Lemjabbar-Alaoui *et al.*, 2006; Liu *et al.*, 2010]. Also, the lungs of *K-ras* transgenic mice were identified with increased Wnt activity [Lee *et al.*, 2009]. Constitutive Wnt activation has increased tumorigenic potential. Several components of the Wnt pathway have been found to overexpress in the NSCLC tissues. β -catenin has been found to express in 94% of the resected SQCC lung cancer samples and also in 51% of adenocarcinoma [Kren *et al.*, 2003]. Accumulation of β -catenin in the nucleus has been correlated to the EGFR mutations within lung tissue [Suzuki *et al.*, 2007] and also, the resistance of NSCLC cell line to gefitinib can be associated to the overexpression of β -catenin [Fang *et al.*, 2014]. The enhanced expression of β -catenin in lung tumor cells can be attributed to the increase in extracellular matrix metalloproteinase inducer (EMMPRIN). Also, β -catenin mediated signalling, cell proliferation, migration and anchorage independent division can be inhibited by silencing the EMMPRIN [Sidhu *et al.*, 2010]. β -catenin may be misregulated in lung cancer cells through its interactive ability with Wolf-Hirschhorn syndrome 1 [Toyokawa *et al.*, 2011]. Although, there have been various reports about lack of β -catenin expression in cell lines still its expression has been interrelated to the increase in tumor size, stage and clinical grade in NSCLC cells [Kren *et al.*, 2003]. Overexpression of β -catenin has been found to directly influence the better prognosis but only in some studies and not all [Stewart, 2014] possibly because there is a negative effect on E-cadherin due to lack of β -catenin [Kren *et al.*, 2003]. Lung cancer tissues [Hommura *et al.*, 2002] or cell lines [Ueda *et al.*, 2001] do not commonly exhibit activating mutations in β -catenin as in the case of colorectal cancer.

Wnt ligand namely *WNT-1* has been found to overexpress in lung cancer cell by the enhanced production of sulfatases or by miR21 or periostin. A study has shown that, out of all NSCLC resected tissues 37-63% showed a positive immunostaining for *WNT-1* [Akiri *et al.*, 2009]. Increased *WNT-1* was correlated with a lower apoptotic index, dysregulated β -catenin expression and enhanced expression of *c-myc*, *cyclin D1*, *VEGF-A*, *MMP-7*, *Ki-67*, *survivin* and density of intratumoral microvessels. Clinically, the Wnt1 overexpression can be attributed to poor prognosis. Inhibition of NSCLC cell line proliferation has been achieved by down regulating WNT-1 through anti WNT-1 antibody or siRNA and this indeed led to the induction of a more differentiated phenotype. Overexpression of other components of Wnt pathway such as *WNT-2*, *WNT-3*, *WNT-5A* and *WNT-11* has also been observed in NSCLC. An association has been found in the overexpression of *WNT-2* and *FZD-8* in NSCLC tumor samples. *WNT-2* activity was also reduced in NSCLC cell lines by ablating it through different routes such as monoclonal antibody, short hairpin RNA, siRNA leading to negative construct induced programme cell death, lowering the colony formation, down regulating the Wnt target genes. *WNT-3A* expression was statistically correlated to the enhanced *ki67*, lowered apoptosis and increased activity of *c-myc* and *survivin* in NSCLC resected tissues [Stewart, 2014].

2.2.9 Genetic variations and their relevance in modulating disease susceptibility and clinical outcomes

The growing field of population genetics has been playing an intricate role in uncovering the pathophysiological facts concerned with a disease. The understanding of the differences in individual's susceptibility towards environmentally induced diseases such as lung cancer has been an impeccable milestone. Human susceptibility has a complex and intricate genetic framework and it is difficult to find instances where an alteration in a single gene or multiple genes can be considered as a hallmark in increasing the sensitivity towards a certain environmental exposure [Christiani, 2006]. Variations in multiple genes may have a cumulative effect which may be moderate or small potentially in affecting the predisposition towards lung cancer. Not only the interactions within the genes, also the synergy between genes and environmental factors along with the epigenetic alterations play a pivotal role in the inducing carcinogenesis of lung [Wu *et al.*, 2004]. In this regard, the single nucleotide polymorphisms (SNP) present across the human genome are used as a tool to conduct candidate gene association studies. The frequency of SNP's in the entire population is 1/3000 base pair and the distribution of SNP's in the human genome varies around 100 times from one region to the other. Although, a polymorphic site is present almost after every 1000 base pair around the genome. Single nucleotide variations in the genomic DNA are identified as the alternative sequences at a particular locus in normal population where in the frequency of the less dominant allelic form is atleast 1% or more. These variations have found in the different regions of the genome. Based on the locus, where the SNP lies its functional attribute can be predicted [Brookes, 1999]. Polymorphic variations found in the coding region can affect certain pertinent proteins and those located in the intronic region may have impact on the splicing mechanism of the m-RNA. Some of the polymorphic sites are found in the promoter region, where they can influence the transcriptional regulation of the gene and hence affecting the gene expression. The extent to which a single nucleotide variation can be reflected in the phenotypic expression can be deciphered and this can aid in elucidating new findings about the polymorphic variant. SNPs, these days are screened for their association with a trait or a disease to find out which are physiologically applicable. Once the polymorphic site is identified and correlated to a particular trait or disease, it is used for screening the different ethnicities to check its distribution and its role in risk profiling [Shastri, 2009]. Large scale epidemiological studies are of great relevance in this regard in order to establish a particular SNP as a potential diagnostic and prognostic tool. Another goal of such studies is to unmask the effect of these naturally occurring genetic differences among individuals in different diseases. These studies help in revealing the disease etiology and outcome by finding the difference the presence of these naturally occurring predisposed variations make [Carlton *et al.*, 2006]. However, in order to find the therapeutic relevance of the genetic profiles accounted on the basis of the general susceptibility factors which affect the abnormalities at the phenotypic level. This process requires a step wise procedure of obtaining the factor controlling the predisposition, characterizing the individuals at higher risk and finding the

detailed disease pathology. Certain molecular evaluations such as gene expression analysis, sequence structure profiling, protein expression coupled with the screening different populations for naturally occurring genomic variations is guiding light of the future research in unravelling the etiology of heterogeneous diseases [Schork *et al.*, 2000].

2.2.10 Genetic variabilities in the Wnt antagonist genes (*sFRP*, *DKK* and *Axin2*) and their correlation with cancer predisposition

Several sequence variations of the Wnt antagonist genes have been evaluated for influencing the cancer risk in different ethnicities. Two widely studied polymorphic sites of *sFRP3* gene namely, rs7775 and rs288326 have been studied in German colorectal cancer patients where only *sFRP3* Arg³²⁴Gly was found to confer an increased risk [Shanmugam *et al.*, 2007]. Another report demonstrated a lack of association of these two variants towards colorectal neoplasia [Berndt *et al.*, 2009]. The polymorphic site rs7775 at codon 324 is located at the C-terminal end of *sFRP3* gene which is a known extracellular glycoprotein and binds to Wnt ligands in order to negatively regulate the Wnt cascade. sFRP proteins interact with Wnt ligands through N-terminal frizzled cysteine-rich domain [Shanmugam *et al.*, 2007]. However, a study has shown that the direct binding between sFRP1 protein and its ligand Wnt was strongly influenced by heparin, and it should be noted that the C-terminus shows heparin-binding activity. Certain studies have also suggested the fact that in order to antagonise the Wnt signal, sFRP protein requires both N- and C-terminal although it is able to bind the Wnt ligand when truncated [Shanmugam *et al.*, 2007]. Polymorphic variant of *sFRP4* rs1802073 was found to be positively correlated with increased risk for rectal cancer and early stage colorectal cancer in German population [Frank *et al.*, 2010]. Furthermore, a study has demonstrated tumor-suppressing activities of sFRP in prostate cancer cells [Zi *et al.*, 2005]. The sFRP^{324Gly} variant and the sFRP^{200Trp/324Gly} double variant haplotype have reduced ability to antagonize Wnt signalling subsequently leading to increased β -catenin levels [Loughlin *et al.*, 2004]. sFRP gene promoter was hypermethylated in more than 80% of mesothelioma primary tissues [Shi *et al.*, 2007]. *sFRP4* Pro³²⁰Thr has been found to be associated with early stage colorectal cancer [Frank *et al.*, 2010] and also been evaluated for its role in renal cancer [Hirata *et al.*, 2009]. Although *sFRP4* rs1802073 was not found be correlated to renal cancer risk but being a non-synonymous substitution (Pro³²⁰Thr), it has been characterized by Polyphen software as a “probably damaging” change which can alter the functionalities of this extracellular glycoprotein [Hirata *et al.*, 2009]. Another polymorphic site is rs1802074, Arg³⁴⁰Lys which has also been explored for its role in process of renal carcinogenesis [Hirata *et al.*, 2009]. These two genetic variants of the *sFRP4* gene have been genotyped and associated with the bone mineral density and body composition in Danish men [Boudin *et al.*, 2012]. A study done in Turkish lung cancer patients reported the lack of functional significance of rs1802074 in influencing the susceptibility towards lung cancer [Yilmaz *et al.*, 2015].

Recently, several different variants of the *DKK4* (rs2073664), *DKK3* (rs2291599, rs3206824 and rs7391689) and *DKK2* (rs447372, rs419558 and rs17037102) have been analyzed in different cancer patients. Another study has investigated the role of *DKK2* (rs17037102), *DKK3* (rs3206824), *DKK3* intron4 G/C (rs7396187), *DKK4* (rs2073664), and *sFRP4* (rs1802074) variants in Turkish lung cancer subjects and none of the sequence variation conferred a significant effect on lung cancer risk. Only the gene combination *DKK3* rs3206824 and *sFRP4* rs1802074 conferred a significantly decreased risk of lung cancer [Yilmaz *et al.*, 2015]. Several SNPs of the *DKK3* gene such as rs2291599, rs3206824 and rs7391689 have been screened in various cancers such as lung [Yilmaz *et al.*, 2015], renal [Hirata *et al.*, 2009], prostate [Kim *et al.*, 2015] and breast cancer [Alanazi *et al.*, 2013]. Similarly, certain non-synonymous substitutions in the *DKK2* gene have been found to be implicated in patients with renal [Hirata *et al.*, 2009], colorectal [Páez *et al.*, 2014] and breast cancer [Alanazi *et al.*, 2013]. Also, the polymorphic site rs2073664 (C/T) in *DKK4* gene has been studied in renal carcinoma patients [Hirata *et al.*, 2009] and another SNP (rs3763511) located within the *DKK4* gene has been investigated in breast cancer subjects of Saudi Arabia [Alanazi *et al.*, 2013]. *DKK4* is an established inhibitor of Wnt signalling which is ligand induced, therefore it can be considered for having antitumor potential. Antitumor properties of *DKK4* pertain to its ability of activating glycogen synthase kinase- β (GSK-3 β) and the subsequent degradation of cytoplasmic β -catenin [Neihrs, 2006]. The polymorphic site rs2073664 lies in exon4 of the *DKK4* gene and results in synonymous valine substitution at 169th position. However, the functional relevance of this sequence variation is yet unclear [Yilmaz *et al.*, 2015]. Very few reports are available in regard of this SNP where its effect on prognosis of cancer patients was investigated. The constitutive activation of the Wnt pathway may be attributed to the sequence variation in the genes coding the Wnt signal regulating proteins namely Dickkopf proteins which prevent the nuclear translocation of β -catenin.

Several studies have been carried out exploring the effect of certain polymorphic sequence differences on the disease susceptibility [Marshall and Christiani, 2013]. Several genetic factors intricately influence the individual's predisposition to acquire lung cancer. There have been a few reports where the sequence variation within the *Axin2* gene was analyzed to find out its effect on cancer risk [Wu *et al.*, 2015]. One of the commonly studied single nucleotide polymorphism includes rs22040308, which has been assessed in different populations such as Turkish [Gunes *et al.*, 2009], Japanese [Kanzaki *et al.*, 2006], Polish [Mostowska *et al.*, 2014] and Chinese [Liu *et al.*, 2014]. These studies have been carried out in different carcinomas including gallbladder [Yadav *et al.*, 2016] ovarian [Mostowska *et al.*, 2014], lung [Gunes *et al.*, 2009], prostate [Ma *et al.*, 2014], astrocytoma [Gunes *et al.*, 2010] *etc.* A few studies have also evaluated several other intronic and exonic polymorphic sites including *Axin2* 148 C/T, *Axin2* 1365 G/A, *Axin2* 432 T/C, *Axin2* 956+16 A/G, *Axin2* 1712+19 G/T, *Axin2* 1386 C/T and *Axin2* 2062 C/T in Turkish lung cancer patients [Gunes *et al.*, 2009]. A similar study was done in astrocytoma patients [Gunes *et al.*, 2010]. However, the sample size was very low and further studies are required to establish the role of these genetic variants within *Axin2* gene in

modulating towards lung cancer susceptibility. Comprehensive information about the genetic variants of Wnt antagonist gene on the basis of above cited literature has been provided in table 2.6.

2.2.11 Association of genetic variations within the Wnt antagonist genes with overall survival and prognosis of cancer patients

In lieu of the above cited literature about the Wnt pathway and its functional implication in altering the various processes such as metastasis and cancer stem cell maintenance which cater to the progression of cancer, it becomes mandatory to evaluate its association with the clinical manifestations of the disease. There have been few reports focusing on the implication of genetic variants of Wnt pathway in predicting the prognosis of particularly lung cancer patients undergoing chemotherapeutic treatment, toxicity and survival. However, SNP variants are an emerging tool for predicting prognosis and survival of lung cancer patients rather than m-RNA profiling. SNP genotyping is a non-invasive technique unlike m-RNA analysis as it does not require any tissue biopsy which is slightly complicated in case of lung tumors. Most of the previously carried studies have utilized immuno staining techniques [Ryu *et al.*, 2004]. A very few studies have actually looked into the sequence variations and their effect on the clinical outcomes. Exploration of Wnt signalling pathway as a prognostic and predictive marker has been carried out in various malignancies such as hematologic [Ashihara *et al.*, 2015], breast [Kazi *et al.*, 2016], ovarian [Bodnar *et al.*, 2014], pancreatic ductal adenocarcinoma [Nakamoto *et al.*, 2014] and esophageal squamous cell carcinoma [Deng *et al.* 2015]. The genetic variations of the Wnt pathway have been evaluated in case of gall bladder cancer patients to find out their predictive role in survival of these patients [Yadav *et al.*, 2016]. One study done in Korean population examined the different sequence variations of *WNT-5A*, *β-catenin*, *WNT-16* and *FZD4* *etc.* and reported a lack of any significant association with survival of lung cancer patients [Yoo *et al.*, 2016]. Another study done in early NSCLC patients, have explored around 40 SNPs of twenty genes of Wnt pathway where some of the genetic variations in *WNT-16* and *catenin* showed a significant association with survival of these patients before and after surgical treatment [Cosico *et al.*, 2014]. A study done in renal cancer patients negated the role of the polymorphic sites in Wnt antagonist genes in predicting their survival after radical nephrectomy [Hirata *et al.*, 2009].

Table 2.6: Genetic variations of the negative regulators of Wnt signalling cascade in different human cancers

S.No	Type of Molecule	Wnt antagonists	SNP ID	Polymorphism	Functional role/ genetic association
1.	Secretory Wnt antagonists	<i>sFRP3</i>	rs288326	Arg 200 Trp	Increased b-catenin levels, Colorectal cancer
			rs7775	Arg 324 Gly	Increased b-catenin levels, Colorectal cancer
		<i>sFRP4</i>	rs1802073	Non-synonymous; Exon 6 P320T	SNP judged ‘damaging’ with PolyPhen, Rectal cancer risk, early stage colon rectal risk
			rs1802074	Non-synonymous; Exon 6 R340K	Lung, renal
		<i>DKK2</i>	rs17037102	Non-synonymous; Exon 3 Arg146Glu	SNP judged ‘benign’ with PolyPhen, Renal, colon
			rs419558	3’ UTR	Renal , colon
		<i>DKK3</i>	rs447372	Intron 1	Renal ,colon
			rs3206824	Non-synonymous; Exon 7 Arg335Gly	SNP judged ‘benign’ with PolyPhen/ Renal, prostate ,breast, lung
		<i>DKK4</i>	rs11022095	Intron 6 A/G	Renal, prostate ,breast, lung
			rs1472189	3’ Near gene C/T	Renal, prostate ,breast, lung
			rs7396187	Intron 4 G/C	Renal, prostate ,breast, lung
			rs2291599	Intron 4 C/T	Renal, prostate ,breast, lung
			rs2073664	Synonymous; Exon 4 V169V	Renal, breast
			rs3763511		Breast cancer
		<i>Axin2</i>	148 C/T	Exon1, Pro50Ser	Protective effect in Lung, astrocytoma, ovarian, prostate
			432 T/C	Exon1, Ile144Ile	Lung, astrocytoma, ovarian, prostate
			956+16 A/G	-Intron2	Lung, astrocytoma, ovarian, prostate
			1365 G/A	Exon 5, Pro455Pro	Lung, astrocytoma, ovarian, prostate
			1386 C/T	Exon5, Pro462Pro	Lung, astrocytoma, ovarian, prostate
			1712+19 G/T	Intron 5	Lung, astrocytoma, ovarian, prostate
			2062 C/T	Exon7, Leu688Leu	Tooth agenesis, Lung, astrocytoma, ovarian, prostate
			2141+73 G/A	Intron 7	Lung, astrocytoma, ovarian, prostate

2.2.12 Epigenetic alterations in the Wnt signalling pathway and their role in carcinogenesis

In order to sustain the involvement of a gene in pathogenesis of cancer formation, certain frequently occurring genetic alteration or mutations can clearly provide the evidence. Apparently, the epigenetic modulation in genes and chromatin modelling may also cater to the process of tumor development and spread by disabling the antagonist and activating the agonists of a signal transduction cascade [Rodríguez-Paredes and Esteller, 2011]. Laird *et al.*, [1995] reported the implication of methylation of gene in oncogenesis driven by Wnt signalling. They demonstrated the pharmacological inhibition and genetic depletion of Methyl transferase enzymes evidently led to the reduction in tumor burden in APC min mouse.

Further studies done using this model also showed that lack of CpG binding repressor Mbd2 also attributed to the decrease in tumor formation through the attenuation of Wnt cascade [Pheesse *et al.*, 2008]. These reports indicated towards the role of hypermethylation of Wnt antagonists in their silencing as shown in fig 2.13. A genome wide study done on medulloblastoma, pancreatic cancer and colon cancer revealed that epigenetic modification of Wnt antagonists are the common targets in these tumors [Kongkham *et al.*, 2010; Vincent *et al.*, 2011; Suzuki *et al.*, 2002]. Functional loss of Wnt inhibitors through DNA methylation and histone modification has been reported in numerous cancers. Epigenetic silencing of Wnt genes has promoted the formation of tumor favourable phenotypes of cells [Serman *et al.*, 2014]. Epigenetic inactivation would help to find out the predisposition of stem or progenitor cells to undergo abnormal clonal expansion without getting through the path of differentiation and maturity [Staal, 2016].

Particularly, the promoter region of *sFRP1* gene which interacts with the wnt ligands to antagonise the pathway was identified to be epigenetically modified in these reports [Caldwell *et al.*, 2004; Björklund *et al.*, 2007]. Other counterparts of the sFRP family have also been reported to be silenced through hypermethylation in different cancers such as gastric [Nojima *et al.*, 2007], hepatocellular [Suzuki *et al.*, 2008], melanoma [Ekstrom *et al.*, 2011] and breast cancer [Veeck *et al.*, 2008]. Recent studies have showed that impairment in the regulation of Wnt signalling inhibitors such as *sFRP*, *WIF-1*, *HDPRI* and *DKK3* through increased methylation of promoter region is evident in many human cancers [Ying and Tao, 2009]. Other proteins of the Wnt signalling pathway such as Wnt ligands Wnt-2, Wnt-3A and Wnt-5A are found to show an increased expression in several human malignancies such as breast, colon, lung and prostate [Ying and Tao, 2009]. Hence, they are identified to have the oncogenic potential of Wnt cascade. Wnt-5A was reported to act as a tumor suppressor as it inhibits the cellular proliferation by negatively regulating the β -catenin mediated Wnt pathway. It is silenced due to epigenetic alteration in hematological cancer [Ying *et al.*, 2007]. Similarly, Wnt-7A and Wnt-9A show epigenetic inactivation in pancreatic [Sato *et al.*, 2003], acute lymphoid leukemia and colon cancer [Shu *et al.*, 2006].

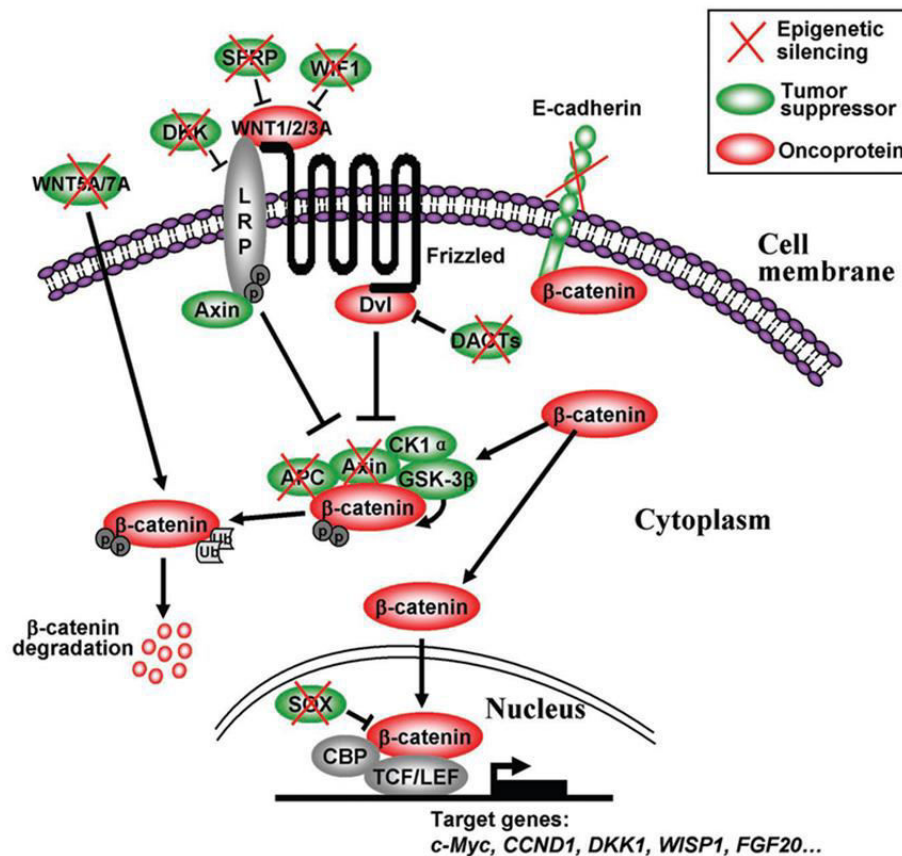


Fig 2.13: Epigenetic silencing of negative WNT regulators contributes to the aberrant activation of WNT/ β-catenin signaling in human cancers. (Source Ying and Tao, 2009)

Veeck *et al.*, [2009] elucidated that association of shorter survival of breast cancer patients with the modulated promoter methylation of *DKK3* gene. A specific tumor biomarker, *sFRP2* for early breast cancer was also reported as the aberrant promoter methylation was directly related to the breast cancer progression [Veeck *et al.*, 2008]. Similarly, *WIF1* was identified as a poor prognostic marker in case of acute promyelocytic leukemia [Chim *et al.*, 2006]. Tseng *et al.*, [2008] showed that epigenetic silencing of *Axin2* in the degradation complex lead to wild-type beta-catenin nuclear accumulation in non-small cell lung cancer tumorigenesis. Dysregulation of Wnt antagonist genes at the epigenetic level have found a strong implication in tumor progression, detection of cancer, determination of prognosis. Epigenetic silencing of the Wnt antagonists is certainly found to trigger the chronic activation of β-catenin mediated pathway and it is the earliest stage of carcinogenesis process even before the key mutations come into play. Dysregulated expression of Wnt antagonists, such as *sFRP* and *DKK*, due to epigenetic alteration results in anomalous Wnt signalling and this has been evidently reported in several human malignancies, including NSCLC [Ying and Tao, 2009] and also it was found that *DKK3* methylation is significantly associated with a poor prognosis in lung adenocarcinomas [Suzuki *et al.*, 2007]. Table 2.7 summarizes the reports available for epigenetic modulation of Wnt antagonists in various cancers.

Table 2.7: Dysregulated Epigenetic activation of the genes of Wnt signalling cascade in different human cancers (Source: Ying and Tao, 2009)

S.No	Type of Molecule	Wnt antagonists	Epigenetic change	Type of tumor
1.	Extracellular Wnt antagonists	<i>sFRP</i> family (<i>sFRP1</i> , 2, 3, 4 5)	Methylation of promoter region	Lung cancer, leukemia, breast cancer, colorectal cancer, renal carcinoma, hepatocellular cancer, esophageal cancer, cervical cancer
		<i>WIF</i>	Methylation of promoter region	Breast cancer, mesothelioma, lung cancer, bladder cancer, gastric cancer, nasopharyngeal cancer, esophageal cancer,
		<i>DKK</i> family (<i>DKK1</i> , 2, 3)	Methylation of promoter region	Colorectal cancer, leukemia, gastric cancer, bladder cancer, breast cancer, prostate cancer
2.	Cytoplasmic Wnt inhibitors	<i>DACT</i> family (<i>DACT1</i> , 2, 3)	Methylation of promoter region and modification of the histone protein	Colorectal cancer
		<i>Axin2</i>	Methylation of promoter region	Colorectal cancer
		<i>APC</i>	Methylation of promoter region	Leukemia, breast cancer, colorectal cancer
3.	Nuclear proteins	<i>SOX</i> family (<i>SOX 7</i> , 17)	Methylation of promoter region	Prostate cancer, breast cancer, colorectal cancer
4.	Wnt ligands	<i>Wnt-5A</i> , 7A, 9A	Methylation of promoter region	Colorectal cancer, Nasopharyngeal cancer, pancreatic cancer, haematological malignancies
5.	Epithelial adhesion proteins	<i>E-cadherin</i>	Methylation of promoter region	Breast cancer, colorectal cancer, gastric cancer

CHAPTER 3

GAPS IN STUDY AND OBJECTIVES

Gaps in study

To the best of our knowledge, there have been no reports regarding Wnt pathway-related gene polymorphisms in lung cancer in North Indian population. Also, there have been no reports regarding the combined role of the Wnt antagonist genes polymorphism towards lung cancer risk. This study might provide inputs about how genetic variations in signalling pathways lead to differences in human susceptibility towards lung cancer risk and also factors such as smoking which might have a synergistic role in interaction with these polymorphic variants in these pathways would make it possible to answer questions like why certain groups of people have a higher incidence of lung cancer and others do not. Thus, the identification of new lung cancer susceptibility genes would yield new insight into lung carcinogenesis/tumorigenesis, and also help in screening particular population who would be at higher risk to develop lung cancer. This study also focuses on the prognostic relevance of the polymorphic sites of Wnt antagonist genes and their potential in predicting clinical outcome. Epigenetic silencing of several Wnt pathway genes has been reported in lung cancer. But the epigenetic silencing has not studied in the plasma of North Indian patients. This help in witnessing the relevance of circulating tumor DNA for predicting the same.

Objectives of the study

To study the single nucleotide polymorphic variants of the Wnt antagonist/signaling molecules by Polymerase chain Reaction-Restriction Fragment Length polymorphism (PCR-RFLP).

1. To perform methylation-specific PCR (MSP) for examining the methylation status of the Wnt antagonist/signaling genes *i.e.* Dickkopf (*DKKs*), Axis inhibition protein-2 (*Axin2*) and secreted frizzled-related protein (*sFRP*) genes in lung tumors.
2. Use of statistical analysis to measure the risk of the different genotypes of the Wnt and *Axin2* genes towards lung cancer development and also to analyze the co-relation between the methylation status of the *DKKs*, *Axin2* and *sFRP* genes with the SNPs of the Wnt genes in relation to clinico-pathological status of the patient. Furthermore we shall evaluate the role of smoking on the gene polymorphism of Wnt also to explore whether smoking plays any role in the promoter methylation of the above said genes.

CHAPTER 4

MATERIALS AND METHODS

4.1 Design of the study and study subjects

A case-control, retrospective and analytical study design was chosen to evaluate the north Indian population for lung cancer predisposition and its association with the genetic variants of Wnt antagonist genes. A total of 320 lung cancer patients were recruited for this study from the Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER) Chandigarh, India. This study has been reviewed and approved by the Institute ethics committee of PGIMER and therefore has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki and its later amendments. Ethical clearance letter has attached in the Appendix-I. Informed written consent was obtained from all participants or their representatives as attached in Appendix-II.

In detail, eligible cases included all patients who were newly diagnosed with primary lung cancer. Further, the histological classification of the patients into SQCC, ADCC and SCLC was also recorded on the basis of their diagnosis. There was no age, gender, smoking, histological, or TNM stage restriction. The exclusion criteria included the following considerations:

- Patients having prior history of any other carcinoma.
- Active infection or immunosuppression (HIV).
- Patients receiving any kind of systemic steroids.
- Patients suffering from chronic form of diarrhoea due to any cause.
- Patients who are not undergoing chemotherapy.

However, before recruiting the various physiological tests such as complete blood count which determines neutrophil count and other biochemical tests such as liver and renal function tests were performed. Also, radiological examination of the patient through chest X-ray and Contrast enhanced computed tomography (CT) scan of thorax and upper abdomen region was done.

Alongside, 340 unrelated individuals with no evidence of lung or any other cancer who visited the hospital for normal health check-ups were recruited as control group. An attempt was made to match each control on basis of age (± 10 years), sex and smoking parameters with respect to cases. A detailed questionnaire was filled for each case and control by a trained interviewer. The questionnaire comprised of major demographic details such as age, gender, region from where the subject belongs, smoking and drinking habits. Smokers reported tobacco habits such as smoking of cigarette and/or beedi (a native cigarette like stick of coarse tobacco hand-rolled in a dry tembuhurni leaf). As an indication of cumulative smoking exposure, pack-years were calculated by the following formula:

[(cigarettes or beedis per day/20) X years smoked]. The essential medical information of cases including histology, TMN classification, clinical staging, and performance status (KPS and ECOG) chemotherapeutic regimen and objective response in accordance to the Response Evaluation Criteria in Solid Tumor Group (RECIST) criteria was obtained from medical records of the hospital. The follow up data of the patients was obtained telephonically using the information of the patient as mentioned in the clinical records.

4.1.1 Chemotherapeutic Regimen

Doublet platinum- based chemotherapeutic drug combination was administered to all the lung cancer patients enrolled in the present study. This combination comprised of different drugs such as docetaxel, irinotecan or pemetrexed along with a platinum based drug such as cisplatin or carboplatin. These drugs were administered intravenously at a fixed dosage according to the surface area of the body. Where in 75mg/m² docetaxel, 500mg/m² pemetrexed or 75mg/m² irinotecan was given as 1 hr infusion which was then followed by cisplatin 70mg/m² infused for over 3 hrs. Standard folate and Vitamin B12 formulations were given to each patient as supplements. Certain drugs such as dexamethasone, granisetron and ranitidine were also administered as a part of premedication routine. Tumor response assessment was carried out after four cycles of chemotherapy. In certain cases, due to high toxicity and massive disease progression reported by radiological assessment chemotherapy procedure was stopped even before the completion of four cycles. The patients showing a good objective response were administered with additional two cycles of chemotherapy.

4.1.2 Follow up and response assessment

Study subjects were followed up for survival through telephonic conversations. Every two months until death or the end of the study period patients were called and inquired about health. Primarily we assessed overall survival and for that we considered survival time from the date of diagnosis till the date of last follow up or death due to other reasons. A total of 212 patients were enrolled for survival analysis, the STRAD diagram explains the selection criteria of the patients for overall survival analysis as shown in fig 4.1.

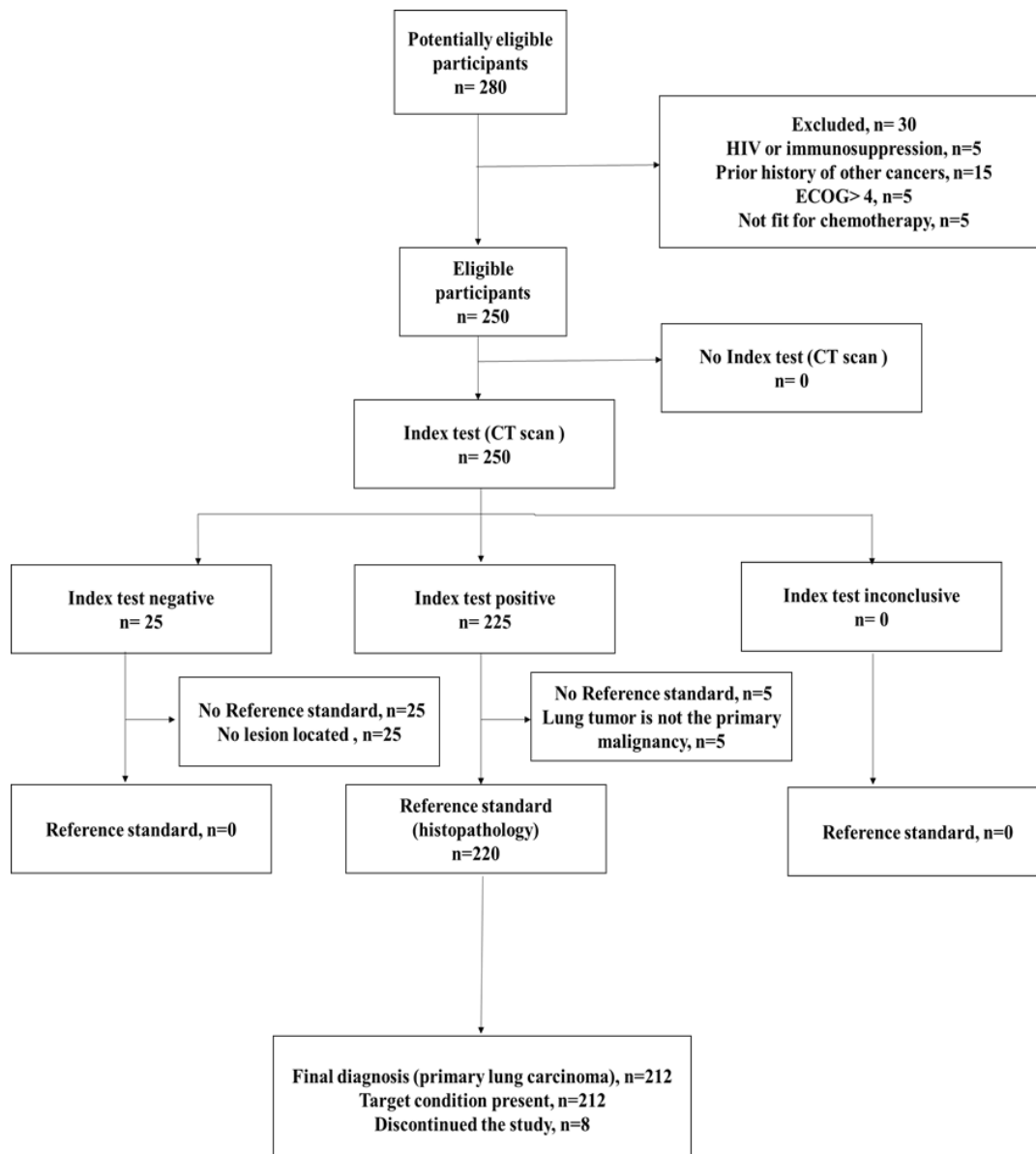


Fig 4.1: STRAD diagram explains the selection criteria of the patients for overall survival analysis

Tumor response was evaluated every 2 cycles of chemotherapy according to RECIST criteria. RECIST criteria characterizes the patients into four groups based on the target lesion after treatment namely complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD).

- CR: Complete disappearance of all target lesions.
- PR: At least a 30% decrease in the sum of diameters of target lesions, taking as reference the baseline sum diameter.
- PD: At least a 20% increase in the sum of diameters of target lesions, taking as reference the smallest sum on study (this includes the baseline sum if that is the smallest on study). In addition to the relative increase of 20%, the sum must also demonstrate an absolute increase

of at least 5 mm. (Note: the appearance of one or more new lesions is also considered progression).

- SD: Neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest sum diameters while on study [Eisenhauer *et al.*, 2009].

4.2 DNA Extraction

4.2.1 Genomic DNA Extraction

Genomic DNA was isolated from peripheral blood samples of cases and controls according to the protocol of Bartlett and Stirling [2003] with slight modifications.

- Firstly, the blood (3ml) was thrice treated with an equal volume of washing buffer containing 320mM sucrose, 1.5% Triton X-100, 20mM Tris-HCl (pH 8.0) and 5mM MgCl₂ in order to separate the red blood cells from the whole blood.
- Further, this RBC free pellet was lysed using the as obtained was lysed at 45⁰C overnight using an equal volume of WBC lysis buffer having 400mM Tris-HCl (pH 8.0), 150mM NaCl, 60mM EDTA, 1% SDS, and 100µg/ml proteinase K.
- Removal of proteins was done using phenol/chloroform/isoamyl alcohol (25:24:1) solution. The mixture was centrifuged at 8000 rpm for 7min at 4°C, to separate the aqueous phase into a new vial. This step was repeated twice and aqueous layer was separated.
- To the aqueous layer, an equal volume of chloroform/isoamyl alcohol (24:1) solution was added. The mixture was centrifuged at 6500 rpm for 5 min, to separate the aqueous layer into another vial.
- DNA was finally precipitated using the equal volume of chilled isopropanol which was obtained after centrifugation at 10,000rpm for 5min.
- The pellet was washed with 1ml of ice-cold 70% ethanol by centrifuging at 10,000 for 5 min. The pellet was then air-dried.
- Finally, it was resuspended in 200µl of sterile Tris-EDTA buffer.

4.2.2 DNA Extraction from Plasma

DNA extraction was also carried from the blood plasma of lung cancer patients for carrying out the methylation specific PCR. A slight modification in the protocol given by Stirling [2003] was done and the following protocol was used.

- 500µl of blood plasma is treated with an equal volume of solution containing 10% SDS and 5mg/ml of proteinase K. This mixture was incubated overnight at 55⁰C.
- In the next step, the removal of proteins was done using phenol/chloroform/isoamyl alcohol (25:24:1) solution. The mixture was vortexed for 30 seconds and centrifuged at 8000 rpm for

7min, to separate the aqueous phase into a new vial. This step was repeated twice and aqueous layer was separated.

- The aqueous layer is mixed with chloroform/isoamyl alcohol (24:1) solution. The mixture was centrifuged at 6500 rpm for 5 min, to separate the aqueous layer into another vial.
- An equal volume of chilled isopropanol was added to the aqueous layer which leads to the precipitation of plasma DNA and this solution is centrifuged at 10,000 rpm for 10 min to obtain the plasma DNA pellet.
- DNA pellet was washed with ice-cold 70% ethanol and centrifuged at 10,000 rpm for 5 min.
- The pellet is air dried and resuspended in 100 µl of Tris-EDTA buffer.

4.2.3 Qualitative and quantitative estimation of genomic DNA

The quality of the isolated DNA was assessed using agarose gel electrophoresis. 0.8% agarose gel with 0.5µg/ml ethidium bromide immersed in 0.5X TBE was used to visualize the isolated DNA sample. The sample was mixed with 6X loading dye and diluted with sterile water before loading in the wells. The gel was visualized under UV transilluminator showing the presence or absence of DNA [Johansson, 1972].

Quantity of DNA was determined by spectrophotometric analysis using Nanodrop (NanoDrop ND-1000 spectrophotometer). A single unit of absorbance (1OD) of double stranded DNA at 260 nm is equivalent to 50ng/µl of DNA. Henceforth, the amount of DNA in each sample is calculated by estimating its absorbance at 260 nm. Also, the purity of the sample was estimated by obtaining the ratio of absorbance at 260 nm to absorbance at 280 nm. For a pure DNA sample, the ratio should come out to be 1.8. A ratio less than 1.8 indicates protein contamination in the DNA sample and a ratio of more than 1.8 indicates the presence of RNA or phenol contaminants [Desjardins and Conklin, 2010]. After complete examination of concentration and purity, DNA was diluted to 100ng/µl and was stored at -20°C for further use as template.

4.3 Genotyping of polymorphic sites of *DKK*, *sFRP* and *Axin2* gene

4.3.1 Genotyping for *DKK* variants

Genotyping analysis was performed using PCR- RFLP technique using the primers, annealing temperature and restriction enzymes as previously reported by Hirata *et al.*, [2009] which are tabulated in table 4.1. A nested PCR approach was used for all the polymorphic sites. A nested PCR is used a detection tool as it leads a highly specific amplification of the target region of the gene. PCR mixture of 25µl consists of 1× PCR buffer, 1.5mM MgCl₂ with 0.5µM of both forward and reverse primer, 200µM of each dNTP's, 100µg/ml bovine serum albumin (BSA) and 1U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 200ng DNA. The amplicon obtained after using first set of primers is further diluted 100 times using TE and then it is used as a

template for the second set of primers which provides the desired amplicon. PCR conditions which were used for both the cycles of amplification is demonstrated in fig 4.2.

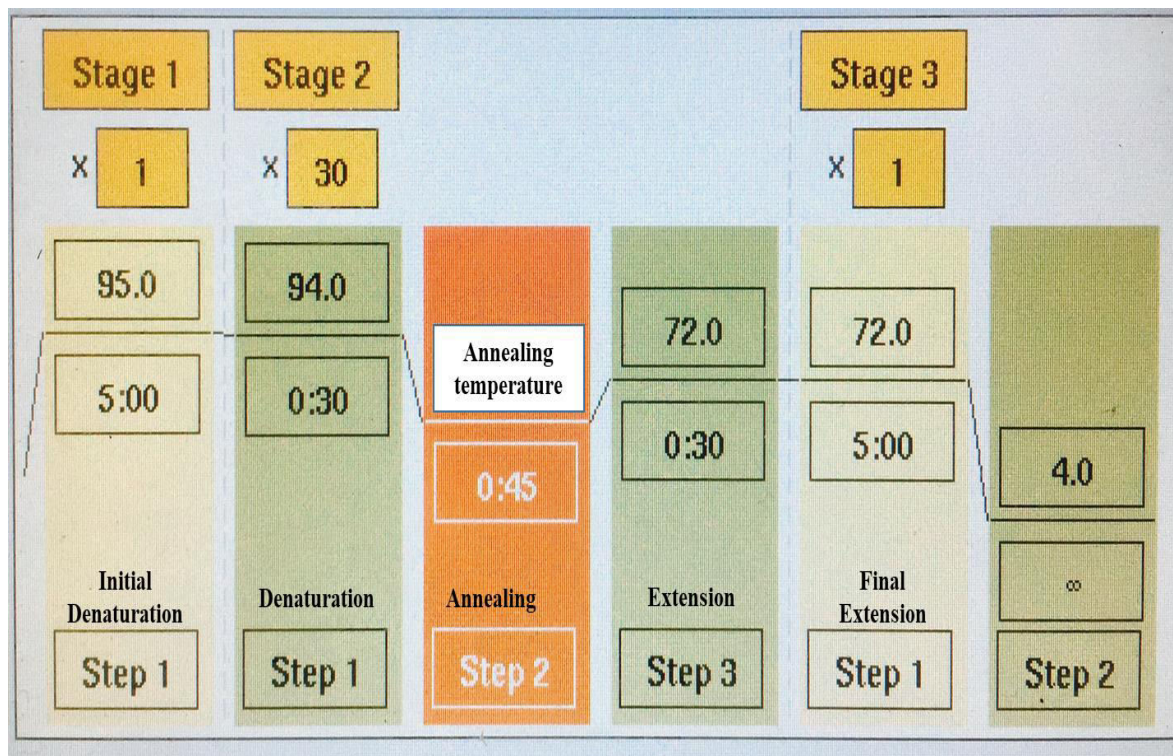


Fig 4.2: Schematic representation of thermal cycling conditions used; Stage 1: Initial Denaturation at 95°C for 5 min; Stage 2: 30 cycles of Denaturation at 94°C for 30 sec, Annealing at annealing temperature for 45 sec and Extension at 72°C for 30 sec; Stage 3: Final extension at 72°C for 5 min

Table 4.1: List of primers and restriction enzymes used for *DKK* genetic variants

Gene, SNP ID	Substitution	Primer sequence: forward and reverse	Annealing temperature	Product size (bp)	Restriction enzyme	Restriction pattern
<i>DKK4</i>, rs2073664 C/T	Synonymous	First: F5-GCC ATG GCA TTA CTG CTT TT-3; R5-ATT GCT GGT CAA TTG GCT TC-3	52°C		<i>EcoNI</i>	
	Exon 4 V ¹⁶⁹ V	Second: F5-CTG CGT GCT GTG TCT GTT TT-3; R5-AAC GCT GGA AGA TTT CTG GA-3	58°C	292		CC (223, 69bp), CT (292,223,69bp), TT(292 bp)
<i>DKK3</i>, rs2291599 C/T	Intron 4 C/T	First: F5-CAG AGG ACA TGG GGT GGA T-3; R5-CTG CTG CTC TGC TCT CCA TT-3	52°C		<i>AvaI</i>	
		Second: F5-CTT GTC CTC CAG GAG TCA GC-3; R5-CAT CAT CGA CGA GGA CTG TG-3	58°C	250		CC (120, 130bp),CT(250,120,130 bp),TT (250 bp)
<i>DKK3</i>, rs3206824, G/A	Nonsynonymous	First: F5-GAG GTC CCC GAT GAG TAT GA-3; R5-TAG GAA GAA GCC TGG TCA GC-3	52°C		<i>DdeI</i>	
	Exon7 Arg ³³⁵ Gly	Second: F5-GGT CCC CGA TGA GTA TGA AG-3; R5-AGC ACA CAC CTG GGG AAA TA-3	58°C	209		GG (209 bp), AG (209, 114, 97 bp), AA (114, 95 bp)
<i>DKK3</i>, rs7396187 G/C	Intron 4 G/C	First:F5-TTC CTT AGG TCC CTA GGT CCA-3; R5-AGG GCA AAG GAG ACT CTT CA-3	52°C		<i>Fnu4HI</i>	
		Second: F5-ACA GGG CAT GGC AGT TAG AG-3; R5-CTC TTC ACC CAA CAG GCA TT-3	58°C	245		GG (245 bp), GC (245,170, 75bp), CC (170, 75 bp)
<i>DKK2</i>, rs447372 A/G	Intron 1	First: F5-TGC TCA AAA AGC AGC ATC TC-3; R5-AGC TAA CGG TCA CTC AAC CAA-3	52°C		<i>EcoRV</i>	
		Second: F5-TCA AAA AGC AGC ATC TCA GG-3; R5-CTC CAA ATT TTG CCC TAC CA-3	58°C	249		AA (249 bp), AG (249, 119,130 bp), GG (119, 130 bp)
<i>DKK2</i>, rs419558 C/T	3'-UTR	First: F5- TACGGACACAGGACCTCACA -3; R5-TCC AAT ATA TGT GGG AAA AGA GC-3	52°C		<i>AflIII</i>	
		Second: F5-TCC CCT GGT TTC AAA GAT GA-3; R5-TGG GAA AAG AGC TAA CAG AGA GA-3	58°C	251		CC (251 bp), TT (121, 130 bp), CT (251, 121, 130 bp)
<i>DKK2</i>, rs17037102 G/A	Nonsynonymous	First: F5-TGG CTT CAT ATT TCA CAT CAA GA-3; R5-TGT GTG GTC TTC CTA GAT TCT GC-3	52°C		<i>DdeI</i>	
	Exon 3 Arg ¹⁴⁶ Glu	Second: F5-TGA TCA TCT CCA GGC ATC TG-3; R5-ATT CTG CCA TCC CAA GTC AT-3	58°C	132		GG (132 bp), GA (132,76,66bp), AA(76, 66 bp)

Restriction fragment length polymorphism

After the successful amplification, the PCR product was treated with the respective restriction enzymes which are obtained from New England Biolabs, Ipswich, MA, USA. 10µl of each PCR product was treated with 5 units of enzyme in the presence of 1X respective NEB buffer in a 20µl reaction. This reaction was incubated at the desired temperature which is 37°C for some and 65°C for others for 10-12 hours.

The digested products are resolved on 2.5% agarose gel dissolved in 0.5X TBE with 0.5µg/ml ethidium bromide. The electrophoresis is conducted in 0.5X TBE at 5 volts/cm. The digested samples are mixed with 2µl of 6X loading buffer containing sucrose, bromophenol blue and xylene cyanol. The resolved gel is visualized under the UV trans-illuminator and restriction pattern is observed [Rasmussen, 2012]. The gel documentation is done Gel Doc™ XR+ Gel Documentation System by Bio rad. Scoring of gels was done by two independent observers so to avoid any biasness.

This scoring done was according to the restriction pattern for every SNP which was based on the generation or removal of the restriction site in the presence or absence of the variant allele. The genotyping analysis was carried out without knowing the status of the sample whether it is a case or control. 15% samples were randomly genotyped twice to confirm the reproducibility of the results which was found to be 100%.

4.3.2 Genotyping for *Axin2* variants

The genotyping of the seven polymorphic variants of *Axin2* gene was carried out by PCR-RFLP. The primers and restriction enzymes used in this protocol have been previously detailed by Gunes *et al*, [2009]. The PCR protocol used for the amplification of these genotypic variants is as follows. The 25 µl reaction mixture comprised of 1X PCR buffer, 1.5mM MgCl₂, with 0.5µM of both forward and reverse primer, 200µM of each dNTP's, 100µg/ml bovine serum albumin (BSA) and 1U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 200ng DNA. The annealing temperature was optimized for each variant and their respective restriction enzymes have been mentioned in table 4.2. The cycling conditions used in this case are similar as shown in fig 4.1. The amplified products were digested with the respective restriction enzymes (5U) and were incubated at desired temperature overnight. RFLP products were resolved on the agarose gel electrophoresis/polyacrylamide gel electrophoresis depending on the difference in the size of fragments to be separated.

Polyacrylamide gel electrophoresis

- Polyacrylamide gels are used to separate nucleic acids and proteins in a solid support under the presence of electric field in a buffer. PAGE gel is a cross linked polymer of acrylamide and bis-acrylamide. Pore size is determined by the total concentration of monomeric units (acrylamide and bis-acrylamide) and ratio of acrylamide: bis acrylamide. Polymerization is initiated by ammonium per sulphate (APS) and catalysed by TEMED (N, N, N', N'-Tetramethylethylenediamine). Polyacrylamide gels have a rather small range of separation, but very high resolving power. In the case of DNA, polyacrylamide is used for separating fragments of less than about 500 bp. It is used to resolve fragments of DNA differing in length by a single base pair also. The gels were visualised using silver staining. This protocol used was given by Sambrook *et al.*, [2001] with slight modifications.
- Vinyl addition polymerization initiated by a free radical-generating system.
- TEMED accelerates the rate of formation of free radicals from persulfate and these in turn catalyze polymerization.
- The persulfate free radicals convert acrylamide monomers to free radicals which react with inactivated monomers to begin the polymerization chain reaction.
- The elongating polymer chains are randomly cross linked by bisacrylamide, resulting in a gel.
- The chemicals required include 30% Acrylamide solution [(29:1) acrylamide: bis-acrylamide], 5X TBE (Tris borate EDTA), 10% APS, TEMED and 6X loading buffer.

Silver staining

It is a highly sensitive technique, relies on the reduction of silver cations into insoluble silver metal by nucleic acids. Reduced silver grains deposit in the gel around DNA bands, creating a latent image. Image is developed to visibility by soaking the gel in solution of silver cations and reducing agent. Silver granules in latent image catalyze the further reduction and deposition of silver from the solution. Bands manifest as dark brown or black regions. Methanol and Glacial Acetic Acid are used as fixative for the DNA in gel. Formaldehyde acts as reducing agent (Reduces Ag^{2+} to Ag) and Sodium Carbonate is used as developing enhancer. The solutions which were used included:

- Fixative Solution- Water: Methanol: Glacial acetic acid (50:40:10).
- Staining solution-0.1gm of AgNO_3 in 100ml of deionised water followed by addition of 150 μl of 37% formaldehyde.
- Developing solution-3gm of Sodium Carbonate in 100ml deionised water and 150 μl of 37% formaldehyde and 20 μl of 10mg/ml sodium thiosulfate.

After putting the gel in fixative for 30 minutes, the gel was washed with sterile deionised water at least thrice and then, it was kept in staining solution for 30 min in dark. After staining, the gel was carefully washed with deionised water three times and then the developing solution was added and gel was kept on shaking. The bands were observed [Goldman and Merril, 1982].

Table 4.2: List of primers and restriction enzymes used for *Axin2* genetic variants

Gene, SNP ID <i>Axin2</i>	Substitution	Primer sequence: forward and reverse	Annealing temperature	Product (bp)	Restriction enzyme	Restriction pattern
rs2240308, 148 C/T	Exon1	F: CCA CGC CGA TTG CTG AGA GG R: TTC CGC CTG GTG TTG GAA GAG ACA T	60°C	242	<i>Mph1103I</i> C	C (218 +24) T (242)
rs2240308, 432 T/C	Exon1	F: CCT GGA GAG GGA GAA ATG C R: CAT CAC CGA CTG GAT CTC G	62°C	253	<i>Esp3I</i> T	T (253) C (145+108)
rs35285779, 956+16 A/G	Intron2	F: GGG TTT CTG GTG AGG GTC AGT R: CAA GCA AGC CCA CGG AAG G	57°C	228	<i>MlsI</i>	A (205+23) G (228)
rs9915936, 1365 G/A	Exon5	F: ACG TCT TCC CTT TCA GGA TG R: GGT ACT GCG AAT GGT GGT G	56°C	248	<i>MspI</i>	G(100+80+36+32) A(116+100+32)
rs1133683, 1386 C/T	Exon5	F: TGC GTA GGG AGC CGA ATG TTG R: GTG GTC CGG GGA GCG GGA TC	62°C	294	<i>TaqI</i>	C (294) T (274+20)
1712+19 G/T	Intron5	F: CAC CACCAC TAC ATC CAC CAC R: GCT CCC ACC TCA CAC CTG	55°C	249	<i>CpoI</i>	G (184+65) T (249)
2062 C/T	Exon7	F: TCC TTC TGT TTT CTC TCT GCT CAT TCC R: AGC GTG TTG GGT GGG GTC G	62°C	209	<i>TaqI</i>	C (209) T (191+18)

4.3.3 Genotyping for *sFRP* variants

Two SNP of the *sFRP3* gene (rs288326 and rs7775) were genotyped using the PCR-RFLP technique which was earlier reported by Shanguman *et al.*, [2007] For the other two SNPs of the *sFRP4* gene, the primer sequences and restriction enzymes were used in accordance to those which were previously detailed by Hirata *et al.*, [2009] where a nested-PCR approach was used to amplify the desired product. The primer sequences and the product size have been reported in table 4.3 along with the restriction enzymes and annealing temperature also. The rest of the protocol for PCR amplification and RFLP was similar as mentioned in case of *DKK* variants. PCR cycling conditions were same as illustrated in fig 4.2.

Table 4.3: List of primers and restriction enzymes used for *sFRP* genetic variants

Gene, SNP ID	Substitution	Primer sequence: forward and reverse	Annealing temperature	Product size (bp)	Restriction enzyme	Restriction pattern
<i>sFRP4</i> , rs1802073 C/A	Nonsynonymous	First: F5-GAGCACCATAAAGGGGTGAG-3; R5-GGGCACATGGCCTTACATAG-3	52°C			
	Exon 6 Pro ³²⁰ Thr	Second: F5-ACAGCGGAGAACAGTTCAGG-3; R5-TGGCCTTACATAGGCTGTCC-3	58°C	246	<i>HinfI</i>	CC (246bp), CA (246, 165,81bp), AA(165, 81bp)
<i>sFRP4</i> , rs1802074 G/A	Nonsynonymous	First: F5-AAGAGAGGCTGCAGGAACAG-3; R5-TCTGTACCAAAGGGCAAACC-3	52°C			
	Exon 6 Arg ³⁴⁰ Lys	Second: F5-AGAGCGGAGAACAGTTCAGG-3; R5-TGGCCTTACATAGGCTGTCC-3	58°C	246	<i>EaeI</i>	GG(246bp), GA (246, 135, 121bp), AA (135,121bp)
<i>sFRP3</i> , rs7775 C/G	Exon 6 Arg ³²⁴ Gly	F5-TGG TGA TAT GTG CTT GTG AC-3; R5-AGT CCA CTG TGT TAC TTT TTG TAT TTC GGG ATT TAG TCG C [†] -3	45°C	193	<i>BstUI</i>	CC (154, 39bp), CT (193,154,39 bp), TT(193 bp)
<i>sFRP3</i> , rs288326 C/T	Exon4 Arg ²⁰⁰ Trp	F5-TGT TCC ATT CGT TGA ATA TC-3; R5-ACA TCA TGG CAC TTA GTC TTT ATC TCT TTA ACT TTA GCT C [‡] -3	45°C	170	<i>TaqI</i>	CC (130, 40bp), CT (170,130,40 bp), TT(170 bp)

[†]The underlined nucleotide is a T as opposed to the C that exists in the *sFRP3* sequence. This forced change creates a *TaqI* restriction enzyme site in the C allele of the SNP, which is absent in the T allele.

[‡]The underlined nucleotide is a C as opposed to the T that exists in the *sFRP3* sequence. This forced change creates *BstUI* restriction enzyme site in the C allele of the SNP which is absent in the G allele.

4.4 Statistical analysis

Statistical analysis for Demographic parameters: Demographic variables such as age, gender, smoking status and pack years were classified into categorical and continuous variables. The continuous variables such as age and pack years were compared among cases and controls using t-test statistics and categorical variables such as gender and smoking status were analysed to find the difference between cases and controls using χ^2 test [Begg and Zhang, 1994].

Statistical analysis for Genotypic and Allelic frequencies: The Hardy–Weinberg equilibrium (HWE) theory ($p^2+2pq+q^2=1$; where p is the frequency of the wild-type allele and q is the frequency of the variant allele) was used for controls to calculate the genotype frequencies of all the polymorphisms using χ^2 test. The minor allele frequency was also calculated for each polymorphic site. Pearson's χ^2 test was used to compare the allelic and genotypic frequencies between cases and controls to find whether significant difference occurs between the two groups [Thakkestian *et al.*, 2005].

Association analysis: Unconditional Multivariate logistic regression was used to find out the adjusted odds ratio along with 95% Confidence Intervals (CI) associated with each genotype and various genotypic combinations. The odds ratio was adjusted for age, gender and smoking as these might act as confounding factors. All the variables used in the study were dichotomous except the continuous variables such as age. In the logistic regression analysis, the occurrence of lung cancer was taken as the dependent variable and genotype was selected as the independent variable which can predict the susceptibility for lung cancer along with various confounding factors such as age, sex and smoking status. The wild genotype was used as a reference group in case of each polymorphic site [Begg and Zhang, 1994; Karimi *et al.*, 2013].

Stratified analysis: The same approach was used for evaluating the risk associated with these polymorphic sites in the various sub groups formed on the basis of histology (SQCC, ADCC and SCLC), smoking status (smokers and non-smokers) and clinic- pathological parameters such as stage, primary tumor extension, lymph node invasion and metastasis. Also, the different genotypic combinations including the double and triple inter SNP combinations were evaluated using the same approach. All p values were two sided, and a p-value of <0.05 was considered statistically significant [Begg and Zhang, 1994]. All the statistical analyses were performed with Medcalc version 16.8.1 (Medcalc Software, Ostend, Belgium) and SPSS Version 20.0.

Linkage Disequilibrium and haplotype analysis: Haplotype and linkage disequilibrium (LD) analysis was carried out using the SHEsis software (<http://202.120.31.177/myAnalysis.php>) [Shi and He, 2009]. The polymorphic sites which share a common chromosomal loci and belong to the same gene were

evaluated for LD in order to obtain D' and r^2 values. Only those haplotype blocks were evaluated where the case and control frequency was more than 0.03, rest of the haplotypes were not considered for finding their correlation with lung cancer predisposition. Also, an evaluation was made to find out the association of different Hapblocks to confer their effect on lung cancer predisposition on the basis of the frequency of these haplotypes among cases and controls. The difference in the frequencies was calculated using global chi square test. The association was estimated by finding out the odds ratio for each haplotype.

False discovery rate analysis: Multiple comparison corrections were carried out for genotypic combinations (SNP- SNP) using FDR (False discovery rate) method by online software based on Benjamini and Hochberg approach (<http://sdmproject.com/utilities/?show=FDR>). The adjusted p values are mentioned and FDR adjusted $p < 0.1$ was considered as significant [Benjamini and Hochberg, 1995].

MDR: The major goal of the present study was to uncover the hidden interactions taking place in the presence of multiple factors which cannot be determined by traditional statistical methods. Nonparametric approach was applied by using Multifactor dimensionality reduction method in order to analyse various gene-gene and gene to environment interactions that contribute maximally to the scenario of lung cancer predisposition. MDR method involves the reduction of data to detect the multilocus genotypic combinations that can help in predicting the risk for a complex disease such as lung cancer. It reduces the multidimensional data into one dimension by pooling the genotypes into high and low risk groups. The interactions model are assessed on the basis of Cross validation Consistency (CVC), which is number of times a model is identified as the best model across the cross-validation sets. Higher CVC signifies the higher support for the strength of the model. Average prediction error is also calculated (1- testing accuracy). Other important parameter is permutation testing (p -value) which determines the significance of the hypothesis generated [Ritchie and Motsinger, 2005]. The software used to apply MDR is version 0.5.1 of the open-source MDR software package that is available online (<http://www.epistasis.org>).

CART analysis: Another high order interaction between all the polymorphic variants was analyzed using CART (classification and regression tree analysis). CART was applied using the CART software (6.0, Salford Systems, San Diego, California). The CART software uses the data mining approach to create a decision tree which depicts the interaction of the all the variables under study through a binary recursive partitioning method to identify sub-groups at higher risk, which are not easily revealed by the traditional logistic regression models. The analysis is conducted in such a manner where the most significant predictor is used to split the sample into subgroups and continues until the differences are not significant. The tree obtained has various nodes and sub nodes depicting their role in modulating the disease

susceptibility. The node with a lower case rate is used as a reference to find out adjusted odds ratio and 95% C.I for all other genotypes depicted in other nodes [Gupta *et al.*, 2016].

Overall survival and Hazards ratio analysis: For finding out the role of these polymorphisms in predicting the survival of the lung cancer patients, firstly the survival time (in months) of each patient was calculated. Survival time can be also be censored when there is a follow-up time but the event which is death of the patient in this case has not yet occurred or is not known to have occurred. The concept of Kaplan Meier was used to find out the Median survival time (MST) and survival curve for each category in all the polymorphic sites. Univariate analysis called log rank test was used as a statistical method to compare the MST of patients based on the genotype of that particular variant. In order to whether these polymorphic variants can act as independent predictor variable, multivariate Cox proportional Hazard was used where Hazard ratio along with 95% C.I adjusted with various covariates such as age, sex, smoking, stage, KPS, ECOG and histology was obtained. Cox's proportional hazards model is analogous to a multiple regression model and enables the difference between survival times of particular groups of patients to be tested while allowing for other factors [Mahimkar *et al.*, 2012]. In this model, the response (dependent) variable is the 'hazard'. The hazard is the probability of dying that patients have survived up to a given point in time or the risk for death at that moment. Kaplan-Meier and Cox Regression method was also used in various subgroups stratified on the basis of sex, smoking, histology and performance status [Efron, 1988]. All the statistical analysis was performed with Medcalc version 16.8.1 (Medcalc Software, Ostend, Belgium).

STREE analysis: Further, high order gene interactions were taken into account by using the recursive partitioning approach. This approach produces the initial spilt taking into consideration the primary factor affecting the OS of patients. The software used for making the decision tree is STREE program (<http://c2s2.yale.edu/software/stree>). Log rank method was employed as the splitting criteria. The recursive method keeps on producing the offspring nodes till statistically insignificant spilt occurs. The resulting tree was binary, and each terminal node represented a subgroup of patients carrying a particular genotypic combination and hence shows a varied survival time and prognosis. HRs and their 95% CIs for each terminal node were calculated using the multivariate Cox proportional hazard model adjusted for sex, age, histology, KPS, ECOG, smoking and stage of tumor.

4.5 Methylation-Specific Polymerase Chain Reaction (MSP) Analysis for *DKK*, *sFRP* and *Axin2* promoter region

The isolated plasma DNA is further modified using Bisulphite conversion method. This technique involves treating methylated DNA with bisulfite, which converts unmethylated cytosines into uracil.

Methylated cytosines remain unchanged during the treatment. Once converted, the methylation profile of the DNA can be determined by PCR. This bisulphite conversion and preparation of the DNA sample for the amplification and detection of methylation profile is a three step process. The process was carried out by using EZ methylation kit (D5001) which simplifies this process of converting the unmethylated cytosines into uracil. The product's innovative in-column desulphonation technology eliminates otherwise cumbersome precipitations. The kit is designed to reduce template degradation, minimize DNA loss during treatment and cleanup, while ensuring complete conversion of the DNA. Purified, converted DNA is ideal for PCR amplification. The reagents used during this conversion approach are:

- CT conversion reagent: CT conversion reagent was dissolved in 750 µl of water and 210µl of M-dilution buffer. This reagent was freshly prepared prior to use and it could be stored overnight at 4°C or for a month at -20°C.
- M-wash buffer: This buffer was reconstituted by adding 24ml of 100% ethanol to the already provided 6ml solution.
- M-dilution buffer
- M-desulphonation buffer
- M-elution buffer
- M-binding buffer

The detailed procedure followed during the conversion of DNA has several steps as mentioned below:

1. 5µl of M-dilution buffer was added to the DNA sample (2µg) and the final volume was adjusted to 50µl with water. The sample was mixed by either flickering or pipetting up and down.
2. This sample was incubated at 37°C.
3. After this, the sample was mixed with 100 µl of the prepared CT conversion reagent.
4. It was kept in incubation at 50°C for 12-16 hours in dark. After, the completion of this incubation the samples were transferred to 4°C for 10 minutes.
5. 400 µl of M-binding buffer was added to a Zymo-Spin™ IC column and it was placed into a provided collection tube.
6. Now, the sample (in step 4) was loaded into the Zymo-Spin™ IC column containing the M-binding buffer. The caps of the column were closed and they were inverted several times to facilitate mixing.
7. Then, the columns were centrifuged at full speed (>10,000 x g) for 30 seconds. The flow through in the collection tube was discarded.

8. Now, 100 µl of M-wash buffer was added to the column and was again centrifuged at full speed for 30 seconds.
9. 200 µl of M-desulphonation buffer was added to the column and it was kept at room temperature for 15-20 minutes. After, incubation centrifugation was done at full speed for 30 seconds.
10. Then, 200 µl of M-wash buffer was added to the column and it was centrifuged at full speed for 30 seconds. Another 200 µl of M-wash buffer was added to the column and it was centrifuged at full speed for 30 seconds.
11. 1.5 ml microcentrifuge tube was placed below the column and 10 µl of M-elution buffer was directly added to this column matrix. It was then centrifuged for 30 seconds at full speed in order to elute the DNA. The concentration of eluted DNA came out to be 200ng/µl.

The modified DNA was used for Methylation-specific PCR, promoter-associated CpG Island was amplified using primers pairs that specifically amplify either methylated or unmethylated sequences of the *DKK*, *sFRP* and *Axin2* promoter regions. DNA from healthy donors shall be used as a negative control, and a Methyl-amp Universal Methylated DNA Kit can be used as a positive control for methylation-specific assays.

For the *DKK2* gene, promoter methylation was analyzed using primers as previously described by Suzuki *et al.*, [2008]. Briefly, both methylated primers and unmethylated primers corresponding to the promoter region of *DKK2* gene were used. *DKK3* gene promoter methylation was analyzed by using both methylated and unmethylated primers as described by Veeck *et al.*, [2009] for amplifying a region between 1529 and 1917 that specifically corresponds to the promoter of the *DKK3* gene. Secreted Frizzled related protein 1 (*sFRP1*) methylation was studied as described by Fukui *et al.*, [2005]. *Axin2* gene promoter methylation will be analyzed using primers as previously described by Tseng *et al.*, [2008]. The list of the primers along with the length of amplified methylated and unmethylated product for each gene and their annealing temperatures is summarized in table 4.4.

A hot start PCR approach was used for detecting the methylation status of these genes. 25 µl reaction mixture comprised of 1X PCR buffer, 1.5mM MgCl₂, with 0.5µM of both forward and reverse primer, 200µM of each dNTP's, 100µg/ml bovine serum albumin (BSA) and 1.5U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 200ng DNA. DNA polymerase was added after initial denaturation. The PCR conditions included an initial melting step of 94°C for 5 min and followed by 35 cycles of 94°C for 45s, variable annealing temperature for 30s and 72°C for 45s and final extension at 72°C for 7 min for all the four genes. MSP products will be analyzed by electrophoresis on 2% agarose gel.

Table 4.4: List of primers used for Methylation specific PCR

Gene	Promoter region	Primer sequence: forward and reverse	Annealing temperature	Product (bp)
<i>sFRP1</i>	Methylated	F:5-TGTAGT TTT CGG AGT TAG TGT CGC GC-3 R: 5-CCTACGATC GAA AAC GAC GCG AAC G-3	60°C	126
	Unmethylated	F:5-GTTTTGTAG TTT TTGGAGTTA GTG TTG TGT-3 R:5-CTCAACCTACAATCAAAAACA ACA CAA ACA-3	60°C	135
<i>DKK2</i>	Methylated	F:5-AGAGTTAAATCGTCGAGATTTC-3 R:5-CTAAAAACAATCAAATACGAAACG-3	60°C	146
	Unmethylated	F:5-GGAGAGTTAAATTGTTGAGATTTT-3 R:5-ACTAAAAACAATCAAATACAAAACA-3	58°C	149
<i>DKK3</i>	Methylated	F:5- GGGCGGGCGGCGGGGC-3 R:5-ACATCTCCGCTCTACGCCCCG-3	58°C	120
	Unmethylated	F:5-TTAGGGGTGGGTGGTGGGGT-3 R:5-CTACATCTCCACTCTACACCCA-3	58°C	126
<i>Axin2</i>	Methylated	F:5- GGGCGCGTAGCGAATAT-3 R:5- ACGTACCCTTCCCCGACAT-3	55°C	202
	Unmethylated	F:5-GGTAGGGGGTGTGTAGTGAA-3 F:5-ACATACCCTTCCCCAACATC-3	55°C	208

CHAPTER-5

RESULTS

5.1 Genomic DNA isolation

Qualitative estimation of genomic DNA was carried out by running 0.8% agarose gel. The bands were observed to examine the integrity and shearing of the genomic DNA. The representative example of a good quality intact genomic DNA is illustrated in fig 5.1. Similarly, the DNA isolated from the blood plasma of lung cancer patients was visualized by agarose gel electrophoresis as shown in fig 5.2 and its qualitative evaluation was carried out.



Fig 5.1: Representative example of agarose gel electrophoresis showing genomic DNA bands

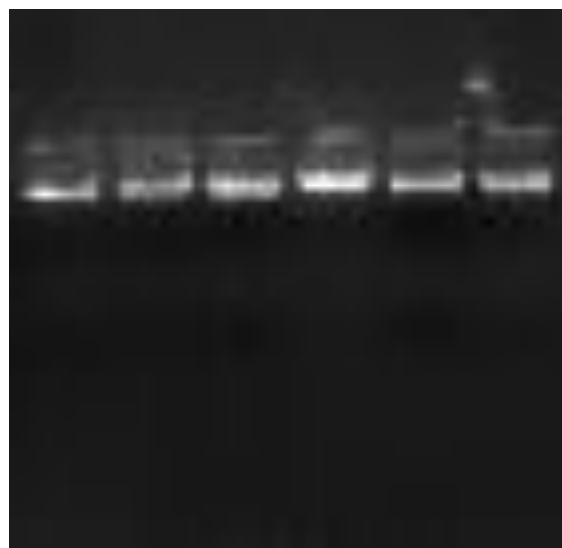


Fig 5.2: Representative example of agarose gel electrophoresis showing plasma DNA bands

5.2 PCR amplification

PCR amplification for each SNP was confirmed by 1.7% agarose gel electrophoresis to ensure successful amplification of the desired length which was confirmed using marker DNA of either 100 bp or 50 bp size. A representative example of agarose gel electrophoresis is shown in fig 5.3, demonstrating the amplification of genomic DNA for *Axin2* 1365 G/A. An amplified product of 248 bp is obtained as indicated by the 100 bp ladder run along.

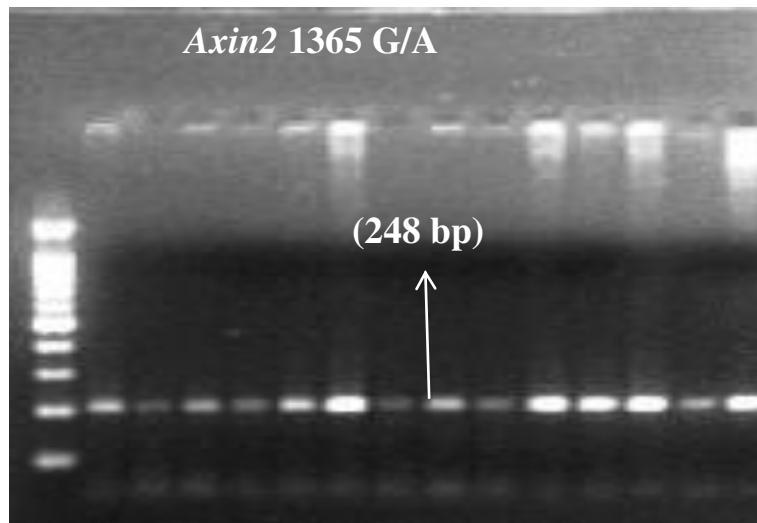


Fig 5.3: Representative example of agarose gel electrophoresis showing PCR amplification for *Axin2* 1365 G/A

5.3 Genotyping by RFLP

5.3.1 Genotyping for *DKK* genetic variants

The agarose gel pictures elucidating the PCR amplified product and restriction pattern of all the *DKK* variants are shown below in fig 5.4. The SNP of *DKK3* gene analyzed was rs3206824 G/A which is located in the exonic region of the gene; this polymorphism leads to a non-synonymous Arg³³⁵Gly substitution in exon 7. The amplified PCR product is of 209 bp in size. The restriction enzyme used for genotyping is *DdeI*, the wild genotype gives a band of GG (209 bp), AG (209, 114, 97 bp), AA (114, 95 bp) as depicted in the fig 5.4 (D). The other polymorphic site of *DKK3* was also studied, which is, rs2291599 C/T. this lies in the intron 4 of the *DKK3* gene. The size of the PCR amplicon is 250 bp, the restriction analysis was carried out by the enzyme *AvaI*. The wild genotype CC (120, 130bp) whereas mutant type remains uncut and heterozygous form gives 250,120,130 bp bands as shown in fig 5.4 (E). Wnt antagonist SNP which we have worked on is the *DKK4* rs2073664 SNP. This SNP is located in exon4 of the *DKK4* gene and results in C→T [V¹⁶⁹V] substitution. After PCR, a 292bp amplicon was subjected to restriction digestion by *EcoNI* enzyme. The wild genotype C/C was cut into two products of 223, 69bp, the mutant genotype T/T remained uncut (292bp) and the heterozygous had three bands of 292, 223, 69bp. The *DKK2*, rs447372 SNP was also studied. This SNP is located in the intron1 of the *DKK2* gene and results in A→G substitution. Nested PCR approach was used to amplify the product of 249 bp, for carrying out the genotyping the amplified product was digested with *EcorV* enzyme giving a pattern where AA (249 bp), AG (249, 119,130 bp) and GG (119, 130 bp). The next SNP of *DKK2* which was studied is rs419558 C/T, where the amplified product of 251 bp was digested by *AflIII* and gave this

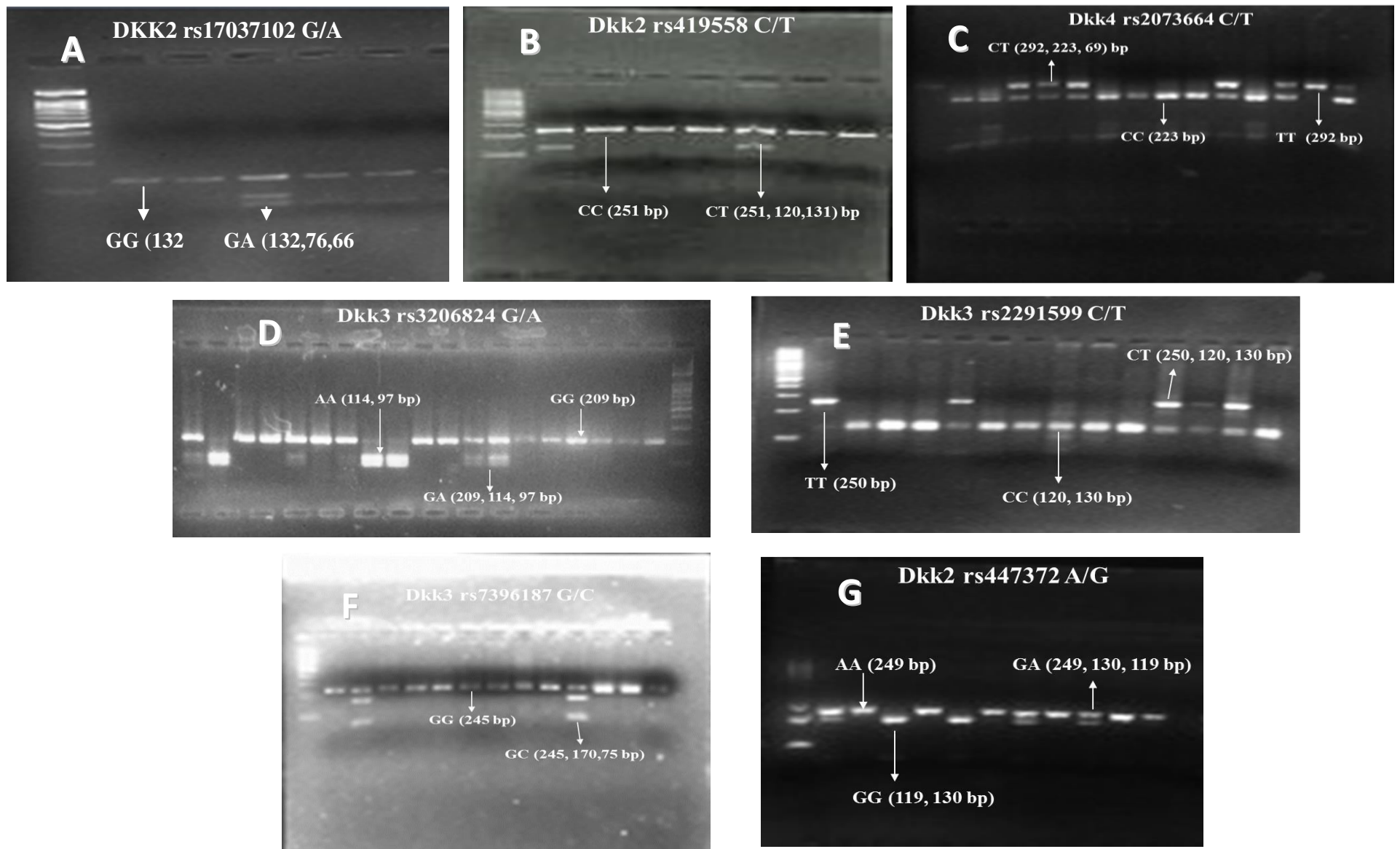


Fig 5.4: Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for *DKK* variants; A: *DKK2* rs17037102 G/A, B: *DKK2* rs419558 C/T, C: *DKK4* rs2073664 C/T, D: *DKK3* rs3206824 G/A, E: *DKK3* rs2291599 C/T, F: *DKK2* rs447372 A/G

pattern: CC (251 bp), TT (121, 130 bp), CT (251, 121, 130 bp). The last SNP studied is *DKK2* rs17037102 G/A, the pattern observed was GG (76,66bp), GA (132, 76,66bp), AA (76, 66 bp)

5.3.2 Genotyping for *Axin2* genetic variants

The first is the 148 C→T rs2240308 SNP at exon 1 of the *Axin2* gene. The amplified product is a 242bp in size, and which is then cut by *Mph1103IC* restriction enzyme. The wild type genotype is cut into 218 and 24 bp. The mutant genotype is uncut and the heterozygous genotype has 242, 218 and 24 bp product which was analyzed on 6% PAGE gel and then visualized by silver staining as shown in fig 5.5 (C). Another SNP (rs1133683) of the *Axin2* gene in exon5 at position 1386 where there is G to C substitution. The amplified product is a 294bp product which creates a C substitution instead of G and hence is cut by *TaqI* enzyme. The normal GG genotype is uncut whereas the mutant genotype is cut in two fragments of 274 and 20 bp. Whereas heterozygous has three bands of 294, 274 and 20bp which is visualized in PAGE as illustrated in fig 5.5 (B). Next SNP is in the exon7 and results in C→T substitution at position 2062. The amplified product is a 209bp product which creates a C substitution instead of T and hence is cut by *TaqI* enzyme. The normal CC genotype is uncut whereas the mutant genotype is cut in two fragments of 191 and 18 bp. Whereas heterozygous has three bands of 209, 191 and 18bp which is visualized by silver staining as demonstrated in fig 5.5 (A). We have also analyzed the *Axin2* gene *i.e.* rs9915936 which is in exon5 of *Axin2* gene and results in 1365 G→A substitution leading to Pro⁴⁵⁵Pro. The amplified product is of 248 bp, where GG wild type gives band of 100,80,36,32 bp and mutant genotype AA gives 116, 100, 32 bp band pattern on digestion with *MspI* as shown in fig 5.5 (E). We have also analyzed the *Axin2* 1712+19 G/T SNP which is in intron 5 of *Axin2* gene and results in G to T substitution. The amplified product is of 249 bp which was digested with *RsrII* isochizomer of *CpoI*. The wild genotype GG gives band of 184, 65 bp whereas the mutant TT genotype gives a 249 bp band and heterozygote GT gives three band of 249, 184 and 65 bp as illustrated in fig 5.5 (G). For *Axin2* rs35285779 (956+16 A/G) which is located in the intron 2 region. The amplified product is of 228 bp. The restriction pattern is wild type (205, 23 bp), mutant type (228 bp) hetero type (228, 205, 23 bp) which was visualized on agarose gel as shown in fig 5.5 (D). *Axin2* 432 T/C was also analyzed where the amplified product is of 253 bp which was digested with *Esp3I*. 2.5% Agarose is used to separate the bands. The restriction pattern followed by wild type (253 bp), mutant type (145, 108 bp) and hetero type (253, 145, 108 bp). The gel pictures have been elucidated in fig 5.5.

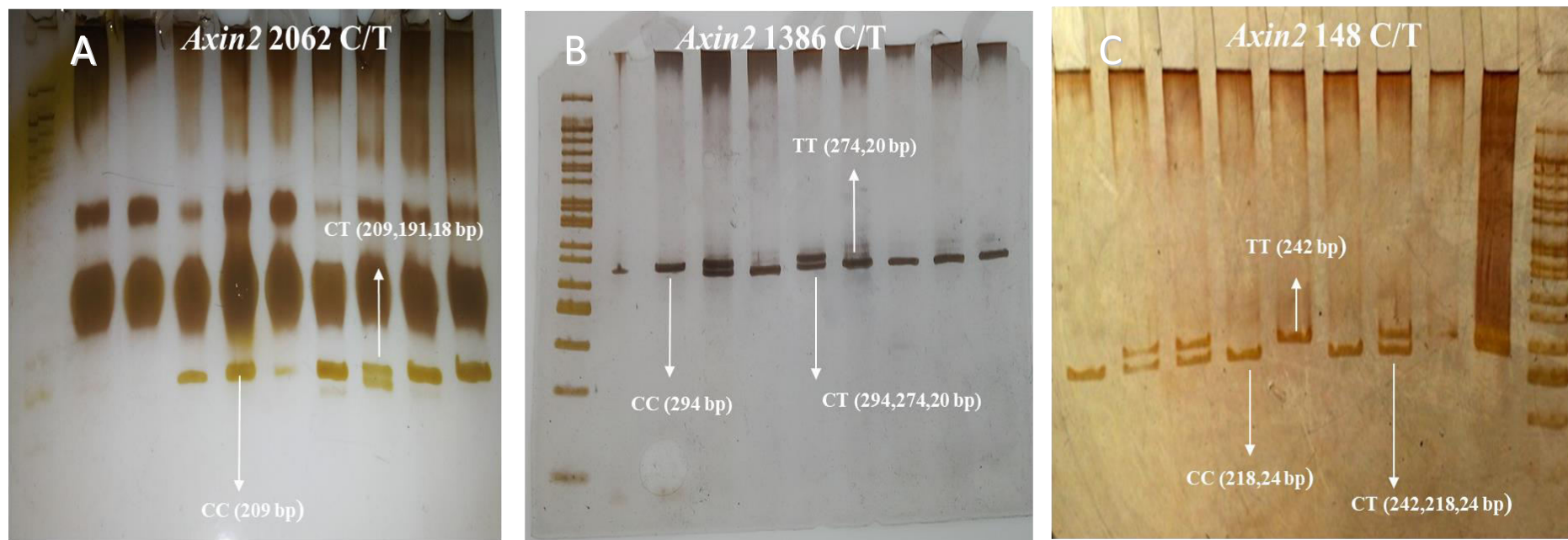


Fig 5.5: Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for *Axin2* variants; A: *Axin2* 2062 C/T, B: *Axin2* 1386 C/T, C: *Axin2* 148 C/T

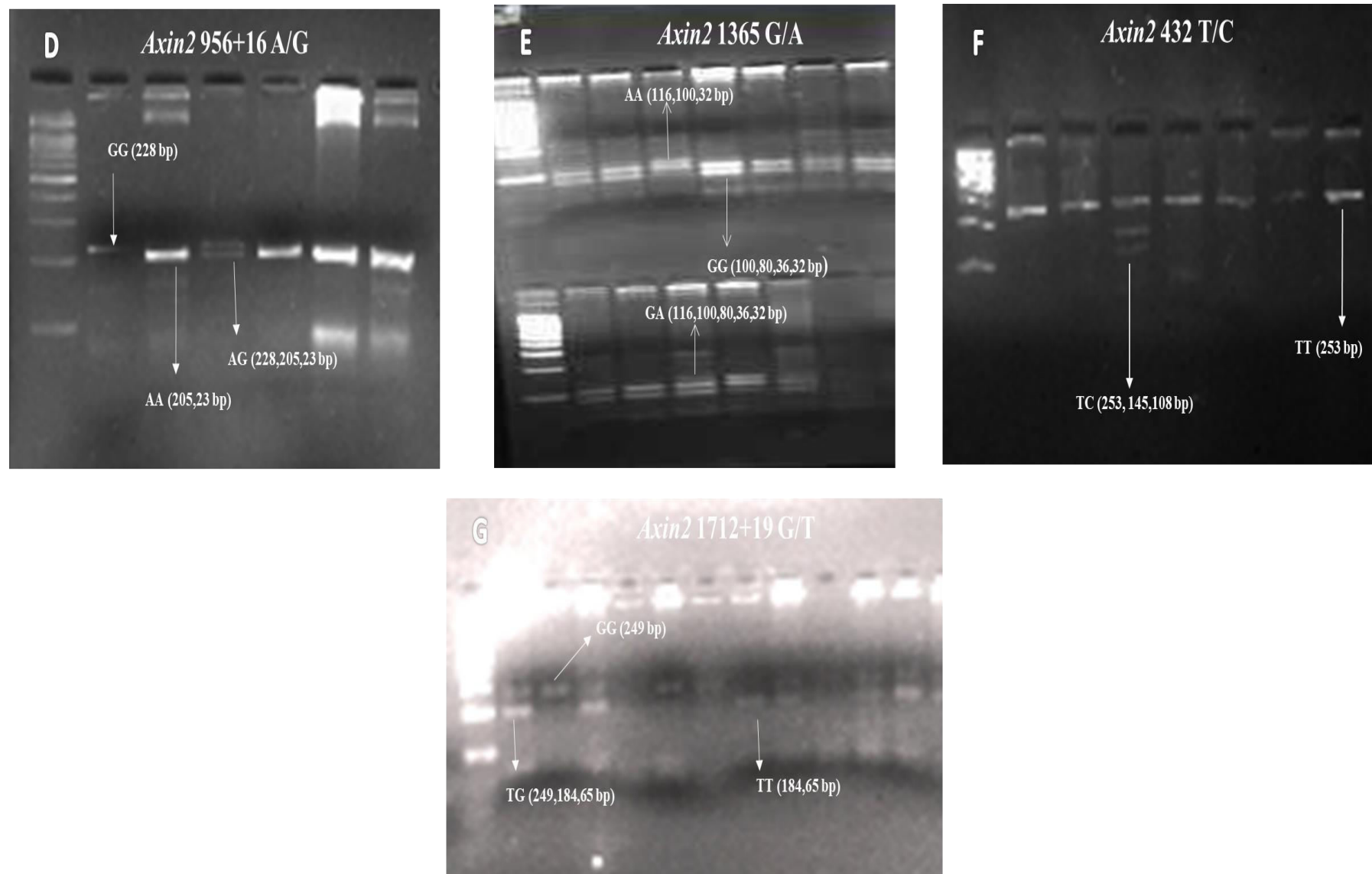


Fig 5.5: Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for *Axin2* variants; D: *Axin2* 956+16 A/G, E: *Axin2* 1365 G/A.

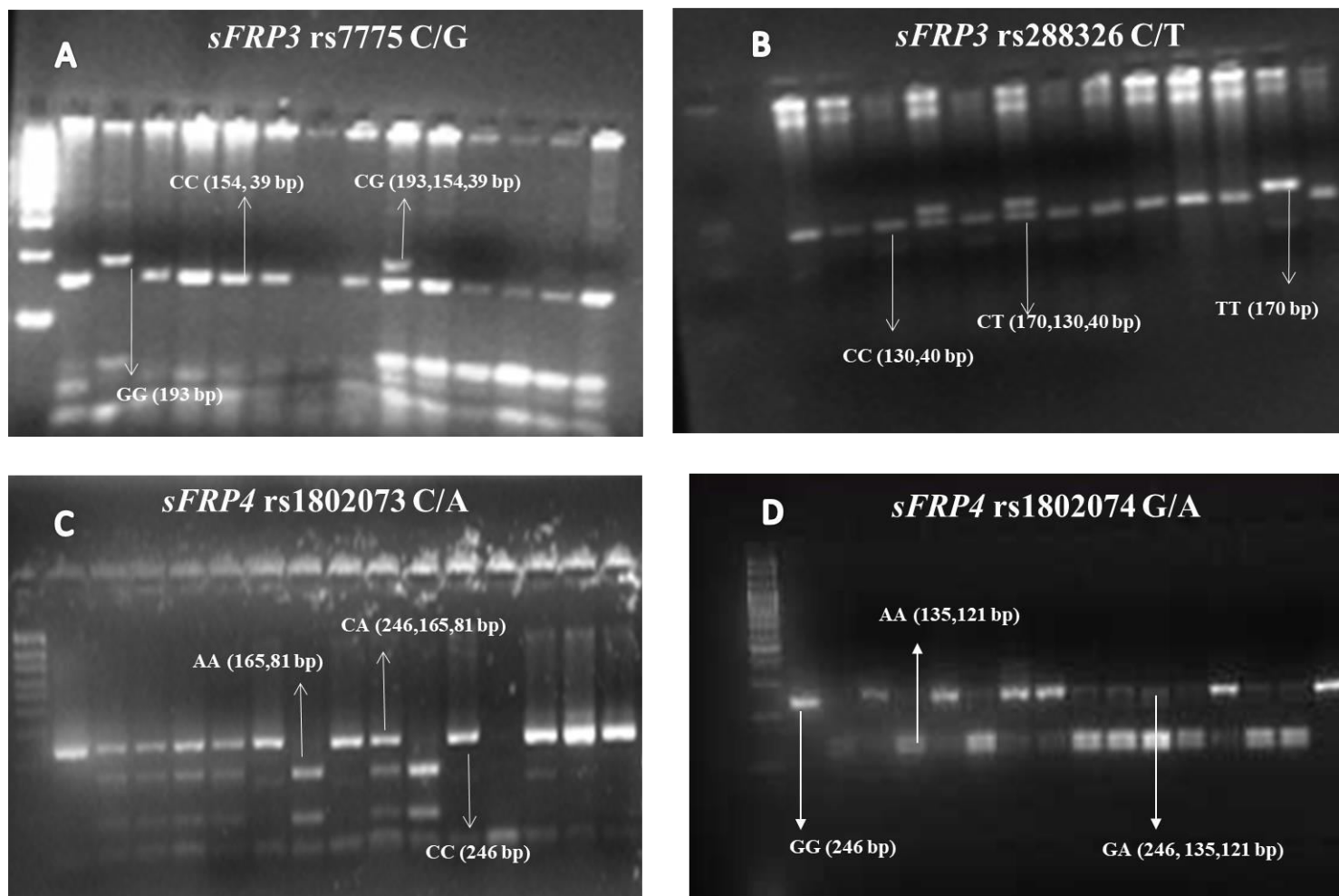


Fig 5.6: Representative example of agarose gel electrophoresis showing PCR-RFLP pattern for *sFRP* variants; A: *sFRP3* rs7775 C/G, B: *sFRP3* rs288326 C/T, C: *sFRP4* rs1802073 C/A, D: *sFRP4* rs1802074 G/A

5.3.3 Genotyping for *sFRP* genetic variants

For *sFRP3* rs288326 C/T, C allele generates two fragments of 130 bp and 40 bp, T allele remains uncut. Similarly for *sFRP3* rs7775 C/G, C allele generates two fragments of 154 bp and 39 bp, G allele remains uncut. In case of *sFRP4* rs1802073 C/A, the wild genotype CC was uncut and the mutant genotype AA gave two bands of 165,81bp and the heterozygous had three bands of 246, 165 and 81bp. In the similar manner, the genotyping for *sFRP4* rs1802074 G/A was also done, where presence of heterozygote was identified by the presence of three bands *i.e.* 246, 135 and 121bp. The representative examples of the RFLP patterns of these four SNPs have been elucidated in fig 5.6.

5.4 Epidemiological statistics of North Indian population under study

5.4.1 Regional Distribution of Lung cancer patients

The collection of blood samples was carried out from lung cancer patients visiting the Department of Pulmonary Medicine, PGIMER Chandigarh. PGIMER is a centrally located specialized government hospital where patients belonging to the different regions of North India are referred by local hospitals. The major regions included Punjab, Haryana, Himachal Pradesh, Uttar Pradesh, Uttarakhand, Jammu & Kashmir and Chandigarh itself. A total of 320 patients were enrolled for the present study, out of which the maximum proportion of patients belonged to Punjab (26%), Himachal Pradesh (23.1%) and Haryana (22.8%) as shown in fig 5.7. The remaining lung cancer patients enrolled for the study belonged to regions such as Uttar Pradesh (11.5%), Uttarakhand (3.7%) and Jammu & Kashmir (1.2%).

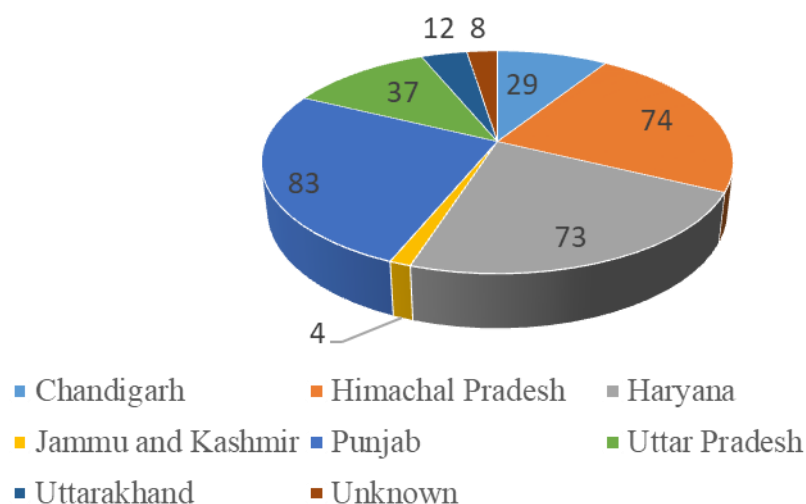


Fig 5.7: Pie chart illustrating regional distribution of North Indian lung cancer patients

The subjects were also analyzed for their regional distribution on the basis of gender. The study population comprises of 276 (86.2%) males and 44 (13.8%) females in cases. Out of 340 control subjects, 282 (82.9%) were recorded as males and rest 58 (17.1%) as females. The diagrammatic representation of regional distributional of males and female study population is shown in fig 5.8. The maximum proportion of males and females belonged to Punjab and Himachal Pradesh respectively. The percentage of males from Haryana and Punjab was same (23.2%). The proportion of females from Punjab and Haryana was similar (20.4%).

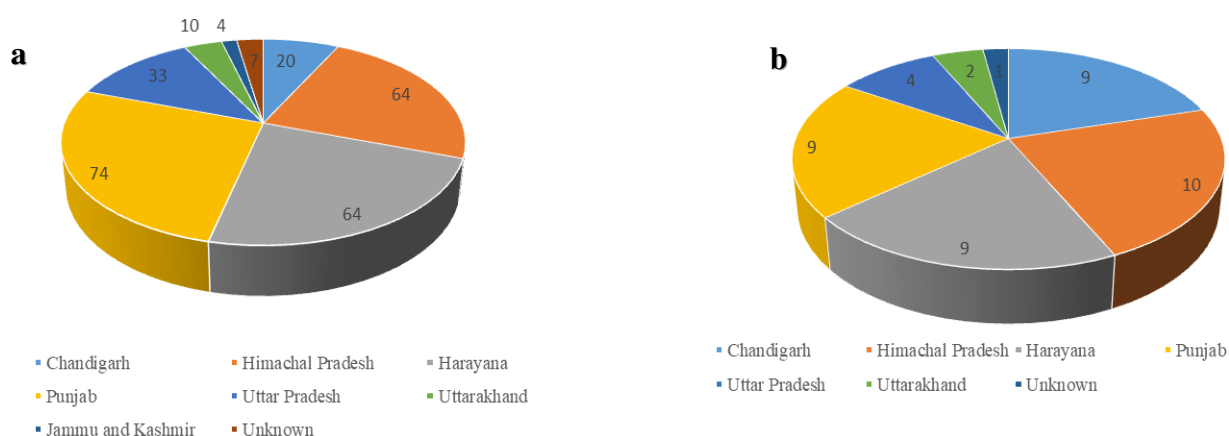


Fig 5.8: Pie chart illustrating regional distribution of North Indian Lung cancer patients
a: regional distribution of males; b: regional distribution of females.

5.4.2 Distribution of study subjects on the basis of smoking:

One of the most important epidemiological data which was collected for all the cases and controls was the information about their smoking status. The study population was characterized as smokers and non-smokers on the basis of their smoking habit. As smoking is an established causative agent for lung cancer, henceforth the proportion of smokers was higher as compared to non-smokers among cases. Similarly, in order to avoid sample bias an attempt was made to match the controls in terms of smoking status. The percentage of smokers among cases and controls was 78% and 69% respectively as shown in fig 5.9. The quantitative estimation of smoking was done by calculating the pack years for each study subject. The mean pack years was higher in the case group when compared to the control group (27.6 ± 33.6 vs 19.4 ± 19.72). This clearly indicates the prevalence of smokers is much higher in cases as compared to controls.

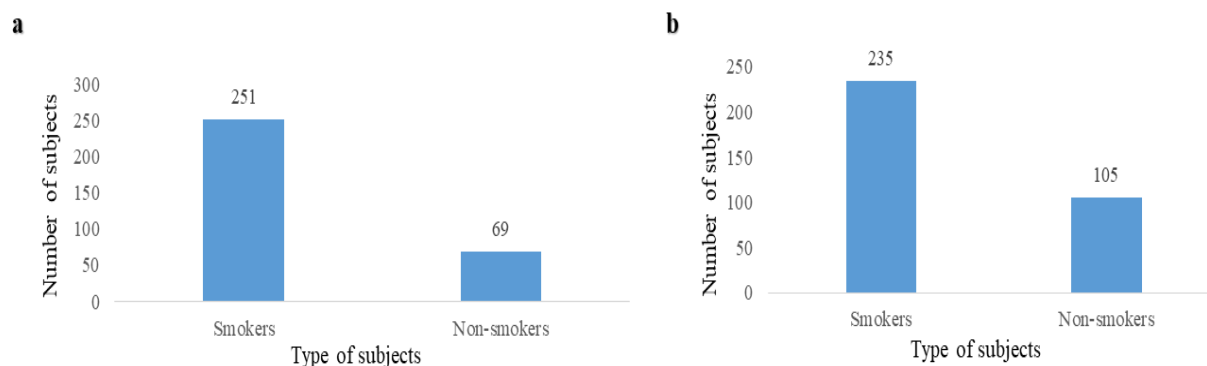


Fig 5.9: Graph showing the distribution of study subjects on the basis of smoking status. a: Among cases, b: Among controls

Further, the detailed information about the type of smoking was also recorded. This data elucidates that the different types of smoking carried out by study subjects included beedi, cigarette, and a combined intake of beedi and cigarette both. However, the major proportion of the study subjects were reported to be exposed with beedi in both cases (86%) and controls (85.5%). The smoking pattern of the study subjects was also analysed on the basis of their gender. The percentage of smokers among males in cases (95.2%) is higher as compared to those in control group (82.9%) as shown in fig 5.10.

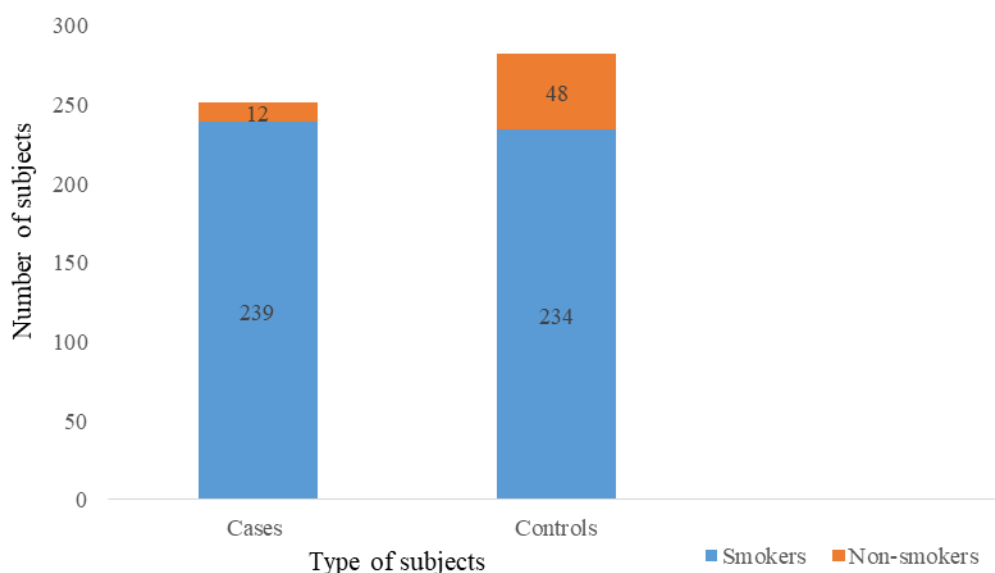


Fig 5.10: Graph showing the distribution of male study subjects on the basis of smoking status.

5.4.3 Distribution of Study subjects on the basis of age

Age is another important demographic parameter, which has been correlated with the onset of many diseases. The age of all study subjects was recorded among both cases and controls. The mean age of cases was estimated as 57.7 ± 10.81 years with a range of 29-86 years. The median age was estimated as 59 years. A sincere attempt was made to match the age of cases and controls with (± 10 years) difference. The mean age calculated for control subjects was found to be 53.05 ± 10.22 years with a range of 23-83 years and the median age was found to be 53 years. The diagrammatic representation of distribution of study subjects on the basis of their gender is shown in fig 5.11. A slightly higher percentage of males were reported to have age ≤ 59 years as compared to those having age less than 59 years (82.8% vs 89.5%) among lung cancer patients. Similarly, 10.1 % females were found to have age ≤ 59 years and 17.1% were reported with age less than 59 years.

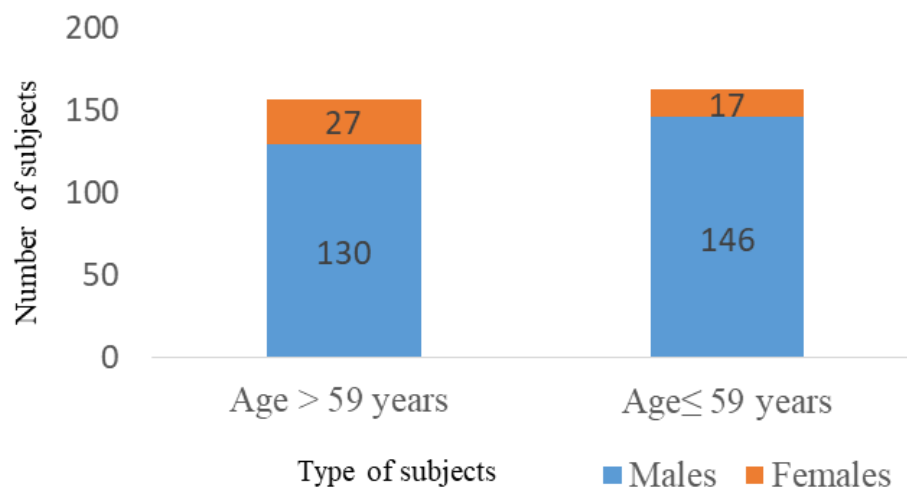


Fig 5.11: Graph showing the age wise distribution of lung cancer on the basis of gender

5.4.4 Distribution of lung cancer patients on the basis of stage of tumor and TNM

Clinico- pathological parameters of the lung cancer subjects were recorded from the medical records of the hospital. The patients were classified on the basis of clinical stage of the tumor into stage I, stage II, stage III and stage IV. The proportion of stage I and stage II patients was found to be very less as compared to stage III and stage IV as shown in fig 5.12. Therefore, major proportion of the patients were reported to possess advanced etiology of the disease. The percentage of stage I (0.93%) and stage II (4.06%) patients is much less than stage III (47.5%) and stage IV (41.2%). Further, the patients were also stratified on the basis of the various parameters of TNM such as primary tumor extension (T), lymph node involvement (N) and metastasis (M). The detailed distribution of patients on the basis of tumor size is as

follows: TX+T1 (7.5%), T2 (12.5%), T3 (25.3%) and T4 (44.4%). Regarding, the lymph node involvement, the patients are distributed as NX+N0 (15.3%), N1 (11.6%), N2 (39.1%) and N3 (23.7%). Similarly, the on basis of metastatic status, patients were classified into M0 (51.3%) and M1 (38.4%). For the remaining 10.3% cases the TNM parameters were not available.

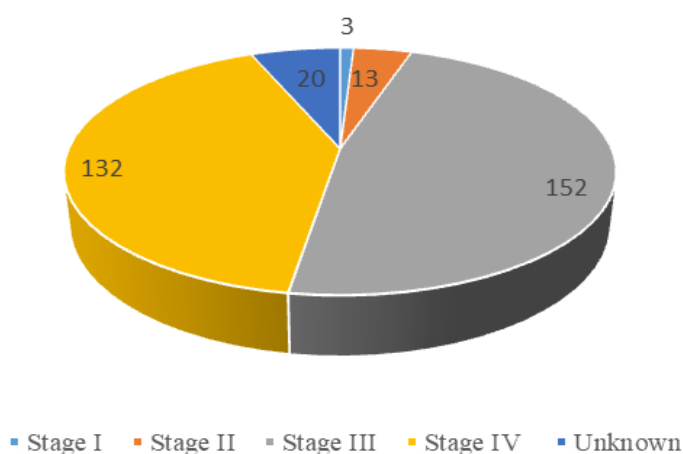


Fig 5.12: Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of clinical stage

5.4.5 Distribution of lung cancer patients on the basis of histology of tumor

Lung cancer is broadly classified on the basis of histopathological specifications into SQCC, ADCC and SCLC. The patients enrolled in this study were stratified on the same basis. In the present study, the prevalence of SQCC patients (44.4%) was highest among all other histological subtypes followed by ADCC (31.5%) and SCLC (22.8%) as illustrated in fig 5.13. The rest 1.3% of the patients were identified with miscellaneous histological subtypes such as large cell carcinoma, mesothelial and NSCLC undifferentiated. Therefore, it is evident that SQCC is the most frequently occurring type of lung cancer in north Indians. An evaluation of number of smokers in each histological category was carried out which revealed that SCLC subgroup has 91.7% smokers, followed by SQCC having 89.4% smokers and ADCC subgroup has the least percentage of smokers (54.4%). Hence, it is evident that smoking is the most prevalent risk factor in case of SCLC patients.

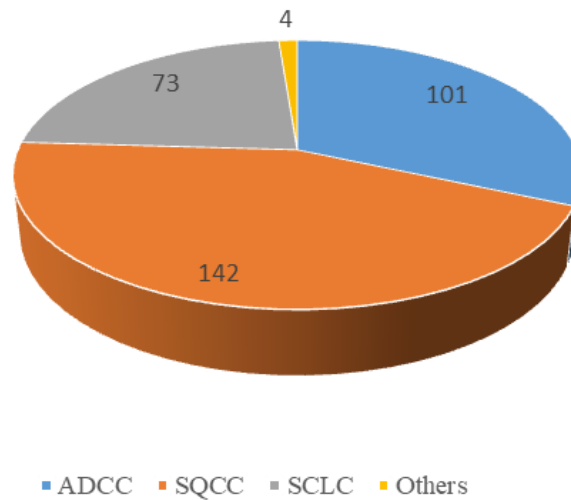


Fig 5.13: Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of histological subtypes.

5.4.6 Distribution of lung cancer patients on the basis of performance status

In order to evaluate the performance of cancer patients during the treatment and to forecast the prognosis of the disease, the patients are classified using two common methods Karnofsky performance status (KPS) and Eastern Cooperative Oncology Group-performance status (ECOG-PS). These methods help to assess the functional condition of the patients. The distribution of patients in accordance to their KPS showed that 57.5% had KPS score of (60-80), 35.6% had KPS of (90-100) and the least percentage (21.8%) of patients were found to have KPS score of (30-50) indicating poor performance status as illustrated in fig 5.14. Similarly, the patients are also stratified according to ECOG performance status and it was observed that 15.3% of patients had ECOG status (0), 13.1% had ECOG status (1), 35.5% had ECOG status (2) and 34.1% had ECOG status (3-4) as shown in fig 5.15.

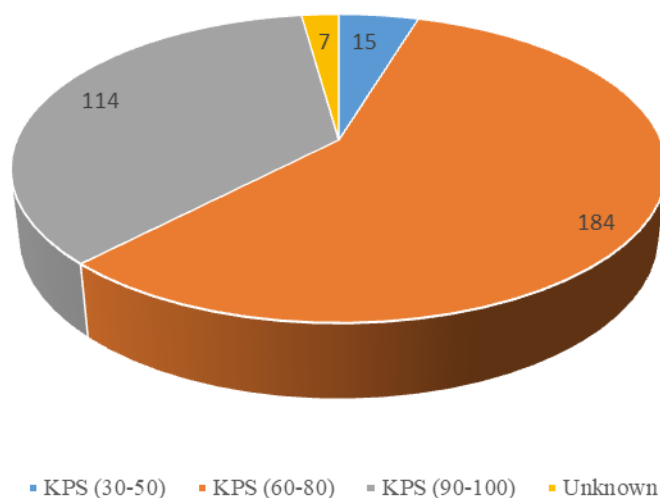


Fig 5.14: Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of Karnofsky performance status (KPS)

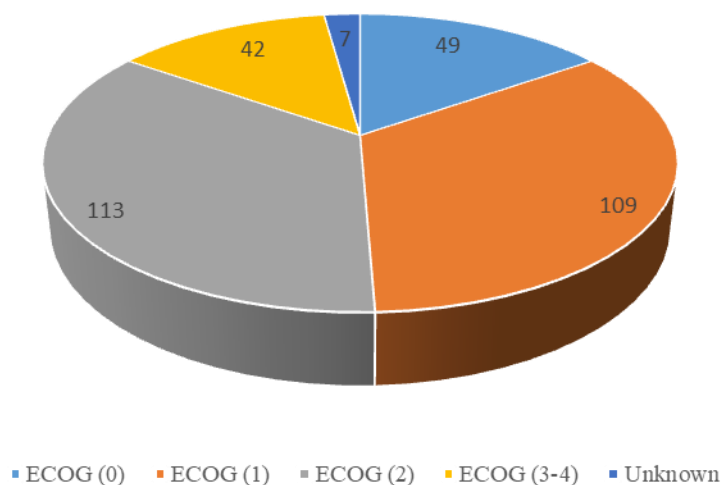


Fig 5.15: Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of Eastern Cooperative Oncology Group -performance status (ECOG-PS).

5.4.7 Distribution of lung cancer patients on the basis of doublet platinum based chemotherapeutic regimen:

Depending upon the histological subtype of lung cancer and taking other basic clinical parameters into consideration, the patients were treated with different combination of chemotherapeutic drugs. The combination comprises a platinum based drug such as cisplatin or carboplatin along with other drugs such

as docetaxel, paclitaxel, pemetrexed, gemcitabine and irinotecan. The highest proportion (22.5%) of patients were treated with docetaxel and cis/carboplatin regimen followed by irinotecan with cis/carboplatin (17.5%) and pemetrexed with cis/carboplatin (16.2%). A very few patients were treated with paclitaxel and cis/carboplatin (5%) followed by gemcitabine along with cis/carboplatin (1.3%) as shown in fig 5.16. Due to unavailability of records about the chemotherapeutic regimen of some patients, 34.6% had fallen in the unknown category.

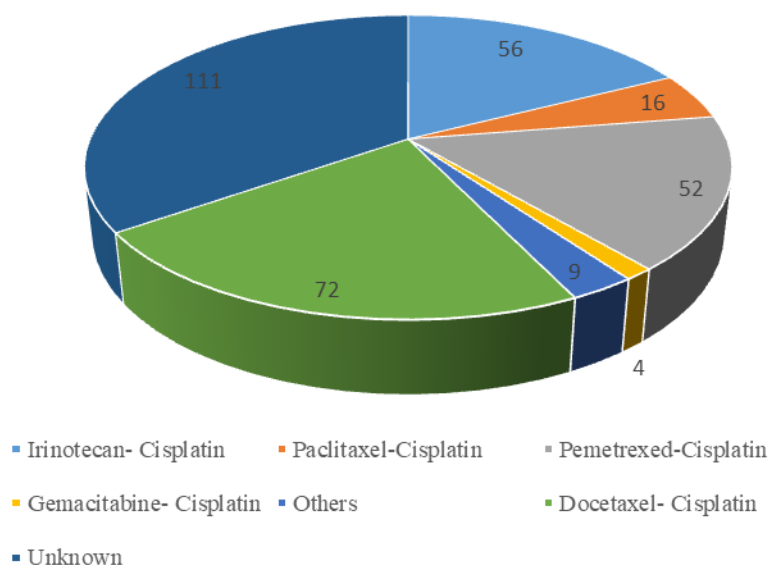


Fig 5.16: Pie chart illustrating distribution of North Indian Lung cancer patients on the basis of Doublet platinum based chemotherapeutic regimen

5.5 Statistical Analysis for polymorphic sites of *DDK* gene

5.5.1 Demographic characteristics

The sample population can be characterized by various epidemiological and demographic parameters including age, gender, smoking status, and the clinical specifications of the cases as tabulated in table 5.1. The mean age came out to be 57.3 ± 10.74 years and 53.23 ± 10.44 years in cases and controls respectively. The study was performed on 300 cases and 300 controls including 258 (86%) males in cases and 253 (84.3%) males in controls whereas 42 (14%) and 47 (15.7%) were females both in case and control groups respectively. No significant difference was found between the distribution of males and females in cases and controls ($p=0.64$), hence there is no sampling bias based on gender. The number of smokers and non-smokers were also compared between the two study groups. The case group comprised 78% smokers

whereas controls had only 71% smokers, a significant difference was observed in the proportion of smokers and non-smokers ($p=0.01$). The attempt to adequately match the number of smokers and non-smokers was made but as established smoking is a significant parameter responsible for occurrence of lung, same was observed in this sample population also. Smoking was quantified in terms of pack years [(cigarettes or beedis per day/20) X years smoked], the means of pack years were also compared among cases and controls and a significant difference was observed ($p=0.003$). The parameters such as age, gender and smoking were adjusted for genotype analysis using unconditional logistic regression as they can act as confounding factors. Further, the different clinical specifications such as histology, clinical stage, tumor size extensions, lymph node invasion and distant metastasis were also recorded from medical reports. Out of the 300 cases, 128 (42.8%) were histologically identified as SQCC, 97 (32.3%) as ADCC and 70 (23.3%) as SCLC. The TNM data was available for 270 cases and when classified on the basis of clinical stage, only few cases were reported with stage I, II whereas most of them were categorized into stage III and IV respectively. (I: 3, II:12, III:140, IV:123) as shown in table 5.1.

Table 5.1: Demographic characteristics among cases and controls

Variable	Cases, n (%) N=300	Controls, n (%) N=300	p^a – value
Age (years)			
Mean $\pm$ SD	57.38 $\pm$ 10.74	53.23 $\pm$ 10.44	$p < 0.0001$
Range			
Gender			
Male	258 (86)	253 (84.3)	0.645
Female	42 (14)	47 (15.7)	
Smoking Status			
Smokers	235 (78.3)	208 (71.4)	0.01
Non-Smokers	65 (21.7)	92 (28.6)	
Pack Years			
Mean $\pm$ SD	28.09 $\pm$ 34.09	19.4 $\pm$ 19.56	0.0003
Histological Types			
SQCC	128 (42.8)		
ADCC	97 (32.3)		
SCLC	70 (23.3)		
Others	4 (1.3)		
Unknown	1 (0.3)		
TNM Staging			
I	3 (1)		
II	12 (4)		
III	140 (46.7)		
IV	123 (41)		
Unclassified	22 (7.3)		
Tumor Size			
Tx	10 (3.3)		
T1	14 (4.7)		
T2	39 (13)		
T3	79 (26.3)		
T4	128 (42.7)		
Unknown	30 (10)		
Lymph Node			
N0	48 (16)		

N1	36 (12)
N2	112 (37.3)
N3	72 (24)
Unknown	30 (10)
Metastasis	
M0	154 (51.3)
M1	116 (38.7)
Unknown	30 (10)

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects ^a*p*-values were derived from Pearson Chi – square test except age; Student t-test was used for age. All *p*-values are two – sided. *p*< 0.05 was considered statistically significant

5.5.2 Genotypic and allelic frequencies of *DKK* polymorphic variants among cases and controls and their role in modifying the susceptibility towards lung cancer

Table 5.2 summarizes the allelic and genotypic frequencies for all the seven variants of *DKK* gene including *DKK2*, *DKK3* and *DKK4*. In case of *DKK4* rs2073664 (C/T) no significant difference between the distribution of genotypic frequencies was found between cases and controls. After applying HWE, it was observed that controls did not deviate from the same ($\chi^2=3.7, df=2, p=0.15$). The percentage of subjects carrying heterozygous CT genotype was found to be over represented in controls (47.7%) as compared to cases (45.6%). The minor allele frequency (MAF) was found out to be 0.28 and 0.26 in controls and cases respectively. When the odds were calculated using wild genotype (CC) as reference, no significant association was seen in subjects carrying one variant allele (CT) genotype (OR=0.95, 95% C.I:0.68- 1.34, *p*=0.80) however, a protective effect was observed in subjects having both variant alleles (TT) genotype (OR=0.37, 95% C.I: 0.14-0.95, *p*=0.03).

Various polymorphic sites of *DKK3* gene including rs2291599, rs3206824 and rs7391689 were also investigated for their role in modulating lung cancer susceptibility as shown in table 5.2. For *DKK3* rs2291599 (C/T), the genotypic and allelic frequency of control subjects was found to follow HWE ($\chi^2=1.85, df=2, p=0.17$). The proportion of subjects with wild (CC) genotype was found to be similar in both cases and controls (65%) however, the percentage of subjects carrying heterozygous (CT) genotype were found to be over represented in controls as compared to the cases (29.7% vs 27%). The difference between the genotypic frequencies between cases and controls was not found to be significant ($\chi^2=1.97, df=2, p=0.37$). The MAF was 0.21 in cases and 0.20 in controls. The association analysis revealed that subjects with mutant (TT) genotype were at a slightly higher risk (OR= 1.46, 95% C.I: 0.74-2.88, *p*=0.26) of acquiring lung cancer but it wasn't statistically significant. The next variant of *DKK3* gene which was analyzed is rs3206824 (G/A). The genotypic frequencies for *DKK3* rs3206824 (G/A) in control group were following the HWE ($\chi^2=0.72, df=2, p=0.39$). No significant difference in genotypic frequencies among cases and controls ($\chi^2=0.80, df=2, p=0.66$) was observed. The MAF was found to be 0.17 and 0.15 in cases and controls respectively. The

percentage of subjects having wild (GG) genotype among cases and controls was 68% and 71.3% respectively. Also, the proportion of study subjects carrying heterozygous (30% vs 27%) and mutant genotype (3% vs 1.7%) were almost similar among cases and controls respectively. The logistic regression approach revealed that the subjects carrying mutant (AA) genotype did not have any significant risk of developing lung cancer (OR=1.29, 95% C.I:0.38-4.39, $p=0.68$). The genotypic frequencies for the last variant of *DKK3* gene rs7391689 (G/C) were also tabulated. Since the frequency of the mutant (CC) genotype was less represented (0.3%), we combined the heterozygous and mutant as a single genotype (GC+CC). The proportion of subjects with variant rs7391689 alleles (GC+CC) were found to be slightly higher in the control group as compared to cases (36.6% vs 27.3%). The distribution of the genotypic frequencies did show a significant difference between cases and controls ($\chi^2=5.1$, df=1, $p=0.02$).

Further, the variants of *DKK2* gene were also analyzed in order to investigate their role in modifying the predisposition towards lung cancer as shown in table 5.2. The first polymorphic site which was investigated was *DKK2* rs447372 (A/G). The genotypic frequencies of the control population showed a slight fluctuation from the HWE ($\chi^2=5.32$, df=1, $p=0.02$). The percentage of subjects having heterozygous AG genotype were over represented in controls in comparison to cases (47% vs 41%). However, the greater proportion of subjects was found to have mutant (GG) genotype in cases than in controls (9% vs 5%). The difference in the genotypic frequencies was also not found to be significant ($\chi^2=3.5$, df=2, $p=0.17$). The MAF in both cases and controls was found to be 0.29.

The allelic and genotypic frequencies of the second polymorphic site of *DKK2* (rs419558 C/T) were in accordance to HWE ($\chi^2=3.78$, df=1, $p=0.05$). The percentage of subjects with heterozygous CT genotype was over represented in cases as compared to controls (32.3% vs 20.7%). There was a significant difference in the genotypic frequencies between cases and controls ($\chi^2=9.8$, df=1, $p=0.001$). The MAF was found out to be 0.10 in controls and 0.16 in cases. Taking CC genotype as reference, logistic regression was applied to reveal that the subjects with the heterozygous CT genotype were at a higher risk of developing lung cancer (OR= 1.77, 95% C.I: 1.21-2.58, **$p=0.002$**) when compared with subjects carrying wild (CC) genotype. The third single nucleotide variant of *DKK2* gene analyzed was rs17037102 (G/A). The genotypic and allelic frequencies were in accordance with HWE ($\chi^2=3.01$, df=1, $p=0.08$). The proportion of subjects with heterozygous (GA) genotype were found to be over represented in cases as compared to controls (28.3% vs 18.3%) and subjects carrying the wild genotype were higher in controls as compared to cases (81.7% vs 71.7%). There was significant difference among the genotypic frequencies of cases and controls ($\chi^2=7.8$, df=1, $p=0.005$). The MAF was found to be 0.09 in controls and 0.14 in cases. The association analysis showed that the individuals with heterozygous (GA) genotype (OR=1.62 95% C.I: 1.09-2.40, **$p=0.01$**) and the combined variant (GA+AA) genotype (OR=1.69 95% C.I: 1.14-2.50, **$p=0.008$**) were at marginally higher

risk of developing lung cancer which was found to be statistically significant. The heterozygous (GA) and mutant (AA) genotype were combined as single genotype since the percentage of subjects with mutant genotype was very low.

Table 5.2: Genotypic and allelic distribution of the *DKK* genetic variants and their association with risk of lung cancer

Genotype	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	<i>p</i>^a
<i>DKK4</i>, rs2073664 C/T				
CC	147 (49)	150 (50)	1.00 (Reference)	
CT	137 (45.7)	143 (47.7)	0.95 (0.68- 1.34)	0.80
TT	16 (5.3)	7 (2.3)	0.37 (0.14-0.95)	0.03
CT+TT	143 (51)	150 (50)	0.89 (0.64-1.24)	0.50
C (Allele)	431	443		
T (Allele)	169	157		
MAF	0.28	0.26		
HWE	$\chi^2=3.7, df=2$			0.15
<i>DKK3</i>, rs2291599 C/T	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	<i>p</i>^a
CC	195 (65)	195 (65)	1.00 (Reference)	
CT	89 (29.7)	81 (27)	0.88 (0.61-1.28)	0.53
TT	16 (5.3)	24 (8)	1.46 (0.74-2.88)	0.26
CT+TT	105 (35)	105 (35)	0.97 (0.68-1.37)	0.87
C (Allele)	479	471		
T (Allele)	121	129		
MAF	0.20	0.21		
HWE	$\chi^2=1.85, df=2$			0.17
<i>DKK3</i>, rs3206824 G/A	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	<i>p</i>^a
GG	214 (71.3)	204 (68)	1.00 (Reference)	
GA	81 (27)	90 (30)	1.14 (0.79-1.64)	0.47
AA	5 (1.7)	6 (3)	1.29 (0.38-4.39)	0.68
GA+AA	86 (28.7)	96 (32)	1.15 (0.80-1.64)	0.42
G (Allele)	509	498		
A (Allele)	91	102		
MAF	0.15	0.17		
HWE	$\chi^2=0.72, df=2$			0.39
<i>DKK3</i>, rs7396187 G/C	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	<i>p</i>^a
GG	190 (63.3)	217 (72.3)	----	
GC	110 (36.7)	82 (27.4)		
CC	...	1 (0.3)		
GC+CC	110 (36.7)	83 (27.7)		
G (Allele)	490	516		
C (Allele)	110	84		
MAF	0.18	0.14		
HWE	$\chi^2=15.11, df=2$			0.001
<i>DKK2</i>, rs447372 A/G	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	<i>p</i>^a
AA	143 (48)	150 (50)	-----	
AG	140 (47)	123 (41)		
GG	17 (5)	27 (9)		
AG+GG	157 (52)	150 (50)		
A (Allele)	426	423		

G (Allele)	174	177		
MAF	0.29	0.295		
HWE	$\chi^2=5.32$, df=2			0.02
DKK2, rs419558 C/T	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	p^a
CC	238 (79.3)	203 (67.7)	1.00 (Reference)	
CT	62 (20.7)	97 (32.3)	1.77 (1.21-2.58)	0.002
C (Allele)	538	503		
T (Allele)	62	97		
MAF	0.10	0.16		
HWE	$\chi^2=3.78$, df=1			0.05
DKK2, rs17037102 G/A	Controls (300) N (%)	Cases (300) N (%)	AOR (95% CI)^b	p^a
GG	245 (81.7)	215 (71.7)	1.00 (Reference)	
GA	55 (18.3)	82 (27.3)	1.62 (1.09-2.40)	0.01
AA	...	3 (1)	...	
GA+AA	55 (18.3)	85 (28.3)	1.69 (1.14-2.50)	0.008
G (Allele)	545	512		
A (Allele)	55	88		
MAF	0.09	0.14		
HWE	$\chi^2=3.01$, df=2			0.08

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

5.5.3 Genotypic and allelic distribution of the *DKK2*, *DKK4* and *DKK3* gene and risk of lung cancer on basis of histological subtypes

In order to segregate the population into high risk subgroups, the patients were divided into categories based upon their histological subtypes. The comparison of the genotypic frequency of the *DKK* variants was carried to find out the prevalence of the high risk subgroups from all the categories as shown in table 5.3. In case of *DKK4* rs2073664 (C/T), the genotypic frequencies for mutant (TT) genotype in different histological types namely SQCC, ADCC and SCLC were found to be similar to the control group (2.3, 3.1, 1.4 vs 5.3% respectively) as shown in table 5.3. When odds ratio was calculated a trend for protective effect is observed in case of SQCC and SCLC patients having TT genotype when compared to wild (CC) genotype, however it was not found to be statistically significant. Hence, none of the histological subgroup showed prominent effect on the risk of developing lung cancer. For *DKK3* rs2291599 (C/T), an increased frequency of subjects with mutant (TT) genotype was observed in SQCC (7.8%) and SCLC (10%) patients as compared to controls (5.3%) as shown in table 5.3. When compared with the wild (CC) genotype, a risk was found to be associated both in SQCC and SCLC patients having mutant (TT) genotype but the risk was not statistically significant, nonetheless a visible trend was observed. In case of *DKK3* rs3206824 (G/A) it was observed that the frequency of mutant (AA) genotype was present in higher frequency in SQCC patients as compared to controls (3.1% vs 1.7%) but no significant association was seen, although a two-fold higher odds ratio was

observed in these patients having AA genotype (OR=2.04 95% C.I.: 0.51-8.05, $p=0.30$). In case of ADCC and SCLC subtypes, no visible difference was observed in the frequency of AA genotype subjects in comparison to control group (1%, 1.4% vs 1.7% respectively). Therefore, no significant correlation was observed in the any of the histological subgroup.

DKK2 rs419558 (C/T) polymorphic site did confer a modulation in the predisposition towards lung cancer in all the histological subtypes. The percentage frequency of heterozygous (CT) genotype was higher in all histological subgroups namely SQCC, ADCC and SCLC as compared to controls (32.8%, 30.9%, and 32.9% vs 20.7% respectively). A similar increase in risk towards lung cancer was observed in all the three histological sub-groups, where patients with heterozygous CT genotype were at approximately 1.8 times higher risk of developing lung cancer. The adjusted odds ratio obtained in case of each histological subgroup was found to be statistically significant with $p=0.02$ in both SQCC and ADCC category and 0.04 in SCLC subgroup. On the other hand, in case of *DKK2* rs17370102 (G/A), a significant increase in frequency of (GA+AA) genotype patients was observed in SQCC group (33.6%) as compared to controls (18.3%). This difference in frequencies ultimately lead to the disposition of 2.2 fold higher risk for SQCC patients carrying variant genotype (OR=2.28, 95% C.I.: 1.40-3.72, $p=0.0009$). Also, SQCC patients carrying heterozygous (GA) genotype showed a two-fold risk of developing lung cancer with p -value of 0.002. However, the frequency of ADCC patients carrying combined variant genotype was 26.8% but the risk estimated was found to be statistically not significant. Similar was the case in SCLC subgroup, where the frequency of subjects having heterozygous genotype was 22.9% and hence, the other two histological subgroups did not show a significantly increased frequency of patients with GA or GA+AA genotype, therefore no significant association was seen in SCLC and ADCC group of patients.

Table 5.3: Genotypic distribution of the *DKK* genetic variants and their association with risk of lung cancer based on histology

Genotype		ADCC			SQCC			SCLC		
<i>DKK4</i> , rs207366 C/T	Controls (300) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (70) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	147 (49)	65 (50.8)	1.00(Reference)		52 (53.6)	1.00 (Reference)		31 (44.3)	1.00 (Reference)	
CT	137 (45.7)	60 (46.9)	0.91 (0.59-1.41)	0.69	42 (43.3)	0.92 (0.57-1.49)	0.74	38 (54.3)	1.14 (0.66-1.98)	0.62
TT	16 (5.3)	3 (2.3)	0.34 (0.09-1.24)	0.10	3 (3.1)	0.55 (0.15-2.02)	0.37	1 (1.4)	0.14 (0.013-1.63)	0.11
CT+TT	143 (51)	63 (49.2)	0.84 (0.54-1.29)	0.44	45 (46.4)	0.88 (0.15-1.41)	0.60	39 (55.7)	1.02 (0.59-1.76)	0.93
		ADCC			SQCC			SCLC		
<i>DKK3</i> , rs2291599 C/T	Controls (300) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (70) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	195 (65)	79 (61.7)	1.00(Reference)		63 (65)	1.00 (Reference)		50 (71.4)	1.00 (Reference)	
CT	89 (29.7)	39 (30.5)	1.18 (0.73-1.91)	0.47	28 (28.9)	0.81 (0.48-1.39)	0.46	13 (18.6)	0.61 (0.30-1.22)	0.16
TT	16 (5.3)	10 (7.8)	1.50 (0.64-3.54)	0.34	6 (6.1)	1.19 (0.44-3.25)	0.72	7 (10)	1.83 (0.65-5.19)	0.25
CT+TT	105 (35)	49 (38.3)	1.22 (0.78-1.91)	0.36	34 (35)	0.85 (0.52-1.41)	0.55	20 (28.6)	0.78 (0.43-1.43)	0.43
		ADCC			SQCC			SCLC		
<i>DKK3</i> , rs3206824 G/A	Controls (300) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (70) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
GG	214 (71.3)	79 (61.7)	1.00 (Reference)		70 (72.2)	1.00 (Reference)		52 (74.3)	1.00 (Reference)	
GA	81 (27)	45 (35.2)	1.42 (0.89-2.26)	0.13	26 (26.8)	0.95 (0.56-1.62)	0.86	17 (24.3)	0.81 (0.43-1.52)	0.52
AA	5 (1.7)	4 (3.1)	2.04 (0.51-8.05)	0.30	1 (1)			1 (1.4)		...
GA+AA	86 (28.7)	49 (38.3)	1.46 (0.93-2.29)	0.09	27 (27.8)	0.95 (0.56-1.60)	0.85	18 (25.7)	0.81 (0.44-1.50)	0.51
		ADCC			SQCC			SCLC		
<i>DKK2</i> , rs419558 C/T	Controls (300) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Case (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (70) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	238 (79.3)	86 (67.2)	1.00(Reference)		67 (69.1)	1.00 (Reference)		47 (67.1)	1.00 (Reference)	
CT	62 (20.7)	42 (32.8)	1.71 (1.05-2.76)	0.02	30 (30.9)	1.87 (1.10-3.18)	0.02	23 (32.9)	1.84 (1.02-3.34)	0.04
		ADCC			SQCC			SCLC		
<i>DKK2</i> , rs17037102 G/A	Controls (300) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (70) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
GG	245 (81.7)	85 (66.4)	1.00 (Reference)		71 (73.2)	1.00 (Reference)		54 (77.1)	1.00 (Reference)	
GA	55 (18.3)	41 (32)	2.13 (1.29-3.49)	0.0028	25 (25.8)	1.54 (0.87-2.69)	0.13	16 (22.9)	1.07 (0.55-2.06)	0.83
AA	...	2 (1.6)		...	1 (1)	...		...		
GA+AA	55 (18.3)	43(33.6)	2.28 (1.40-3.72)	0.0009	26 (26.8)	1.60 (0.92-2.79)	0.09	...		

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.5.4 Effect of smoking status on the association of *DKK* gene polymorphisms and lung cancer risk

In order to study gene-environment interaction, patients were subdivided into smokers and non-smokers based upon their smoking status. The genotypic frequencies for smokers and non-smokers have been summarized in table 5.4 (a) and table 5.4 (b) respectively along with the associated odds ratio, 95% C.I. and *p*-values. *DKK4* rs2073664 (C/T) variant did not confer any contribution in increasing the lung cancer risk neither in smokers nor in non-smokers as shown in table 5.4 (a) and 5.4 (b). However, a decreased frequency of patients with mutant (TT) genotype was observed among smokers when compared to controls (2.5% vs. 6.2%) and therefore, a strong protective effect was observed (OR=0.32, 95% C.I.: 0.11-0.94, *p*=0.03). Also, in case of non-smokers a lower frequency of subjects having the mutant (TT) genotype was observed among cases (1.5%) in comparison to controls (3.3%) but the association analysis exhibited a lack of any significant correlation between lung cancer risk and this polymorphic site.

In case of *DKK3* rs2291599, it was observed that patients carrying CT or TT genotype did not show any significantly increased frequency in comparison to controls among both the smoker and non-smoker subgroup. The percentage of mutants among smokers was 8.5%, which was slightly higher as compared to controls (5.3%). As a result, a slight trend of increased risk was observed in case of smokers which carry mutant (TT) genotype (OR=1.59, 95% C.I.: 0.72-3.98, *p*=0.24). On the other hand, the mutants among non-smokers were prevalent at similar frequencies in both cases (6.2%) and controls (5.5%). Therefore, no significant association was observed among non-smokers.

In case of *DKK3* rs3206824 (G/A), the frequency of heterozygotes (28.1% vs 28.4%) and mutants (2.5% vs 2.4%) among smoker group was found to be almost similar in case and control group respectively. Hence, it was evident that this polymorphic site was not modulating the lung cancer risk among smokers. On the contrary, as the frequency of heterozygotes was 36.9% among cases and 24.1% among controls henceforth, non-smoker subjects carrying heterozygous (GA) genotype exhibited a two-fold risk of developing lung cancer (OR=1.91, 95% C.I.: 0.93-3.84, *p*=0.04).

DKK2 rs17370102 (G/A) variant also did not affect lung cancer susceptibility in case of smoker subjects as the frequency of subjects having combined variant (GA+AA) genotype was not significantly higher among cases (28.1%) as compared to controls (20.7%) as shown in table 5.4 (a). However, non-smokers with variant (GA+AA) genotype of *DKK2* rs17370102 (G/A) were found at almost 2.7 times higher risk of developing lung cancer (OR=2.69, 95% C.I.: 1.16-6.24, *p*=0.02) as the subjects carrying this genotype were over represented among cases (29.2%) in comparison to controls (13.2%).

In case of *DKK2* rs419558 (C/T) polymorphic site, patients who were smokers and had heterozygous (CT) genotype were found to have a 2-fold risk for developing lung cancer (OR=1.94, 95%CI: 1.60-2.99, ***p*=0.002**). As the controls comprised of only 22.6% heterozygous subjects whereas the case group included 34.8% of them. However, no such significant risk was observed in non-smoker category as no significant difference was evident among both cases and controls as shown in table 5.4 (b).

Table 5.4 (a): Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer (smokers)

Genotype	SMOKERS			
	Controls n (%) N =208	Cases n (%) N=235	AOR (95% CI) ^b	<i>p</i> ^a
<i>DDK4</i>, rs2073664 C/T				
CC	94 (45.2)	113 (48.1)	1.00 (Reference)	
CT	101 (48.6)	116 (49.4)	0.94 (0.63-1.39)	0.77
TT	13 (6.2)	6 (2.5)	0.32 (0.11-0.94)	0.03
CT+TT	114 (54.8)	122 (51.9)	0.87 (0.59-1.27)	0.48
<i>DDK3</i>, rs2291599 C/T				
CC	141 (67.8)	158 (67.2)	1.00 (Reference)	
CT	56 (26.9)	57 (24.3)	0.88 (0.56-1.37)	0.58
TT	11 (5.3)	20 (8.5)	1.59 (0.72-3.48)	0.24
CT+TT	67 (32.2)	77 (32.8)	1.00 (0.66-1.50)	0.99
<i>DDK3</i>, rs3206824 G/A				
GG	144 (69.2)	163 (69.4)	1.00 (Reference)	
GA	59 (28.4)	66 (28.1)	0.98(0.64-1.50)	0.94
AA	5 (2.4)	6 (2.5)	1.35 (0.39-4.60)	0.63
GA+AA	64 (30.8)	72 (30.6)	1.01 (0.66-1.52)	0.96
<i>DDK2</i>, rs17037102 G/A				
GG	165 (79.3)	169 (71.9)	1.00 (Reference)	
GA	43 (20.7)	65 (27.6)	1.43 (0.91-2.24)	0.11
GA+AA	43 (20.7)	66 (28.1)	1.45 (0.92-2.27)	0.10
<i>DDK2</i>, rs419558 C/T				
CC	161 (77.4)	151 (64.2)	1.00 (Reference)	
CT	47 (22.6)	82 (34.8)	1.94 (1.6-2.99)	0.0024

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

Table 5.4 (b): Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer (non-smokers)

Genotype	NON SMOKERS			
<i>DKK4</i> , rs2073664 C/T	Controls n (%) N =92	Cases n (%) N=65	AOR (95% CI) ^b	<i>p</i> ^a
CC	53 (57.1)	37 (57)	1.00 (Reference)	
CT	36 (39.6)	27 (41.5)	1.03 (0.52-2.03)	0.91
TT	3 (3.3)	1 (1.5)	0.53 (0.04-5.72)	0.60
CT+TT	39 (42.8)	28 (43)	0.99 (0.51-1.93)	0.99
<i>DKK3</i> , rs2291599 C/T	Controls n (%) N =92	Cases n (%) N=65	AOR (95% CI) ^b	<i>p</i> ^a
CC	54 (59.3)	37 (56.9)	1.00 (Reference)	
CT	33 (36.2)	24 (36.9)	0.94 (0.46-1.94)	0.86
TT	5 (5.5)	4 (6.2)	1.00 (0.24-4.15)	0.99
CT+TT	38 (41.7)	28 (43.1)	0.91 (0.46-1.79)	0.80
<i>DKK3</i> , rs3206824 G/A	Controls n (%) N =92	Cases n (%) N=65	AOR (95% CI) ^b	<i>p</i> ^a
GG	70 (76.9)	41 (63.1)	1.00 (Reference)	
GA	22 (24.1)	24 (36.9)	1.91 (0.93-3.89)	0.04
AA	...	...		
GA+AA	22 (24.1)	24 (36.9)		
<i>DKK2</i> , rs17037102 G/A	Controls n (%) N =92	Cases n (%) N=65	AOR (95% CI) ^b	<i>p</i> ^a
GG	80 (86.8)	46 (70.8)	1.00 (Reference)	
GA	12 (13.2)	17 (26.2)	2.32 (0.98-5.49)	0.05
GA+AA	12 (13.2)	19 (29.2)	2.69 (1.16-6.24)	0.02
<i>DKK2</i> , rs419558 C/T	Controls n (%) N =92	Cases n (%) N=65	AOR (95% CI) ^b	<i>p</i> ^a
CC	77 (83.5)	50 (76.9)	1.00 (Reference)	
CT	15 (16.5)	15 (23.1)	1.48 (0.65-3.38)	0.34

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

5.5.5 Association between lung cancer risk and different combinations of *DKK4*, *DKK3* and *DKK2* genotypes

Further, in order to investigate gene-gene and SNP-SNP interaction the genotypic combinations of the various *DKK* variant genotypes were analyzed to find out whether there is a significant difference in frequency between cases and controls, the data is tabulated in table 5.5. The genotypic combination comprising *DKK3* rs2295199 (C/T) and *DKK2* rs419558 (C/T) showed a 1.5 fold higher risk of lung cancer however, this was not found to be statistically significant as the frequency of heterozygotes was 16.1% in cases and 10.5% among controls. Out of all the double combinations, a two-fold increased risk of lung cancer was observed in patients with variant (GA+AA+GC) genotype of *DKK3* rs3206824 (G/A) and *DKK2* rs419558 (C/T) (OR=2.32, 95% C.I: 1.23-4.36, ***p*=0.008**) as the proportion of subjects having variant genotype was much higher in cases (20.1%) as compared to controls (9.1%). Also, there were no subjects reported with CC genotype of *DKK2* rs419558 (C/T), so they were not included in the analysis.

Table 5.5: Genotypic Distribution based on different genotypic combinations and their association with lung cancer risk

Genotype					
<i>DKK3</i>, rs2291599 C/T & <i>DKK3</i>, rs3206824 G/A	Cases, n (%) N = 202	Controls, n (%) N = 191	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	156 (77.2)	154 (80.6)	1.00 (Reference)		
1	44 (21.8)	35 (18.3)	1.24 (0.74-2.08)	0.40	0.76
2	2 (1)	2 (1.1)	0.94 (0.12-6.94)	0.95	0.95
3	46 (22.8)	37 (19.4)	1.22 (0.74-2.03)	0.42	0.7
<i>DKK3</i>, rs2291599 C/T & <i>DKK2</i>, rs419558 C/T	Cases, n (%) N = 155	Controls, n (%) N = 172	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	130 (83.9)	154 (89.5)	1.00 (Reference)		
1	25 (16.1)	18 (10.5)	1.56 (0.80-3.02)	0.18	0.87
<i>DKK3</i>, rs2291599 C/T & <i>DKK2</i>, rs17037102 G/A	Cases, n (%) N = 161	Controls, n (%) N = 181	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	139 (86.3)	162 (89.5)	1.00 (Reference)		
1	22 (13.7)	19 (10.5)	1.31 (0.67-2.58)	0.42	0.87
<i>DKK3</i>, rs2291599 C/T & <i>DKK4</i>, rs2073664 C/T	Cases, n (%) N = 150	Controls, n (%) N = 131	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	103 (68.7)	94 (71.7)	1.00 (Reference)		
1	46 (30.7)	36 (27.5)	1.12 (0.65-1.19)	0.67	0.87
2	1 (0.6)	1 (0.8)	0.67 (0.04-11.02)	0.77	0.97
3	47 (31.3)	37 (28.3)	1.10 (0.64-1.87)	0.78	0.84
<i>DKK3</i>, rs3206824 G/A & <i>DKK2</i>, rs419558 C/T	Cases, n (%) N = 179	Controls, n (%) N = 186	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	143 (79.9)	169 (90.9)	1.00 (Reference)		
3	36 (20.1)	17 (9.1)	2.32 (1.23-4.36)	0.008	0.12
<i>DKK3</i>, rs3206824 G/A & <i>DKK2</i>, rs17037102 G/A	Cases, n (%) N = 167	Controls, n (%) N = 154	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	143 (79.9)	136 (88.3)	1.00 (Reference)		
3	24 (20.1)	18 (11.7)	1.28 (0.65-2.52)	0.46	0.7
<i>DKK3</i>, rs3206824 G/A & <i>DKK4</i>, rs2073664 C/T	Cases, n (%) N = 160	Controls, n (%) N = 97	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	109 (68.1)	74 (76.3)	1.00 (Reference)		
1	50 (31.1)	23 (23.7)	1.40 (0.77-2.55)	0.26	0.72
3	51 (31.9)	23 (23.7)			
<i>DKK2</i>, rs17037102 G/A & <i>DKK4</i>, rs2073664 C/T	Cases, n (%) N = 157	Controls, n (%) N = 150	Adjusted OR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	111 (70.7)	121 (80.7)	1.00 (Reference)		
3	46 (29.3)	29 (19.3)	1.52 (0.87-2.67)	0.13	0.30

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype. FDR: False Discovery Rate.

5.5.6 Linkage Disequilibrium and Haplotype Analysis in *DKK2* and *DKK3* gene polymorphisms:

Linkage disequilibrium and haplotype frequencies were calculated and tabulated as shown in table 5.6 for all the three possible combination of SNP's of *DKK2* gene and *DKK3* gene. The haplotype with the frequencies less than 0.03 were not included in the analysis. *DKK3* rs2291599 (C/T) and *DKK3* rs3206824 (G/A) showed linkage disequilibrium with $D'=0.36$, $r^2=0.096$, global $\chi^2=2.90$, d.f=3, $p=0.40$. However, none of the haplotypes showed any significant difference between the frequencies in cases and controls.

Table 5.6a: Haplotype frequencies and lung cancer risk associated with the *DKK3* haplotypes

<i>DKK3</i> , rs2291599 C/T & <i>DKK3</i> , rs3206824 G/A	Case frequency (%)	Control frequency (%)	OR (95% CI) ^b	<i>p</i> -value ^a
C-G	424 (0.70)	428 (0.71)	0.96 (0.75-1.24)	0.80
C-A	47 (0.07)	51 (0.08)	0.91 (0.60-1.38)	0.66
G-T	74 (0.12)	81 (0.13)	0.90 (0.64-1.26)	0.53
A-T	55 (0.09)	40 (0.06)	1.41 (0.92-2.15)	0.10

^aTwo-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls. ^bOdds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis.

The *DKK2* haplotypes were also analyzed to find out significant difference in frequency of various haplotype in cases and controls as shown in table 5.6 (b). *DKK2* rs419558 (C/T) and *DKK2* rs170370102 (G/A) showed linkage disequilibrium with $D'=0.68$, $r^2=0.01$, global $\chi^2=14.8$, df=3, $p=0.0006$. Both haplotypes T-G (OR=1.56; 95%CI: 1.10-2.21, $p=0.01$) and C-A (OR=1.58; 95%CI:1.10-2.27, $p=0.01$) showed a significant association with lung cancer risk. In case of both these haplotypes, an increased frequency is evident among cases when compared to controls.

Table 5.6 (b): Haplotype frequencies and lung cancer risk associated with the *DKK2* haplotypes

<i>DKK2</i> , rs419558 C/T & <i>DKK2</i> , rs170370102 G/A	Case frequency (%)	Control frequency (%)	OR (95% CI) ^b	<i>p</i> -value ^a
C-G	421 (0.70)	483 (0.80)	0.59 (0.45-0.77)	0.0001
T-G	91 (0.15)	5 (0.10)	1.56 (1.10-2.21)	0.01
C-A	82 (0.136)	55 (0.09)	1.58 (1.10-2.27)	0.01
T-A	6 (0.01)	0 (0.00)	...	...

^aTwo-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls. ^bOdds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis.

5.5.7 Relation between *DKK* Genotypic Distribution and Clinico-pathologic Characteristics

Table 5.7 shows the effect of *DKK* variant SNP's on the various clinic-pathological features such as tumor stage, primary tumor extension, lymph node invasion and distant metastasis. The patients were stratified according to the clinical stage into stage III and stage IV patients and also on the basis of tumor extension into TX+T1+T2 and T3+T4 and the data has been summarized in table 5.7(a).

In order to investigate the relationship of different genotypes of *DKK* genes, the genotypic frequencies were calculated for each polymorphic site for both stage III and stage IV patients as shown in table 5.7(a). The comparison in the genotypic frequencies was carried out to find out whether any significant difference lies in the genotypic frequencies among the two subgroups made on the basis of stage. Overall, none of the polymorphic site depicted any significant difference in the genotypic frequencies among the two subgroups pertaining to the lack of evident association between these genotypes and stage of the tumor. This can be clearly seen in case of *DKK3* rs3206824 where, the genotypic frequencies of the subjects carrying combined variant genotype was 31.3% in stage III and 31.7% among stage IV category. Similarly, the genotypic frequencies of heterozygotes in case of *DKK3* rs7396187 were almost identical among stage III (27.1%) and stage IV (26.9%) category.

For the next clinical parameter which is primary tumor extension as mentioned above, the patients were divided into two categories on the basis of tumor size as shown in table 5.7 (a). A higher proportion of lung cancer patients were reported in the T3+T4 category (207) whereas TX+T1+T2 subgroup consists of only 63 patients. The genotypic frequencies were estimated for each polymorphic site. None of the polymorphic site except *DKK4* rs2073664 showed any significant association between the genotype and primary tumor extension. In case of *DKK4* rs2073664, the percentage of heterozygotes (CT) in the TX+T1+T2 subgroup was higher as compared to T3+T4 tumor subgroup (58.8% vs 44%). A strong protective effect is seen in case of patients with heterozygous (OR=0.53, 95% C.I.:0.29-0.97, $p=0.04$) and combined variant genotype (OR=0.55, 95% C.I.:0.30-0.99, $p=0.04$).

In case of all the variants of *DKK3*, the frequency of heterozygotes and combined variant subjects in both the subgroups made on the basis of primary tumor extension were almost comparable. For rs2295199, the percentage of heterozygotes was 30.1% in TX+T1+T2 category and 24.6% in T3+T4 category. Similarly, in case of rs3206824 the percentage of GA+AA genotype carrying patients were 36.5% in TX+T1+T2 category and 30% in T3+T4 category as shown in table 5.7 (a).

The patients carrying GA+AA genotype of rs17037102 showed similar frequencies in T3+T4 category (28.6%) and TX+T1+T2 subgroup (29.5%). Hence, no significant association was observed between the genotype of these polymorphic sites and primary tumor extension.

Table 5.7 (b) depicts the genotypic frequencies of *DKK* variants in relation to clinical parameters such as lymph node involvement and metastasis. On the basis of lymph node invasion, patients were divided into N0 and N1+N2+N3 groups. *DKK4* rs2073664 did not confer any significant modulation towards lymph node involvement. The frequencies of patients having CT+TT genotype among the groups (N0 vs N1+N2+N3) were not significantly different (43.8% vs 51.8%). Therefore, no significant relationship was observed between this genotype and lymph node invasion.

Further, when *DKK3* genetic variants were evaluated not all but one of them did show a role in lymph node invasion as shown in table 5.7 (b). Out of the three SNP's, the genetic variant rs2291599 showed a two-fold increased probability (OR= 2.36 95% C.I: 0.97-.5.74, $p=0.05$) of lymph node invasion in patients having CT genotype for this polymorphic site as the frequency of heterozygous was higher (28.1%) in patients showing presence of lymph node involvement as compared to those lacking lymph node involvement (14.6%). The next variant rs3206824 exhibited nearly similar frequencies of subjects having GA+AA genotype amongst the groups showing positive lymph node involvement (31.9%) and negative lymph node involvement (31.2%).

Further, the next variant rs419558 reported almost similar percentage of heterozygous subjects among those showing positive lymph node invasion (34%) and lacking lymph node invasion (31%). Lastly, in case of rs17037102 the frequency of patients having GA+AA genotype was almost similar in the two groups (29% vs 31.2%). As evident, none of the polymorphic sites located in *DKK2* gene contributed in lymph node invasion in lung cancer patients.

Table 5.7 (b) also demonstrates the relationship of these different genotypes with metastatic status of the patients and further to classify the patients on the basis of distant metastatic status, they were stratified into M0 and M1+M2 group. The genotypic frequencies were estimated for each polymorphic site in accordance to the metastatic status of patients. The patients showing the positive distant metastasis are compared by taking lack of metastasis as a reference group.

In case of *DKK4* rs2073664, the genotypic frequencies of variant subjects were compared among the two groups and it was evident there was no significant difference in between the patients who showed metastasis were 50% and those showing the absence of metastasis were 49.4%. Hence, this polymorphic site did not show any association with spread of lung cancer to other organs.

DKK3 variants including rs2295199 and rs3206824 were also evaluated for their association with distant metastasis as shown in table 5.7 (b). However, none of them showed any significant association with this clinical parameter. In case of rs2295199, the percentage of subjects having CT+TT genotype were 38 and 30.5% in the subgroup positive for metastasis and the one lacking metastasis respectively. Also, when frequency of subjects having GA+AA genotype for rs3206824 was similar in metastatic (31%) and non-metastatic group (31.9%). Further, the genotypic frequencies were also estimated for various *DKK2* variants and classified according to the metastatic status of the tumor. The proportion of heterozygous subjects for rs419558 was equivalent among the two subgroups (32.8% vs 33.8%). Similar observations were recorded in case of rs17037102, where in the frequency of GA+AA genotype patients was 28.5% among both the groups. Hence, none of these polymorphic sites contributed towards the progression of lung cancer among other organs as shown in table 5.7 (b).

Table 5.7 (a): Relationship of different genotypes of *DKK* gene with the clinico-pathological parameters (stage and primary tumor extension)

Genotype	Clinical Stage		AOR (95% CI) ^a	<i>p</i> ^b	Primary Tumor Extension		^b AOR (95% CI) ^b	<i>p</i> ^b
	III	IV			Tx+T1+T2	T3+T4		
<i>DKK4</i>, rs2073664 C/T	140	123			63	207		
CC	72 (51.4)	60 (48.8)	Reference		25 (39.7)	110 (53.1)	Reference	
CT	65 (46.4)	61 (49.6)	1.18 (0.71-1.96)	0.51	37 (58.8)	91 (44)	0.53 (0.29-0.97)	0.04
CT+TT	68 (48.6)	63 (51.2)	1.15 (0.70-1.91)	0.56	38 (60.3)	97 (46.9)	0.55 (0.30-0.99)	0.04
<i>DKK3</i>, rs2291599 C/T	140	123			63	207		
CC	92 (65.7)	78 (63.4)	Reference		40 (63.5)	138 (66.7)	Reference	
CT	33 (23.6)	38 (30.9)	1.32 (0.74-2.38)	0.33	19 (30.1)	51 (24.6)	0.82 (0.42-1.57)	0.55
CT + TT	48 (34.3)	45 (36.6)	1.04 (0.61-1.76)	0.87	23 (36.5)	69 (33.3)	0.92 (0.50-1.69)	0.80
<i>DKK3</i>, rs3206824 G/A	140	123			63	207		
GG	96 (68.6)	84 (68.3)	Reference		40 (63.5)	145 (70)	Reference	
GA	38 (27.1)	39 (31.7)	1.15 (0.66-2.01)	0.60	22 (35)	57 (27.5)	0.73 (0.39-1.35)	0.31
GA+AA	44 (31.3)	39 (31.7)	1.00 (0.58-1.72)	0.99	23 (36.5)	62 (30)	0.79 (0.43-1.45)	0.46
<i>DKK2</i>, rs419558 C/T	140	123			63	207		
CC	94 (67.1)	83 (67.5)	Reference		41 (65)	139 (67.1)	Reference	
CT	46 (32.9)	40 (32.5)	1.08 (0.63-1.85)	0.75	22 (35)	68 (32.9)	0.84 (0.46-1.54)	0.58
<i>DKK2</i>, rs17037102G/A	140	123			63	207		
GG	96 (68.5)	86 (70)	Reference		45 (71.4)	146 (70.5)	Reference	
GA	44 (31.5)	35 (28.5)	0.86 (0.49-1.49)	0.60	17 (27)	59 (28.5)	1.06 (0.56-2.03)	0.83
GA+AA	44 (31.5)	37 (30)	0.90 (0.52-1.55)	0.71	18 (28.6)	61 (29.5)	1.04 (0.55-1.96)	0.90

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*^b-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

Table 5.7 (b): Relationship of different genotypes of *DKK* gene with the clinico-pathological parameters (lymphnode invasion and metastasis)

Genotype	Lymph node Invasion		^b AOR (95% CI) ^b	<i>P</i>	Metastasis		^b AOR (95% CI) ^b	<i>p</i>
	N0	N1+N2+N3			No	Yes		
<i>DKK4</i>, rs2073664 C/T	48	220			154	116		
CC	27 (56.2)	106 (48.1)	Reference		78 (50.6)	58 (50)	Reference	
CT	19 (39.5)	110 (50)	1.38 (0.71-2.65)	0.33	72 (46.7)	55 (47.4)	1.11 (0.67-1.85)	0.67
CT+TT	21 (43.8)	114 (51.8)	1.31 (0.69-2.49)	0.39	76 (49.4)	58 (50)	1.11 (0.67-1.84)	0.67
<i>DKK3</i>, rs2291599 C/T	48	220			154	116		
CC	35 (73)	142 (64.5)	Reference		107 (69.5)	72 (62)	Reference	
CT	7 (14.6)	62 (28.1)	2.36 (0.97-5.74)	0.05	33 (21.4)	37 (31.9)	1.67 (0.93-3.02)	0.08
CT + TT	13 (27)	78 (35.5)	1.53 (0.75-3.10)	0.23	47 (30.5)	44 (38)	1.32 (0.77-2.24)	0.30
<i>DKK3</i>, rs3206824 G/A	48	220			154	116		
GG	33 (68.8)	150 (68.1)	Reference		105 (68.1)	80(69)	Reference	
GA	14 (29.1)	65 (29.5)	1.04 (0.51-2.09)	0.91	43 (28)	36(31)	1.08 (0.62-1.89)	0.76
GA+AA	15 (31.2)	70 (31.9)	1.03 (0.52-2.04)	0.92	49 (31.9)	36(31)	0.96 (0.55-1.65)	0.88
<i>DKK2</i>, rs419558 C/T	48	220			154	116		
CC	33 (69)	145 (66)	Reference		102 (66.2)	78 (67.2)	Reference	
CT	15 (31)	75 (34)	1.07 (0.54-2.12)	0.83	52 (33.8)	38 (32.8)	1.06 (0.62-1.80)	0.82
<i>DKK2</i>, rs17037102G/A	48	220			154	116		
GG	33 (68.8)	156 (71)	Reference		110 (71.5)	83 (71.5)	Reference	
GA	15 (31.2)	61 (27.7)	0.85 (0.43-1.70)	0.66	43 (28)	32 (27.6)	0.95 (0.54-1.67)	0.87
GA+AA	15 (31.2)	64 (29)	0.90 (0.46-1.79)	0.78	44 (28.5)	33 (28.5)	0.91 (0.52 -1.59)	0.75

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding p^b -values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

5.5.8 Multiple dimensionality reduction analysis of genetic variants within *DKK* gene:

Non parametric analysis was carried out using MDR approach so as to find out the interaction of these *DKK* variants and their role in predicting the lung cancer predisposition in north Indians. This approach reveals the possible interaction models comprising the genetic variants and environmental exposure which is smoking in this case. The present analysis was carried out taking seven variants of *DKK* gene and smoking as one of the factor in predicting lung cancer susceptibility. The analysis predicted three interaction models as summarized in table 5.8. Out of the all the polymorphic sites, one factor model comprised of *DKK2* rs419558 having a CVC of 7/10, prediction error= 0.47 and permutation testing $p=0.003$. The two factor model came out to be best interaction model, it consists of *DKK2* rs419558 and *DKK2* rs17037102. This model has the highest, 10/10 CVC along with 0.42 prediction error which is the lowest out of all the other models with a $p=0.003$. The analysis also gave a three factor model which included *DKK2* rs419558, *DKK3* rs3206824 and *DKK2* rs17037102, however this model showed the minimum CVC of 4/10 with the highest prediction error of 0.47 and $p=0.001$. Overall, *DKK2* rs419228 and *DKK2* rs17037102 together came out as the best model and contribute in modulating lung cancer susceptibility.

The entropy dendrogram illustrates the type of interaction taking place between the *DKK* variants which are the important factors of the models predicted by MDR as shown in fig 5.17. The length of the lines connecting these variants signify the intensity of interaction. The colors of the lines connecting these SNP's indicate towards the type of interaction taking place between them. *DKK2* rs419558 is connected to *DKK2* rs17037102 via blue line which indicates redundancy or no interaction.

Table 5.8: Multifactor Dimensionality reduction (MDR) analysis showing association of high order interactions of *DKK* variants with lung cancer

No. of risk factors	Best interaction model	Cross Validation Consistency (CVC)	Prediction error	<i>p</i> -value
1	<i>DKK2</i> rs419558	7/10	0.47	0.003
2	<i>DKK2</i> rs419558, <i>DKK2</i> rs17037102	10/10	0.42	0.003
3	<i>DKK2</i> rs419558, <i>DKK3</i> rs3206824, <i>DKK2</i> rs17037102	4/10	0.47	0.001

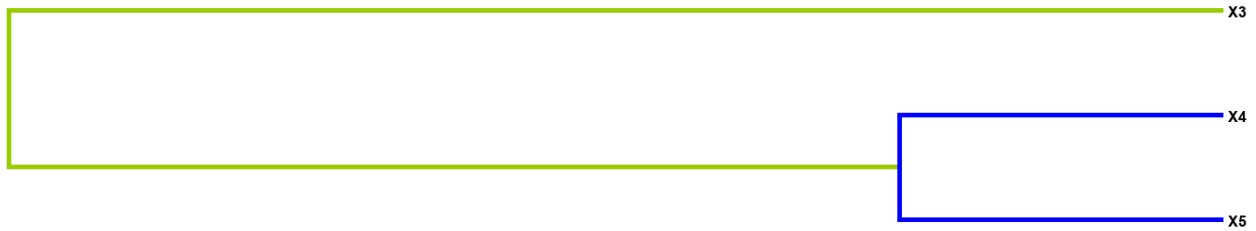


Fig 5.17: MDR interactions with DKK variants; X3: *DKK3* rs3206824 X4: *DKK2* rs419558, X5: *DKK2* rs17037102

5.5.9 Classification and regression tree analysis of genetic variants within *DKK* gene

In order to overcome the limitations of the traditional regression methods, CART was used for deciphering the different genotypic combinations pertaining to low and high risk subgroups. This analysis gave a decision tree where the node forming the first split is considered as the most significant factor in modulating lung cancer predisposition. Case rate was estimated for each terminal node which depicts the percentage of the lung cancer cases among the subjects falling under each terminal node. The terminal node having the lowest case rate is used as the reference node, for estimating the odds ratio associated with other terminal nodes which is adjusted for age, gender and smoking status.

When, all the five variants of *DKK* gene are evaluated through CART analysis, only three terminal nodes were observed as shown in fig 5.18. The adjusted odds ratio were calculated for each terminal node using terminal node 3 (*DKK2*rs17037102 (W)/ *DKK2*rs419558 (W)) as the reference node as it has the lowest case rate of 45.15. Also, *DKK2* rs419558 came out to be most significant factor as it formed the first split and it is the root node in the decision tree. Out of the three terminal nodes, two terminal nodes namely 1 and 2 were characterized as high risk subgroups. These genotypic combinations exhibited a moderately increased risk for lung cancer with significant *p* values as shown in table 5.9. The patients having the genotypic combination of *DKK2*_{rs419558} (M) showed a significant 2 fold risk of acquiring lung cancer. Terminal node 2 also showed an increased risk of lung cancer which was found to be statistically significant.

Table 5.9: Risk estimates based on CART analysis for *DKK* gene

Terminal Node	Genotype for each terminal node	Control(n)	Case(n)	Case Rate ^a	AOR (95% CI)		p ^b
Node 1	<i>DKK2</i> rs419558 (M)	56	95	62.9	2.06	(3.07-1.37)	0.0004
Node 2	<i>DKK2</i> rs17037102 (M)/ <i>DKK2</i> rs419558 (W)	43	62	59.04	1.75	(2.74-1.11)	0.01
Node 3	<i>DKK2</i> rs17037102 (W)/ <i>DKK2</i> rs419558 (W)	164	135	45.15	1.0 ^c		

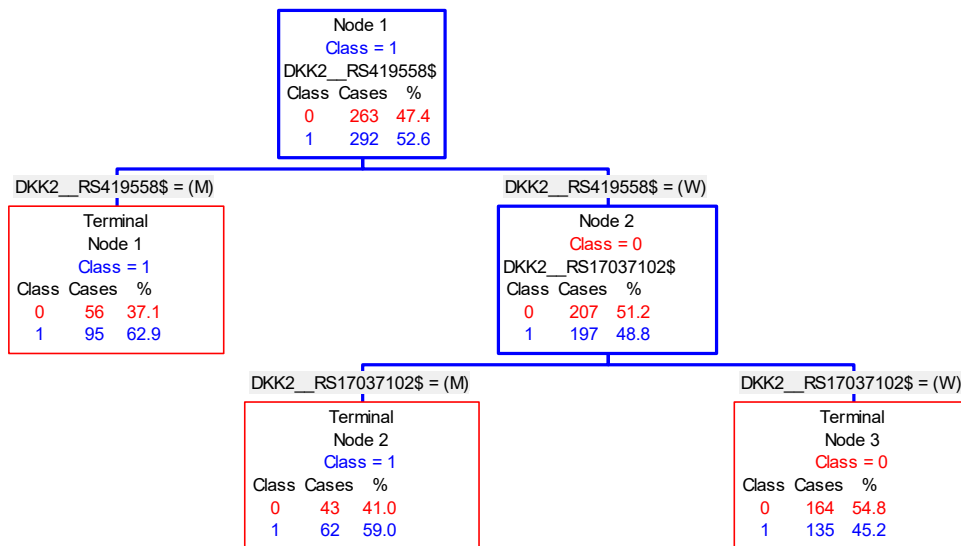


Fig 5.18: CART analysis of *DKK* variants; W: wild type; M: variant type. Red line colour indicates terminal nodes.

5.5.10 Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of *DKK* polymorphism

Table 5.10 summarizes the clinical parameters of the study subjects considered for survival analysis. Out of 212 subjects, 86.8% were found to be males and the remaining (13.2%) were females. As smoking is an established cause for occurrence of lung cancer, so majority of sample population comprised of smokers (80.6%), while 19.4% were characterized as non-smokers. Different clinical parameters such as histology, TNM staging, performance status and chemotherapeutic regimen are also listed in table 5.10. The patients are classified on the basis of their histology into ADCC, SQCC and SCLC. Considering the TNM staging data, maximum proportion (49.5%) of lung cancer patients were diagnosed with stage III of lung cancer. Stage IV patients constituted 41.9% of all the lung cancer patients. In order to evaluate the performance status of the patients, two criteria's namely ECOG and KPS were used. 133 (62.7%) patients were found to have KPS score of 60-80 and 74 (34.9%) patients harbored an ECOG-PS value of 1 and 2. Further, we stratified the sample population on the basis of their chemotherapeutic regimen also, highest percentage (35.2%) of patients were treated with docetaxel and cisplatin doublet combination. SCLC patients were given irinotecan and cisplatin doublet combination constituting 26.2% of the total. Also, the patients receiving paclitaxel and cisplatin were 7.6% and those treated with pemetrexed and cisplatin were 25.6%. Hazards ratio was also calculated for various parameters such as gender, smoking, TNM staging and performance status. It was evident that none of these factors could act as an independent predictor of survival for lung cancer patients. Lung cancer patients having ECOG (1) were found to have a better survival (HR=0.66) as compared those with ECOG (0) but the *p*-value (0.07) was not found to be statistically significant.

5.5.11 Association of *DKK* genetic variants with overall survival of lung cancer patients

Table 5.12 summarizes the univariate and multivariate analysis for the various polymorphic sites of *DKK* gene in relation to overall survival of patients. Median survival time (MST) was calculated using kaplan meir statistics and survival curves were obtained for each genotype. Genotypic influence on the survival of the patients was examined by using log rank tests which compares the MST of different groups taking the wild genotype as reference group. Univariate analysis gave the probability of death (Hazards ratio) whereas multivariate cox regression analysis provided the adjusted hazards ratio. Hazard ratio was adjusted using the cox regression model for various co-variates such as age, sex, smoking, histology, stage, KPS and ECOG in order to find whether a particular SNP can act as an independent predictor and prognostic factor.

Table 5.10: Distribution of Demographic and Clinical parameters

Variables	Total(N)	Cases, n (%)	HR (95% C.I) ^a	<i>p</i>
Age				
Mean ± SD	212	57.87 ± 10.81		
Range		28-86		
Gender				
Male	212	184 (86.8)	0.93 (0.61-1.40)	0.74
Female		28 (13.2)	Reference	
Smoking status				
Smokers	212	171 (80.6)	1.04 (0.72-1.49)	
Non-smokers		41 (19.4)	Reference	0.83
Histology				
ADCC	212	72 (34)		
SCLC		51 (24)		
SQCC		86 (40)		
Others		3 (2)		
TNM Staging				
I		2 (0.9)		
II	212	6 (2.8)		
III		105 (49.5)	Reference	
IV		89 (41.9)	1.23 (0.92-1.65)	0.16
Others		10 (4.7)		
Overall Survival				
Dead		181 (85.3)		
Alive	212	31 (14.6)		
Performance Status				
KPS(30-50)		7 (3.3)	Reference	
KPS(60-80)	212	133 (62.7)	0.66 (0.29-1.49)	0.32
KPS(90-100)		72 (33.9)	0.80 (0.52-1.21)	0.30
ECOG(0)		31 (14.6)	Reference	
ECOG(1)		74 (34.9)	0.66 (0.42-1.03)	0.07
ECOG(2)	212	74 (34.9)	0.84 (0.68-1.05)	0.14
ECOG(3 and 4)		33 (15.5)	0.94 (0.79-1.12)	0.53
Chemotherapy Regimen				
Docetaxel/ Cisplatin		55 (35.2)		
Irinotecan/ Cisplatin		41 (26.2)		
Paclitaxel/ Cisplatin	156	12 (7.6)		
Pemetrexed/ Cisplatin		40 (25.6)		
Gemacitabine/ Cisplatin		1(0.6)		
Others		7		

Abbreviations: SD = Standard Deviation, n = total number of case patients. All *p*-values are two – sided. *p* < 0.05 was considered statistically significant. HR (95% CI)^a depicts hazards ratio with 95% C.I.

On the basis of the survival status of the patients in case of *DKK4* rs2073664, 42.5% patients who were heterozygous for *DKK4* rs2073664 were found to be dead and 61.3% were alive as shown in table 5.11. It was observed that MST of lung cancer patients having heterozygous genotype for *DKK4* rs2073664 was

higher (7.5 months) as compared to wild (6.2 months) as shown in fig 5.19 (a) and mutant (6.5 months) genotype. Better OS was observed for heterozygous patients both in univariate (HR=0.74) and multivariate (HR'=0.75) analysis, but both lacked statistical significance.

None of the three variants (rs2295199, rs3206824, rs7396187) of *DKK3* gene showed any significant role in predicting the survival of lung cancer patients as shown in table 5.11. In case of *DKK3* rs2295199, the percentage of dead and alive patients were 37% and 38.7% respectively in mutant category. The difference in MST in each genotypic category was not found to be significant, although the heterozygous subjects showed a better survival time of 7.5 months as compared to 6.4 months in case of wild subjects. In case of *DKK3* rs3206824, the percentage of dead and alive was 33.1% and 38.7% for subjects having GA+AA genotype. The lung cancer subjects harboring the heterozygous (CT) genotype did have a higher MST (8.8 months, log rank $p=0.07$) with hazard ratio of 0.74 although on adjusting with various covariates, no significant change in death probability was observed (HR'=0.76, 95% C.I.:0.54-1.06, $p=0.10$) as shown in fig 5.19 (B). Out of all the subjects carrying the variant genotype for *DKK3* rs7396187, 29.3% were dead and 32.3% were alive. Median survival time of subjects carrying wild and heterozygotes genotype was 6.3 and 7.8 months respectively. Log rank p was found to be 0.74, therefore it is evident that the overall survival of patients is not associated with genotype of this polymorphic site. A similar trend was evident when hazard's ratio was calculated after adjusting for various confounding factors such as age, gender, smoking, stage, histology, KPS and ECOG.

The three polymorphic sites of *DKK2* gene *i.e.* rs447372, rs17037102, rs419558 also exhibited a lack of association with the prognosis of North Indian lung cancer patients as shown in table 5.11. In case of rs447372, the percentage of dead and alive was 49 and 58% respectively among the subjects carrying AG+GG genotype. On comparing, the MST of each genotypic group it was seen that highest MST (10.3 months) as shown in fig 5.19 (C) was possessed by patients having mutant (GG) genotype, followed by heterozygous group having 8.2 months and least MST of 5.9 months was observed in case of subjects having wild genotype. However, the difference was not found to be significant (log rank $p=0.12$). On the contrary, this polymorphic site did show a slightly good prognostic effect in patients carrying AG+GG genotype which showed a lower hazard's ratio after adjusting (HR'=0.73, 95% C.I.:0.54-0.99, $p=0.04$). Further, in case of exonic polymorphic site *i.e.* rs17037102 the proportion of dead was 29.7 and 19.4% were found to be alive among the subjects carrying combined variant (GA+AA) genotype. Non-significant difference in MST was observed among all the categories of the genotypic distribution of this polymorphic site. Subjects falling under wild category had MST of 6.7 months and those having heterozygous GA genotype showed the survival time of 6.4 months. It was evident that this genetic variant did not affect the death probability of lung cancer patients not even after applying multivariate cox

regression approach. Also, for rs419558 the percentage of dead and alive was 31 and 42% respectively among heterozygous subjects for this site. MST of patients having wild genotype was 6.3 months slightly lower as compared to heterozygotes (7.3 months). Also, the hazard's ratio obtained after both univariate and multivariate analysis were found to lack statistical significance.

On the basis of the survival status of the patients in case of *DKK4* rs2073664, 42.5% patients who were heterozygous for *DKK4* rs2073664 were found to be dead and 61.3% were alive as shown in table 5.11. It was observed that MST of lung cancer patients having heterozygous genotype for *DKK4* rs2073664 was higher (7.5 months) as compared to wild (6.2 months) as shown in fig 5.19 (A) and mutant (6.5 months) genotype. Better OS was observed for heterozygous patients both in univariate (HR=0.74) and multivariate (HR'=0.75) analysis, but both lacked statistical significance.

None of the three variants (rs2295199, rs3206824, rs7396187) of *DKK3* gene showed any significant role in predicting the survival of lung cancer patients as shown in table 5.11. In case of *DKK3* rs2295199, the percentage of dead and alive patients were 37% and 38.7% respectively in mutant category. The difference in MST in each genotypic category was not found to be significant, although the heterozygous subjects showed a better survival time of 7.5 months as compared to 6.4 months in case of wild subjects. In case of *DKK3* rs3206824, the percentage of dead and alive was 33.1% and 38.7% for subjects having GA+AA genotype. The lung cancer subjects harboring the heterozygous (CT) genotype did have a higher MST (8.8 months, log rank $p=0.07$) with hazard ratio of 0.74 although on adjusting with various covariates, no significant change in death probability was observed (HR'=0.76, 95% C.I.:0.54-1.06, $p=0.10$) as shown in fig 5.19 (b). Out of all the subjects carrying the variant genotype for *DKK3* rs7396187, 29.3% were dead and 32.3% were alive. Median survival time of subjects carrying wild and heterozygotes genotype was 6.3 and 7.8 months respectively. Log rank p -value was found to be 0.74, therefore it is evident that the overall survival of patients is not associated with genotype of this polymorphic site. A similar trend was evident when hazard's ratio was calculated after adjusting for various confounding factors such as age, gender, smoking, stage, histology, KPS and ECOG.

The three polymorphic sites of *DKK2* gene *i.e.* rs447372, rs17037102, rs419558 also exhibited a lack of association with the prognosis of North Indian lung cancer patients as shown in table 5.11. In case of rs447372, the percentage of dead and alive was 49 and 58% respectively among the subjects carrying AG+GG genotype. On comparing, the MST of each genotypic group it was seen that highest MST (10.3 months) as shown in fig 5.19 (C) was possessed by patients having mutant (GG) genotype, followed by heterozygous group having 8.2 months and least MST of 5.9 months was observed in case of subjects having wild genotype. However, the difference was not found to be significant (log rank $p=0.12$). On the

contrary, this polymorphic site did show a slightly good prognostic effect in patients carrying AG+GG genotype which showed a lower hazard's ratio after adjusting (HR'=0.73, 95% C.I.:0.54-0.99, $p=0.04$). Further, in case of exonic polymorphic site *i.e.* rs17037102 the proportion of dead was 29.7 and 19.4% were found to be alive among the subjects carrying combined variant (GA+AA) genotype. Non-significant difference in MST was observed among all the categories of the genotypic distribution of this polymorphic site. Subjects falling under wild category had MST of 6.7 months and those having heterozygous GA genotype showed the survival time of 6.4 months. It was evident that this genetic variant did not affect the death probability of lung cancer patients not even after applying multivariate cox regression approach. Also, for rs419558 the percentage of dead and alive was 31 and 42% respectively among heterozygous subjects for this site. MST of patients having wild genotype was 6.3 months slightly lower as compared to heterozygotes (7.3 months). Also, the hazard's ratio obtained after both univariate and multivariate analysis were found to lack statistical significance.

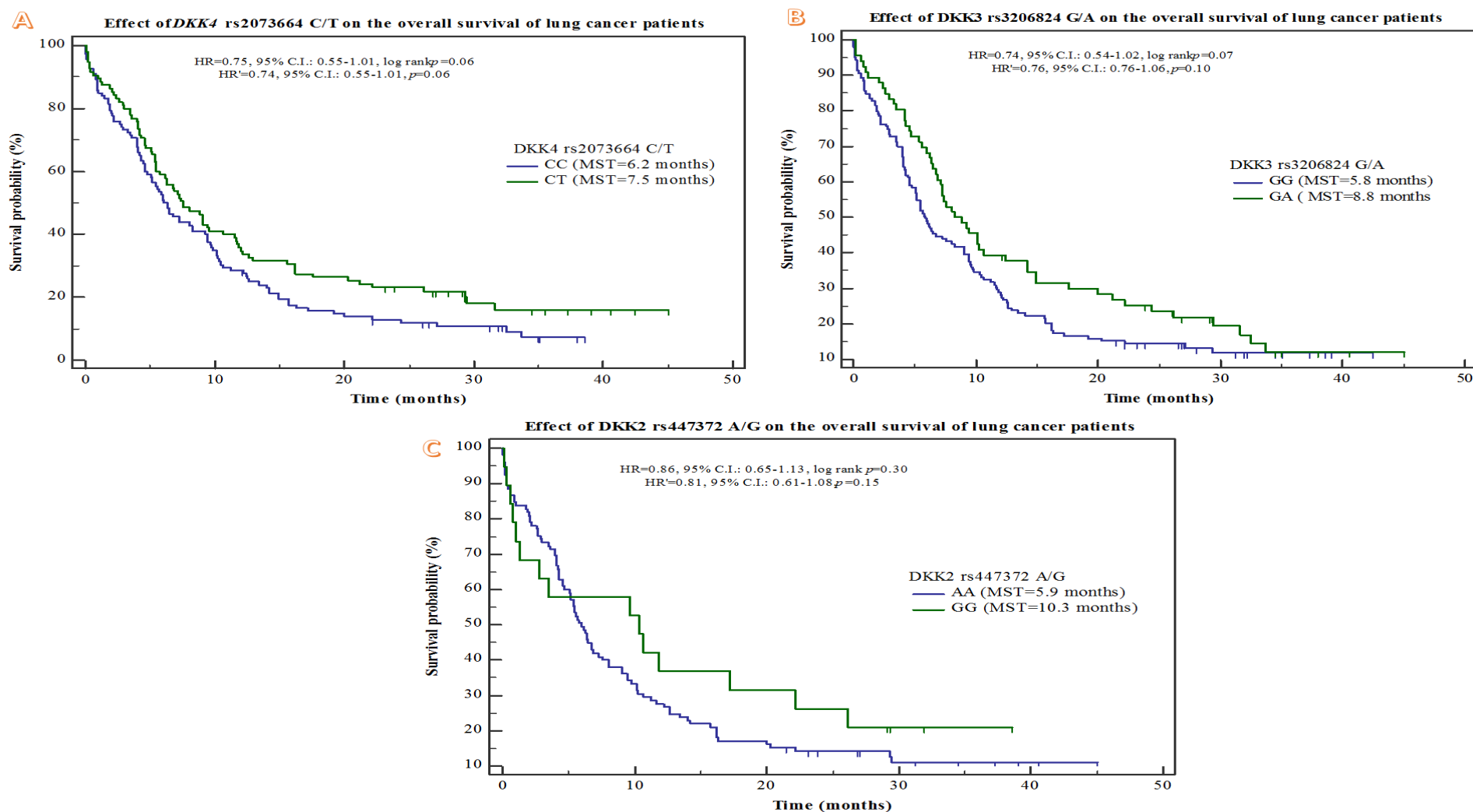


Fig 5.19: Kaplan meir survival curves for *DKK* variants; A: *DKK4* rs2063664 C/T (CC: reference), B: *DKK3* rs3206824 G/A (GG: reference), C: *DKK2* rs447372 A/G (AA: reference) for lung cancer patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time.

Table 5.11: Univariate and Multivariate analysis for Wnt antagonist (*DKK*)

Genotype	CASES n (%), N=212	DEAD n (%), N=181	ALIVE n (%), N=31	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank <i>p</i>	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	<i>p</i>
<i>DKK4</i> , rs2073664 C/T								
CC	112 (52.8)	101(55.8)	11 (35.5)	6.2		Reference	Reference	
CT	96 (45.3)	77 (42.5)	19 (61.3)	7.5	0.06	0.75 (0.55-1.01)	0.74 (0.55-1.01)	0.06
TT	4 (1.9)	3 (1.7)	1 (3.2)	6.5	0.74	0.97 (0.51-1.60)	1.01 (0.56-1.85)	0.95
CT+TT	100 (47.2)	80 (44.2)	20 (64.5)	7.5	0.07	0.76 (0.57-1.02)	0.79 (0.58-1.06)	0.12
<i>DKK3</i> , rs2291599 C/T								
CC	133 (62.7)	114 (63)	19 (61.3)	6.4		Reference	Reference	
CT	63 (29.7)	52(28.7)	11 (35.5)	7.5	0.52	0.89 (0.64-1.24)	0.92 (0.65-1.29)	0.64
TT	16 (7.5)	15 (7)	1 (3.2)	6.7	0.98	1 (0.76-1.31)	1.07 (0.81-1.43)	0.60
CT+TT	79 (37.3)	67 (37)	12 (38.7)	7.3	0.55	0.91 (0.67-1.23)	0.96 (0.70-1.30)	0.96
<i>DKK3</i> , rs3206824 G/A								
GG	140 (66)	121 (66.9)	19 (61.3)	5.8		Reference	Reference	
GA	66 (31.1)	55 (3)	11 (35.5)	8.8	0.07	0.74 (0.54-1.02)	0.76 (0.54-1.06)	0.10
AA	6 (2.8)	5 (2.8)	1 (3.2)	4.9	0.91	1.02 (0.65-1.60)	1.07 (0.67-1.72)	0.76
GA+AA	72 (34)	60 (33.1)	12 (38.7)	8.3	0.11	0.78 (0.58-1.05)	0.81 (0.58-1.12)	0.20
<i>DKK3</i> , rs7396187 G/C								
GG	149 (70.3)	128 (70.7)	21 (67.7)	6.3	0.74	Reference	Reference	
GC	62 (29.2)	52 (28.7)	10 (32.3)	7.8		0.89 (0.65-1.23)	0.89 (0.64-1.23)	0.49
CC	1 (0.4)	1 (0.5)	0	16.7		...	...	
GC+CC	63 (29.7)	53 (29.3)	10 (32.3)	8	0.47	0.89 (0.65-1.23)	0.87 (0.63-1.20)	0.87
<i>DKK2</i> , rs447372 A/G								
AA	105 (49.5)	92 (51)	13 (42)	5.9		Reference	Reference	
AG	87 (41)	74 (40.8)	13 (42)	8.2	0.22	0.82 (0.61-1.12)	0.77 (0.56-1.05)	0.11
GG	20 (9.4)	15 (8.2)	5 (16.1)	10.3	0.30	0.86 (0.65-1.13)	0.81 (0.61 -1.08)	0.15
AG+GG	107 (50.4)	89 (49)	18 (58)	8.3	0.12	0.79 (0.59-1.06)	0.73 (0.54-0.99)	0.04
<i>DKK2</i> , rs17037102 G/A								
GG	152 (71.7)	127 (70.3)	25 (80.6)	6.7	0.78	Reference	Reference	
GA	59 (27.8)	53 (29.2)	6 (19.4)	6.4		1.11 (0.81-1.54)	1.09 (0.78-1.53)	0.59
AA	1 (0.4)	1 (0.5)	0	10.6		...	...	
GA+AA	60 (28.3)	54 (29.7)	6 (19.4)	6.8	0.31	1.17 (0.84-1.63)	1.13 (0.81-1.58)	0.45
<i>DKK2</i> , rs419558 C/T								
CC	143 (67.5)	125 (69)	18 (58)	6.3		Reference	Reference	
CT	69 (32.5)	56 (31)	13 (42)	7.3	0.38	0.87 (0.64-1.18)	0.82 (0.60-1.13)	0.24

^aUnadjusted Hazards ratio for Kaplan meier analysis, ^bhazards ratio for age, sex, smoking, histology, stage, KPS, ECOG

5.5.12 Association of Wnt antagonist (*DKK*) polymorphism and overall survival according to histological subtype of lung cancer patients

In order to evaluate the relevance of these polymorphic sites in predicting the survival of patients in different subgroups stratified on the basis of histology, the association of *DKK* variants with survival of patients was analysed and data was tabulated in table 5.13. The data for each histological subtype has been summarized in tables given below. The genotypic distribution and overall survival of ADCC patients has been tabulated in table 5.12 (a), SQCC patients in table 5.12 (b) and SCLC patients in table 5.12 (c).

In case of *DKK4* rs2073664, the percentage of heterozygous patients were 45.9%, 38.4% and 53% among ADCC, SQCC and SCLC subtype. It was found that ADCC patients carrying the heterozygous (CT) genotype showed a better OS as compared to wild genotype (8.8 months vs 6.3 months). Hazards ratio obtained from both univariate and multivariate analysis was 0.53 with significant *p* values of 0.01 and 0.03 as shown in fig 5.20 (A). Also, the ADCC subjects carrying CT+TT genotype showed a MST of 7.5 months pertaining to a lower hazards ratio in case of univariate analysis (HR=0.58, 95% C.I.:0.35-0.96, log rank *p*=0.03) and multivariate analysis (HR'=0.61, 95% C.I.: 0.35-1.06, *p*=0.08). However, the ADCC patients having mutant (TT) genotype showed the least MST (2.8 months, log rank *p*=0.19) and three fold increased death probability was also observed in these patients (HR'=3.30, 95% C.I.: 1.02-10.60, *p*=0.04) as shown in table 5.12 (a). On the contrary, the MST of SQCC subjects carrying mutant (TT) genotype was 10.1 months that was higher as compared to 5.5 months as in case of subjects having wild (CC) genotype. However, the hazards ratio obtained after both univariate and multivariate analysis was not significantly associated with the genotype. Similarly, SCLC patients having heterozygous genotype showed a slightly better OS as compared to wild type (MST=8.03 months vs 7.3 months, log rank *p*=0.71) and this polymorphic site was found to act as an independent predictor of survival only in case of ADCC patients and not among SQCC and SCLC patients.

Out of the three *DKK3* polymorphic sites, the two SNPs namely rs2291599 and rs7396187 were found to play no role in predicting the prognosis of patients in any of the histological subtypes. The proportion of subjects having CT+TT genotype for rs2291599 was 33.3% in ADCC subtype, 44.2% in SQCC and 31.4% in SCLC subtype. Overall survival analysis in case of rs2291599 revealed that ADCC subjects carrying heterozygous (CT) and CT+TT genotype both had a better OS with MST of 8.3 months as compared to wild type (6.1 months) as shown in table 5.12 (a). However, the difference between MST of these genotypic categories was not found to be significant. Similar trend was observed in case of SQCC subjects having CT+TT genotype where they exhibited a MST of 7.5 months nearly equivalent to wild type patients (6.9 months). However, in case of SCLC patients a lower MST was observed in case of CT+TT genotype carrying subjects (5.8 months) as compared to wild type (8.2 months). Therefore, this

genotypic variant did not act as an independent predictor of survival in case of any of the histological subtypes.

Further, in case of rs3206824 the percentage of GA+AA subjects was 32%, 39.5% and 27.5% among ADCC, SQCC and SCLC category. The overall survival analysis showed that among the ADCC patients, the highest MST was observed in case of heterozygous (GA) subjects (9.2 months) as illustrated in fig 5.20 (B) followed by GA+AA genotype (8.8 months) and 5.9 months in case of wild (GG) type. The least MST was seen in mutant subjects (4.5 months). However, no significant difference was observed among MST of all genotypic classes. On the similar lines, SQCC patients also lacked to a significant association between this polymorphic site and overall survival as shown in table 5.12 (b). Median survival time of individuals carrying either GG, GA, AA and GA+AA was found to be 5.1, 8.3, 7.7 and 9.4 months respectively but the difference was not found to be statistically significant as shown in fig 5.20 (C). Only SCLC patients who were harboring the mutant (AA) genotype for *DKK3* rs3206824 had a lower MST as compared to the wild type genotype (1.8 vs 8.0 months, log rank $p=0.14$) and after applying cox regression SCLC patients exhibited a three-fold increased death ratio ($HR'=3.45$, 95% C.I.: 1.07-11.09, $p=0.03$) as shown in fig 5.20 (D) when compared to the SCLC subjects carrying the wild (GG) genotype.

In case of next variant *DKK3* gene, rs7396187 it was observed that 29%, 31.4% and 25.5% heterozygous subjects were reported among ADCC, SQCC and SCLC categories respectively. Kaplan meier analysis estimated the MST of patients among each category. In case of ADCC patients, the heterozygous (GC) subjects showed a higher MST of 9.2 months as compared to 6.1 months as in case of wild (CC) type subjects as shown in table 5.12 (a). However, the difference was not statistically significant as shown fig 5.20 (E). SQCC patients having heterozygous and wild type genotype were estimated to have a MST of 6.2 and 6.4 months respectively ($HR'=1.13$, 95% C.I.: 0.66-1.96, $p=0.64$). Similar trend was also evident among SCLC patients as the heterozygotes showed almost comparable MST when compared to wild type (6.7 months vs 7.3 months, log rank $p=0.90$). Hence, it was evident that this SNP was unable to predict the prognosis of the lung cancer patients with any histological subtype.

The three genetic variants of *DKK2* gene were also studied for their association with overall survival in subgroups classified on the basis of histology. In case of rs447372, the subjects harboring AG +GG genotype were 57%, 38.4% and 56.9% in ADCC, SQCC and SCLC subtype respectively. Survival analysis reported that MST of mutant (GG) genotype carrying ADCC patients was highest (10.8 months) and subjects carrying wild (AA) genotype was 6.1 months, which was the lowest as shown in table 5.12 (a). The log rank test did not reveal any significant association between these genotypes and OS of ADCC subjects. In case of SQCC patients, a higher MST was observed in case of subjects carrying GG genotype as compared to wild type subjects (13.4 months vs 5.4 months, log rank $p=0.27$) as shown in fig 5.20 (F). On the contrary, SCLC subjects carrying GG genotype had a lower MST as compared to wild type subjects (5.73 months vs 9 months, log rank $p=0.24$) but a non-significant hazards ratio was observed in

case of both univariate (HR=1.83, 95% C.I.: 0.47-7.03) and multivariate analysis (HR'=1.24, 95% C.I.: 0.69-2.25, $p=0.46$) as shown in table 5.12 (c).

The next variant of *DKK2* rs17037102 did not contribute towards predicting the survival of lung cancer patients. ADCC subjects having heterozygous genotype for this polymorphic site reported a better survival of 8.8 months as compared to wild type (6.1 months) subjects. However, no significant association was observed as the hazards ratio was not found to be significant (HR'=1.02, 95% C.I.: 0.99-1.05, $p=0.07$) as shown in table 5.12 (a). On the contrary, in case of SQCC patients MST of heterozygotes was lower than wild type (6.2 months vs 9.4 months, log rank $p=0.21$). Similar trend was observed in case of SCLC patients where the heterozygous (GA) patients showed a MST of 5.8 months lower than the MST of wild type subjects (8 months) as shown in table 5.12 (c). There was only a trend that was evident as that these observations lacked statistical validation. Hence, no significant relationship was observed between this genotype and prognosis of the lung cancer patients possessing any histological subtype.

Further, it was observed that only ADCC patients having heterozygous (CT) genotype for *DKK2* rs419558 showed a better MST as compared to as compared to ADCC patients carrying wild type (CC) genotype (8.3 months vs 6.2 months, log rank $p=0.11$). However, on further adjusting with various covariates the HR was found to be 0.50 ($p=0.03$) as shown in table 5.12 (a). Hence, the heterozygous ADCC patients showed a good prognosis as compared to wild (CC) genotype. No such association was observed in case of SQCC and SCLC patients as the differences in MST of the genotypic categories were not statistically significant.

Table 5.12 (a): Association of *DKK* SNPs on overall survival according to tumor histology (ADCC)

Genotype	ADCC					
<i>DKK4</i>, rs2073664 C/T	Cases N (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	37 (51.4)	6.3		Reference	Reference	
CT	33 (45.9)	8.8	0.01	0.53 (0.32-0.89)	0.53 (0.29-0.94)	0.03
TT	1 (1.4)	2.8	0.19	3.40 (0.09-123)	3.30 (1.02-10.6)	0.04
CT+TT	34 (47.2)	7.5	0.03	0.58 (0.35-0.96)	0.61 (0.35-1.06)	0.08
<i>DKK3</i>, rs2291599 C/T	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	47 (65.3)	6.1		Reference	Reference	
CT	21 (29.2)	8.3	0.75	0.91 (0.52-1.58)	0.80 (0.42-1.53)	0.52
TT	3 (4.1)	7.13	0.68	1.27 (0.34-4.70)	1.30 (0.69-2.44)	0.40
CT+TT	24 (33.3)	8.3	0.80	0.93 (0.55-1.58)	0.88 (0.48-1.61)	0.69
<i>DKK3</i>, rs3206824 G/A	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
GG	49 (68)	5.9		Reference	Reference	
GA	22 (30.5)	9.2	0.28	0.74 (0.44-1.25)	0.79 (0.44-1.43)	0.44
AA	1 (1.4)	4.5	0.40	2.25 (0.11-42.17)	1.40 (0.48-4.09)	0.53
GA+AA	23 (32)	8.8	0.37	0.78 (0.46-1.31)	0.85 (0.47-1.51)	0.58
<i>DKK3</i>, rs7396187 G/C	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
GG	51 (70.9)	6.1		Reference	Reference	
GC	20 (29.1)	9.2	0.23	0.72 (0.42-1.22)	0.56 (0.31-1.03)	0.06
CC	---	---	---	---	---	---
GC+CC	---	---	---	---	---	---
<i>DKK2</i>, rs447372 A/G	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
AA	31 (43)	6.1		Reference	Reference	
AG	32 (44.4)	8.5	0.43	0.80 (0.47-1.38)	0.67 (0.38-1.19)	0.18
GG	8 (12.5)	10.3	0.53	0.77 (0.35-1.68)	0.71 (0.44-1.14)	0.16
AG+GG	40(57)	8.8	0.25	0.74 (0.44-1.26)	0.63 (0.36-1.09)	0.10
<i>DKK2</i>, rs17037102 G/A	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
GG	52 (72.4)	6.1		Reference	Reference	
GA	18 (26.3)	8.8	0.76	0.91 (0.52-1.59)	1.02 (0.99-1.05)	0.07
AA	1 (1.3)	10.6	0.88	0.86 (0.13-5.47)	0.79 (0.28-2.17)	0.64
GA+AA	19 (27.6)	8.8	0.91	0.97 (0.56-1.67)	0.85 (0.46-1.55)	0.60
<i>DKK2</i>, rs419558 C/ T	Cases n (%) N=71	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	52 (72.2)	6.3		Reference	Reference	
CT	19 (27.8)	8.3	0.11	0.62 (0.36-1.07)	0.50 (0.26-0.94)	0.03

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.12 (b): Association of *DKK* SNPs on overall survival according to tumor histology (SQCC)

Genotype	SQCC					
	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
<i>DKK4</i>, rs2073664 C/T						
CC	50 (58.1)	5.5		Reference	Reference	
CT	33 (38.4)	7.2	0.30	0.77 (0.47-1.2)	0.78 (0.47-1.28)	0.33
TT	3 (3.5)	10.1	0.64	0.71 (0.21-2.4)	0.99 (0.45-2.17)	0.99
CT+TT	36 (41.9)	7.1	0.28	0.76 (0.47-1.23)	0.79 (0.48-1.28)	0.34
<i>DKK3</i>, rs2291599 C/T	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	48 (55.8)	5.5		Reference	Reference	
CT	30 (34.9)	6.9	0.65	0.89 (0.53-1.48)	0.84 (0.49-1.46)	0.56
TT	8 (9.3)	8.2	0.72	0.86 (0.40-1.86)	0.84 (0.52-1.36)	0.50
CT+TT	38 (44.2)	7.5	0.54	0.86 (0.53-1.38)	0.82 (0.50-1.36)	0.45
<i>DKK3</i>, rs3206824 G/A	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
GG	52 (60.5)	5.1		Reference	Reference	
GA	30 (34.9)	8.3	0.17	0.70 (0.43-1.15)	0.63(0.37-1.07)	0.09
AA	4 (4.7)	7.7	0.67	0.77 (0.27-2.22)	0.79(0.41-1.54)	0.50
GA+AA	34 (39.5)	9.4	0.15	0.70 (0.44-1.13)	0.64(0.38-1.07)	0.09
<i>DKK3</i>, rs7396187 G/C	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
GG	59 (68.6)	6.4		Reference	Reference	
GC	27 (31.4)	6.2	0.72	1.09 (0.65-1.83)	1.13(0.66-1.96)	0.64
CC	---	---	---	---	---	---
GC+CC	---	---	---	---	---	---
<i>DKK2</i>, rs447372 A/G	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
AA	53 (61.6)	5.4		Reference	Reference	
AG	27 (31.4)	12.1	0.10	0.64 (0.39-1.06)	0.65 (0.37-1.14)	0.13
GG	6 (7)	13.4	0.27	0.57 (0.29-1.57)	0.77 (0.45-1.30)	0.33
AG+GG	33 (38.4)	11.5	0.05	0.61 (0.38-0.99)	0.61 (0.36-1.04)	0.07
<i>DKK2</i>, rs17037102 G/A	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
GG	60(69.8)	9.4		Reference	Reference	
GA	26(30.2)	6.2	0.21	1.36 (0.80-2.32)	1.40 (0.82-2.38)	0.20
AA	---	---	---	---	---	---
GA+AA	---	---	---	---	---	---
<i>DKK2</i>, rs419558 C/ T	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	58(67.4)	6.2		Reference	Reference	
CT	28(32.6)	6.8	0.70	0.90(0.55-1.49)	0.86(0.50-1.46)	0.58

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.12 (c): Association of *DKK* SNPs on overall survival according to tumor histology (SCLC)

Genotype	SCLC					
<i>DKK4</i>, rs2073664 C/T	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
CC	24(47)	7.3		reference	Reference	
CT	27(53)	8.03	0.71	0.90(0.50-1.59)	0.91(0.47-1.74)	0.78
TT	--	--	--	--	--	--
CT+TT	--	--	--	--	--	--
<i>DKK3</i>, rs2291599 C/T	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
CC	35 (68.6)	8.2		Reference	Reference	
CT	11 (21.6)	5.8	0.60	1.20 (0.57-2.53)	1.55 (0.67-1.56)	0.30
TT	5 (9.8)	4.1	0.95	1.02(0.39-2.64)	1.12 (0.60-2.10)	0.70
CT+TT	16 (31.4)	5.8	0.57	1.18(0.63-2.23)	1.48 (0.74-2.95)	0.25
<i>DKK3</i>, rs3206824 G/A	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
GG	37 (72.5)	7.2		Reference	Reference	
GA	13 (25.5)	8.0	0.60	0.84(0.44-1.58)	0.91 (0.45-1.87)	0.81
AA	1 (2)	1.8	0.14	3.84(0.08-175)	3.45 (1.07-11.09)	0.03
GA+AA	14 (27.5)	8.0	0.70	1.12(0.58-2.16)	1.35 (0.67-2.73)	0.40
<i>DKK3</i>, rs7396187 G/C	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
GG	37 (72.5)	7.3		Reference	Reference	
GC	13 (25.5)	6.7	0.90	0.96 (0.50-1.84)	1.19(0.55-2.58)	0.65
CC	1 (2)	16.1	0.41	0.45(0.11-1.84)	0.52(0.18-1.51)	0.23
GC+CC	14(27.5)	8.0	0.83	0.93(0.5-1.75)	1.07(0.51-2.24)	0.85
<i>DKK2</i>, rs447372 A/G	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
AA	22 (43.1)	9		Reference	Reference	
AG	25 (49)	7.23	0.75	1.09 (0.60-1.97)	0.98 (0.46-2.09)	0.96
GG	4 (7.8)	5.73	0.24	1.83 (0.47-7.03)	1.24 (0.69-2.25)	0.46
AG+GG	29 (56.9)	7.1	0.58	1.16 (0.66-2.06)	1.35 (0.67-2.68)	0.39
<i>DKK2</i>, rs17037102 G/A	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
GG	37 (72.5)	8.0		Reference	Reference	
GA	14 (27.5)	5.8	0.90	1.04 (0.54-1.97)	0.92 (0.45-1.88)	0.83
<i>DKK2</i>, rs419558 C/ T	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	H.R^b	<i>P</i>
CC	32 (62.7)	7.2		Reference	Reference	
CT	19 (32.3)	8.2	0.77	1.08 (0.60-1.96)	1.31(0.68-2.53)	0.41

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

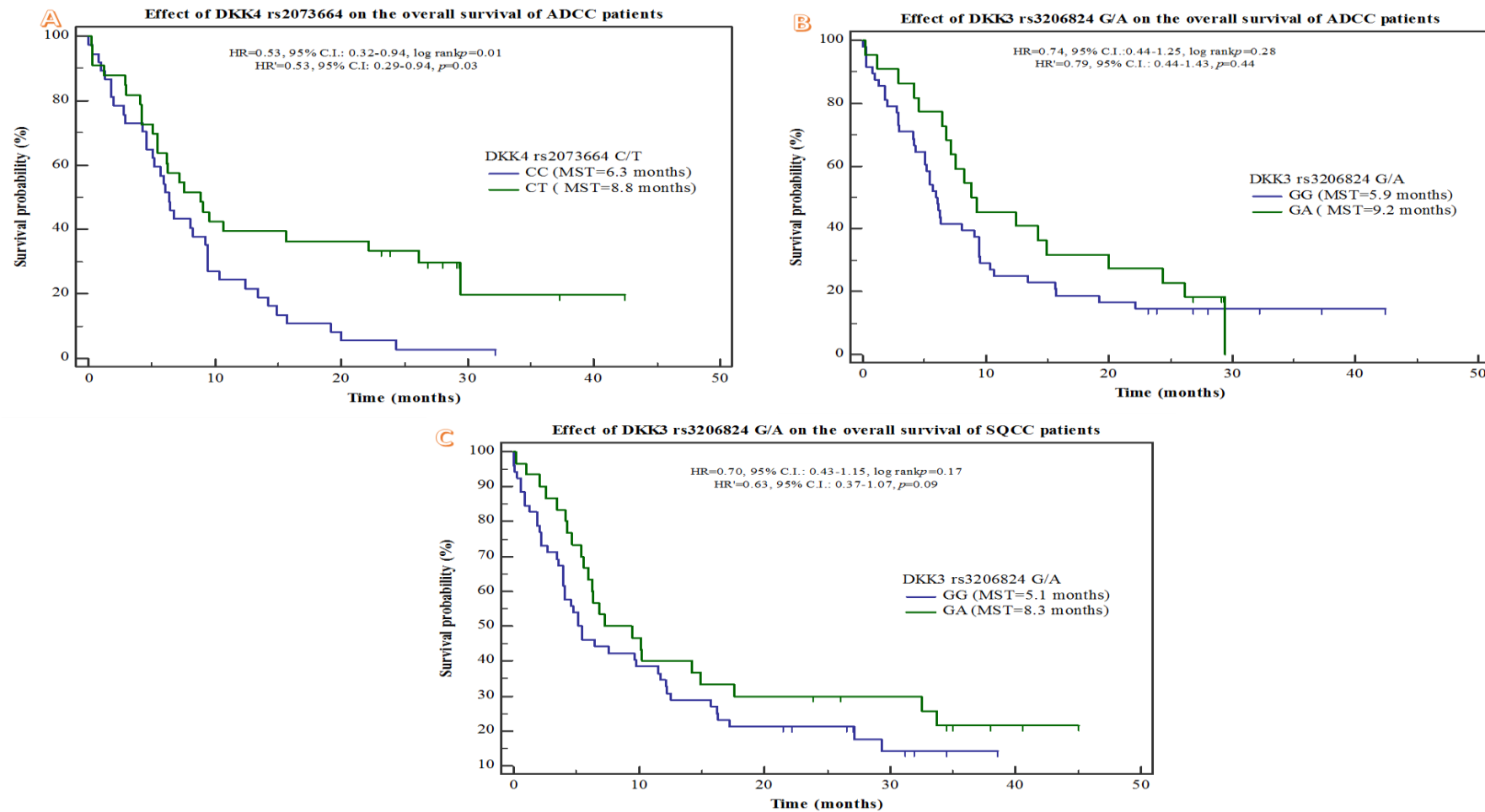


Fig 5.20: Kaplan meir survival curves for *DKK* variants in lung cancer patients stratified according to their histology; A: *DKK4* rs2063664 C/T (CC: reference) in ADCC patients, B: *DKK3* rs3206824 G/A (GG: reference) in ADCC patients, C: *DKK3* rs3206824 G/A (GG: reference) in SQCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time , ADCC: adenocarcinoma, SQCC: squamous cell carcinoma

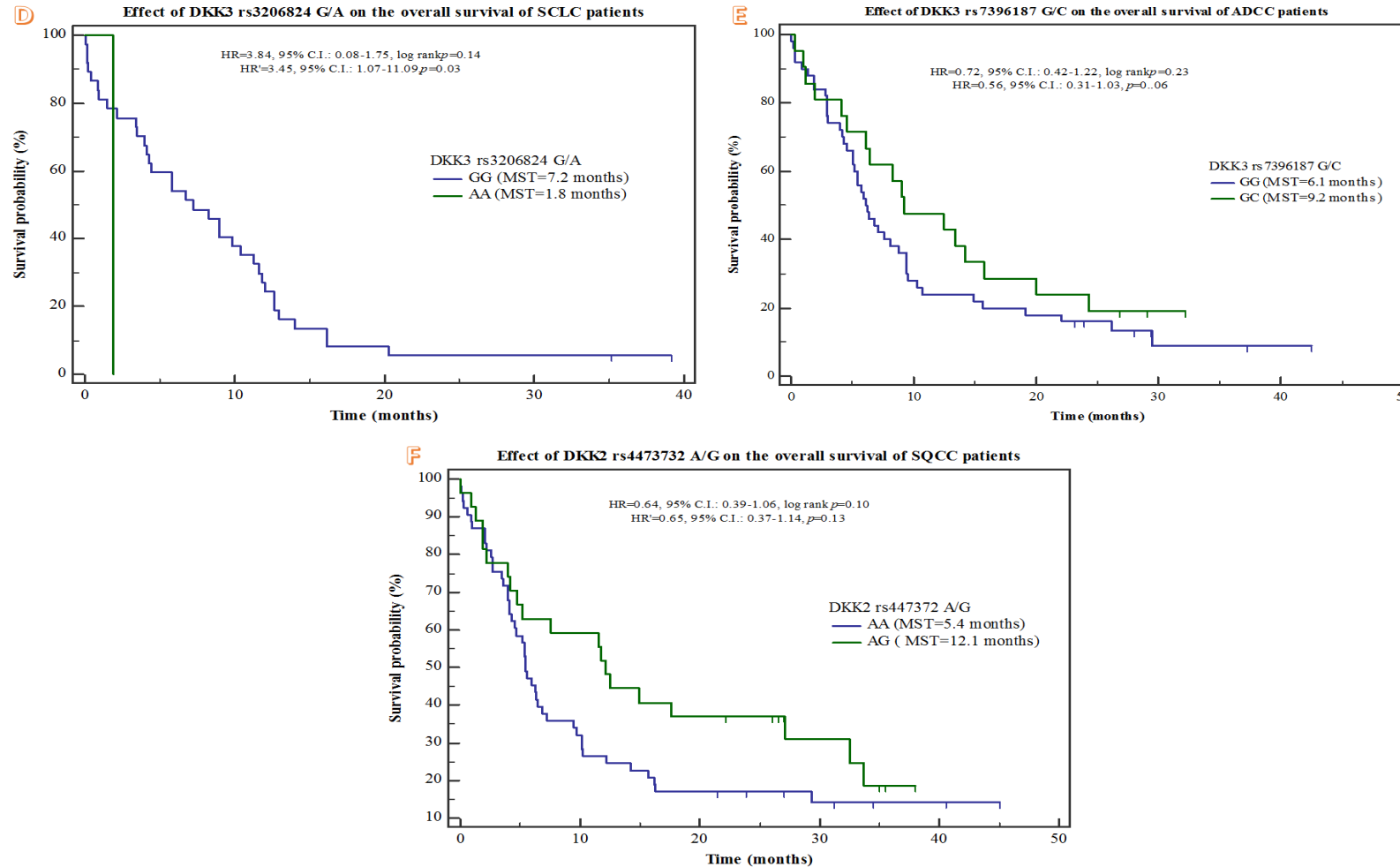


Fig 5.20: Kaplan meir survival curves for *DKK* variants in lung cancer patients stratified according to their histology; D:*DKK3* rs3206824 G/A (GG: reference) in SCLC patients, E: *DKK3* rs7396187 G/C(GG: reference) in ADCC patients, F: *DKK2* rs447372 A/G (AA: reference) in SQCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time , ADCC: adenocarcinoma, SQCC: squamous cell carcinoma, SCLC: Small cell lung carcinoma.

5.5.13 Association of Wnt antagonist (*DKK*) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response

The patients were analysed for the response to the chemotherapeutic regimen in relation to their genotypes for different *DKK* variants as shown in table 5.13. The patients were divided into good and bad responders on the basis of the RECIST criteria where they are annotated into four groups, CR (complete response), PR (partial response), SD (stable disease) and PD (progressive disease). The proportion of good (CR+PR) and bad (SD+PD) responder's was almost similar. However, the adjusted odds ratio was calculated for each polymorphic site taking the wild genotype as the reference and confounding factors taken into consideration were age, gender, smoking, stage, histology, KPS, ECOG and chemotherapeutic regimen. The objective response data was available only for 146 patients in this case. None of all the polymorphic sites of *DKK4*, *DKK3* and *DKK2* gene revealed a significant difference in the frequency of patients showing good and bad response. In case of *DKK4* rs2073664, the representation of patients having CT+TT genotype was similar among good (50%) and bad responders (48.6%). Further in case of *DKK3* rs2295199, the proportion of subjects having CT+TT genotype was slightly higher (36.4%) among good responders as compared to bad responders (30.6%) but the difference was not found to be statistically significant. Also, in case of *DKK3* rs3206824 the percentage of GA+AA was higher among good responders (36.5%) in comparison to bad responders (27.8%). The distribution of next variant rs7396187 showed that the representation of heterozygotes among good and bad responders was almost similar (24.3 and 26.4% respectively). In case of *DKK2* rs447372, the percentage of subjects carrying AG +GG genotype were 51.4% among good responders where they constituted 56.9% of bad responders. Hence, no significant difference in the frequency of variant genotype was observed among the two groups. Similarly, in case of rs17037102, the frequency of heterozygotes was slightly lower among good responders as compared to bad responders (23% vs 29.2%) but it was not showing any significant association. Only in case of *DKK2* rs419558, the heterozygous (CT) patients showing a good response (22.2%) were significantly under represented as compared to those exhibiting a poor response (44.5%). Hence, these subjects showed an adjusted odds AOR=2.36, 95% C.I.:1.05-5.28, $p=0.03$ which shows that the heterozygous (CT) patients are predicted to exhibit a poor prognosis as compared to those having wild (CC) genotype for *DKK2* rs419558 as shown in table 5.13.

Table 5.13: Relationship of different *DKK* genotypes with chemotherapy response

Genotype	Response of chemotherapy		AOR ^a (95% CI) ^b	<i>p</i> ^b
	Good Responders CR+PR (74) n (%)	Bad Responders SD+PD(72) n (%)		
<i>DKK4, rs2073664 C/T</i>				
CC	37 (50)	37 (51.3)	Reference	
CT	35 (47.2)	34 (47.2)	0.92 (0.44-1.95)	0.84
TT	2 (2.8)	1 (1.4)	1.07 (0.28-4.05)	0.92
CT+TT	37 (50)	35 (48.6)	0.90 (0.43-1.90)	0.80
<i>DKK3, rs2291599 C/T</i>				
CC	46 (62.2)	50 (69.4)	Reference	
CT	22 (29.7)	20 (27.7)	0.74 (0.33-1.69)	0.48
TT	5 (6.7)	2 (1.38)	0.48 (0.19-1.19)	0.11
CT+TT	27 (36.4)	22 (30.6)	0.64 (0.29-1.41)	0.27
<i>DKK3, rs3206824 G/A</i>				
GG	47 (63.5)	52 (72.2)	Reference	
GA	24 (32.4)	19 (26.4)	0.75 (0.34-1.66)	0.48
AA	3 (5.1)	1 (1.4)	0.53 (0.15-1.89)	0.33
GA+AA	27(36.5)	20 (27.8)	0.70 (0.33-1.50)	0.36
<i>DKK2, rs447372 A/G</i>				
AA	36 (48.6)	31 (43.1)	Reference	
AG	33 (44.6)	33 (45.8)	1.46 (0.68-3.17)	0.32
GG	5 (6.8)	8 (11.1)	1.06 (0.52-2.14)	0.86
AG+GG	38 (51.4)	41 (56.9)	1.47 (0.70-3.04)	0.30
<i>DKK3, rs7396187 G/C</i>				
GG	55 (74.3)	53 (73.6)	Reference	
GC	18 (24.3)	19 (26.4)	0.96 (0.42-2.02)	0.92
GC+CC	19 (25.7)	19 (26.4)	1.05 (0.46-2.39)	0.89
<i>DKK2, rs17037102 G/A</i>				
GG	57(77)	51 (70.8)	Reference	
GA	16 (21.6)	20 (27.7)	1.38 (0.60-3.17)	0.44
GA+AA	17(23)	21 (29.2)	1.36 (0.60-3.10)	0.45
<i>DKK2, rs419558 C/T</i>				
CC	56(77.8)	40 (55.5)	Reference	
CT	18 (22.2)	32 (44.5)	2.36 (1.05-5.28)	0.03

^a Adjusted Odds ratios, ^b 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, ECOG, KPS, stage, histology and chemotherapeutic regimen. Abbreviations: CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; AOR, Adjusted Odds ratio; CI, confidence interval.

5.6 Statistical Analysis of polymorphic sites of *Axin2* gene

5.6.1 Demographic characteristics

Table 5.14 summarizes the various demographic details of the study subjects that were recruited. The study comprised of 303 cases and 305 controls. The mean age of both the groups has been illustrated in the table 5.14 which came out to be 57.38 ± 10.74 for cases and 53.23 ± 10.44 for controls respectively. The number of males and females were also compared between the two groups and it was observed that there was no significant difference in both cohorts ($p=0.60$). The percentage of males in cases and controls was 86% vs 88% respectively. The next important factor under

consideration was smoking; an adequate attempt was made to match the number of smokers and non-smokers in the both groups. It was seen that number of smokers was slightly higher in the case group as compared to controls (78.8% vs 73.1%). However, the difference was not found to be significant ($p=0.11$). The mean pack years were also calculated for both the study groups. The mean pack years were higher in the cases as compared to controls. (28.09 ± 34.09 vs 19.4 ± 19.56). A significant difference ($p=0.0003$) was found between the pack years of cases and controls pertaining to the fact that smoking came out to be a significant risk factor for lung cancer. Further, considering the clinical parameters of the patients it was found that the major percentage of patients were diagnosed with squamous cell carcinoma (SQCC) (43.23%) followed by adenocarcinoma (ADCC) (32%) and the least percentage of patients were found to suffer from small cell lung cancer (SCLC) (23.76%). The distribution of the patients on the basis of stage is also given table 5.14. 46.7% of the patients were diagnosed with stage III and 41% with stage IV. The proportion of patients diagnosed with stage I and II were 1 and 4% respectively. The other clinical factors under study include the TNM staging where the primary tumor extension is classified from T_0 to T_4 , the presence and extent of lymph node involvement is classified as N_0 , N_1 till N_3 and distant metastasis is classified into M_0 and $M_{1a/1b}$. The percentage distribution of patients according to tumor size, lymph node invasion and distant metastasis has been recorded in table 5.14. Distribution according to primary tumor extension comprised of T_x : 3.3%, T_1 :4.7%, T_2 : 13%, T_3 : 26.3%, T_4 : 42.7%. According to lymph node involvement parameter, the patients are distributed as N_0 : 16%, N_1 : 12%, N_2 : 37.3% and N_3 : 24%. Also, the proportion of patients who did not show distant metastasis was 51.3% as compared to 38.7% who were showing positive metastasis. In all these three categories, 10 % of patients were classified as unknown.

Table 5.14: Demographic characteristics among cases and controls

Variable	Cases, n (%) N=303	Controls, n (%) N=305	<i>p</i> ^a
Age (years)			
Mean ± SD			
Range	57.38 ± 10.74	53.23 ± 10.44	p< 0.0001
Gender			
Male	262 (86.46)	269 (88.19)	0.60
Female	41(13.53)	36 (11.8)	
Smoking Status			
Smokers	239 (78.87)	223 (73.11)	0.11
Non–Smokers	64 (21.12)	82 (26.88)	
Pack Years			
Mean ± SD	28.09 ± 34.09	19.4 ± 19.56	0.0003
Histological Types			
SQCC	131 (43.23)		
ADCC	97 (32.01)		
SCLC	72 (23.76)		
Others	2 (0.66)		
Unknown	1 (0.33)		
TNM Staging			
I	3 (1)		
II	12 (4)		
III	140 (46.7)		
IV	123 (41)		
Unclassified	25 (7.3)		
Tumor Size			
Tx	10 (3.3)		
T1	14 (4.7)		
T2	39 (13)		
T3	79 (26.3)		
T4	128 (42.7)		
Unknown	33 (10)		
Lymph Node			
N0	48 (16)		
N1	36 (12)		
N2	112 (37.3)		
N3	72 (24)		
Unknown	33 (10)		
Metastasis			
M0	154 (51.3)		
M1	116 (38.7)		
Unknown	33 (10)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects ^a*p*-values were derived from Pearson Chi – square test except age; Student t-test was used for age. All *p*-values are two – sided. *p*< 0.05 was considered statistically significant . ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.6.2 Genotypic and allelic frequencies of the *Axin2* variants and their association with overall lung cancer risk

The genotypic and allelic frequencies of the seven different variants of *Axin2* gene have been tabulated in table 5.15. The seven polymorphic sites which have been evaluated in this study include *Axin2* 148 C/T (rs2240308), 1365 G/A (rs9915936), 432 T/C (rs2240308), 956+16 A/G (rs35285779), 1712+19 G/T, 1386 C/T (rs1133683) and 2062 C/T. The genotypic and allelic frequencies have been illustrated for each variant along with the minor allele frequencies in both the study groups. The table also shows the HWE values for control group of each polymorphic site. It can be observed that the following variants were in accordance to HWE namely *Axin2* 148 ($\chi^2=0.94$, df=1, $p=0.33$), 1365 ($\chi^2=1.53$ df=1, $p=0.21$), 956+16 ($\chi^2=0.31$, df=1, $p=0.57$), *Axin2* 432 ($\chi^2=0.22$, df=1, $p=0.63$) and 2062 ($\chi^2=1.84$, df=1, $p=0.17$). Only two variants namely *Axin2* 1386 ($\chi^2=8.79$, df=1, $p=0.003$) and 1712+19 ($\chi^2=175$, df=1, $p=0.000$) showed a certain deviation from HWE.

In case of *Axin2* 148 C/T, the frequency of the heterozygous (CT) genotype was consistent among the cases (49.5%) and controls (47.2%) as shown in table 5.15. The percentage of mutants was 17.8% among cases and 26.2% among controls and a significant difference was observed between the genotypic distribution of the cases and controls ($\chi^2=6.96$, df=2, $p=0.03$). MAF was found to be 0.42 in cases and 0.49 in controls. Using CC genotype as reference, the multivariate logistic regression showed that subjects carrying the variant (TT) genotype was exhibiting a protective effect towards risk for lung cancer (OR=0.55, 95% C.I: 0.34-0.87, $p=0.01$).

In case of *Axin2* 1365 G/A polymorphic site, when a comparison between the genotypic frequencies of cases and controls was done, a significant difference was observed ($\chi^2=6.83$, df=2, $p=0.03$) as shown in table 5.15. The representation of heterozygotes was less among cases (9.6%) as compared to controls (16.7%). Similarly, the subjects carrying GA+AA genotype showed a higher frequency among controls (18.3%) in comparison to cases (11.6%). The minor allele frequency was found to be 0.1 and 0.06 in controls and cases respectively. Taking GG as a reference genotype, logistic regression analysis revealed a protective effect in patients with heterozygous (GA) genotype (OR=0.45, 95% C.I.: 0.27-0.75, $p=0.002$). However, patients with mutant (AA) genotype did not show any such association towards risk for lung cancer.

The next intronic variant of *Axin2* gene evaluated was 956+16 A/G. The genotypic frequencies of cases and controls were compared and it was found that there was no significant difference between the two study groups ($\chi^2=0.89$, df=2, $p=0.63$) as shown in table 5.15. The proportions of subjects carrying the variant (GG) genotype were similar in both cases (0.6%) and controls (0.6%). Also, the frequency of AG+GG carrying subjects was almost similar in cases (15.8%) and controls (18.6%). MAF was comparable in cases (0.08) and controls (0.09). Association analysis was carried out using

AA (wild) as a reference genotype, it was visualized that neither of the variant genotypes *i.e.* AG and GG showed any significant role in modulating the risk towards lung cancer as shown in table 5.15.

For *Axin2* 432 T/C, the proportion of subjects with wild (TT) genotype were considerably over represented in both cases (97.3%) and controls (94.7%) as shown in table 5.15. The percentage of heterozygotes was 2.7% among cases and 5.2 % among controls. However, no significant difference between the genotypic frequencies among cases and controls ($\chi^2=2.07$, $df=1$, $p=0.14$). The minor allele frequency was found to be 0.02 in controls and 0.01 in cases. We also report in this study a lack of subjects having the variant (CC) or mutant genotype both in cases and controls. The individuals carrying the heterozygous (TC) genotype did not show any significant variation in their frequencies among cases and controls. Taking TT genotype as reference, the logistic regression was applied and it showed a trend of protective effect which lacked statistical significance (OR=0.56, 95% C.I.: 0.23-1.35, $p=0.20$)

Another variant lying in exon 7 of the *Axin2* gene namely 2062 C/T polymorphism was also analyzed in the present study as shown in table 5.15. The frequency of wild (CC) (84.8% vs 85.5%) and heterozygous (CT) (15.1% vs 14.4%) was roughly same in between cases and controls respectively. No significant difference was observed between the genotypic distribution of cases and controls ($\chi^2=0.02$, $df=1$, $p=0.88$). MAF of 0.07 was observed in both cases and controls. The frequency of the variant or mutant (TT) genotype was null in both the groups. When odds ratio was calculated using CC (wild) as a reference genotype, subjects carrying the heterozygous (CT) genotype did not exhibit any significant risk of developing lung cancer (OR=0.96, 95% C.I.:0.60-1.52, $p=0.86$).

5.6.3 Genotypic and allelic frequencies of the *Axin2* variants and their association with overall lung cancer risk on basis of histological subtypes

As lung cancer is one of the most heterogeneous solid tumor and the molecular pathology of each subtype is varied. Hence, the analysis for cancer predisposition was also carried out for different

subtypes of lung cancer in relation to the different genotypes of *Axin2* SNP's and the data is summarized in table 5.16.

In case of *Axin2* 148, a decreased frequency of patients carrying variant (TT) genotype was observed in SQCC subgroup (17.5% vs 26.2% in controls) which showcased a decreased risk of lung cancer in them (OR=0.51, 95% C.I: 0.28-0.94, ***p*=0.03**) as shown in table 5.16. No such association was observed in ADCC and SCLC as the frequency of subjects harboring TT genotype among ADCC and SCLC groups was 18.5% and 18%. However, the frequency was lower in both the subgroups in comparison to controls but, the difference was not found to be statistically significant. In case of *Axin2* 1365, stratifying the subjects on the basis of their histological subtypes, it was evident that lung cancer patients carrying the heterozygous and mutant genotype did not exhibit any significant association in SQCC and ADCC subgroups. The percentage of heterozygotes among SQCC and ADCC subgroups were 9.9% and 12.6% which was slightly lower than controls (16.7%). This difference was found to lack statistical significance. On the contrary, the frequency of heterozygotes decreased to 5.6% among SCLC patients and hence, it pertained to show a significant protective effect for heterozygous genotype carrying subjects (OR=0.24, 95% C.I: 0.08-0.72, ***p*=0.01**). The next intronic variant of *Axin2* gene under study was 956+16 A/G. None of the histological subgroup showed any significant inclination towards affecting the risk associated with lung cancer in individuals carrying the combined variant (AG+GG) genotype for this polymorphic site as shown in table 5.16. The frequency of combined variant genotype carrying subjects among controls (18.6%) was slightly higher as compared to SQCC (16.8%), ADCC (13.4%) and SCLC (18.1%) subgroups.

For *Axin2* 432 T/C, when patients were classified on the basis of their histological subtype, none of the subgroups revealed any association with lung cancer risk as shown in table 5.16. As the difference in frequency of heterozygotes was not found to be significant, the lowest frequency was observed among SCLC patients (1.4%), followed by SQCC (3.1) and ADCC (3.1%) and highest was estimated among the control group (5.2%). Therefore, this polymorphic site did not confer any role towards modulating lung cancer susceptibility in any of the histological subtypes.

Table 5.15: Genotypic distribution of the *Axin2* genetic variants and their association with risk of lung cancer

Genotype				
<i>Axin2 148 C/T</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	81 (26.5)	99 (32.6)	1.00 (Reference)	
CT	144 (47.3)	150 (49.6)	0.87 (0.59-1.28)	0.49
TT	80 (26.2)	54 (17.8)	0.55 (0.34-0.87)	0.01
CT+TT	224 (73.5)	204 (67.4)	0.75 (0.52-1.08)	0.13
C (Allele)	153	174		
T (Allele)	152	129		
MAF	0.49	0.42		
HWE	$\chi^2=0.94$, df=1, <i>p</i> =0.33			
<i>Axin2 1365 G/A</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
GG	249 (81.6)	268 (88.4)	1.00 (Reference)	
GA	51 (16.7)	29 (9.6)	0.45 (0.27-0.75)	0.002
AA	5 (1.6)	6 (2)	1.08 (0.37-3.73)	0.90
GA+AA	56 (18.3)	35 (11.6)	0.50 (0.31-0.81)	0.004
G(Allele)	274.5	282.5		
A(Allele)	30.5	20.5		
MAF	0.1	0.06		
HWE	$\chi^2=1.53$ df=1, <i>p</i> =0.21			
<i>Axin2 956+16 A/G</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
AA	248 (81.3)	255 (84.2)	1.00 (Reference)	
AG	55 (18.0)	46 (15.2)	0.83 (0.53-1.29)	0.41
GG	2 (0.6)	2 (0.6)		
AG+GG	57 (18.6)	48 (15.8)	0.83 (0.53-1.28)	0.40
A(Allele)	275.5	278		
G (Allele)	29.5	25		
MAF	0.09	0.08		
HWE	$\chi^2=0.31$, df=1, <i>p</i> =0.57			
<i>Axin2 432 T/C</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
TT	289 (94.8)	295 (97.3)	1.00 (Reference)	
TC	16 (5.2)	8 (2.7)	0.56 (0.23-1.35)	0.20
T(Allele)	297	299		
C (Allele)	8	4		
MAF	0.02	0.01		
HWE	$\chi^2=0.22$, df=1, <i>p</i> =0.63			
<i>Axin2 1712+19 G/T</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
TT	42 (13.8)	87 (28.7)	-----	
GT	263 (86.2)	205 (67.7)		
GG	...	11 (3.6)		
GT+TT	263 (86.2)	216 (71.3)		
T (Allele)	173.3	189.5		
G (Allele)	131.5	113.5		
MAF	0.43	0.37		
HWE	$\chi^2=175$, df=1, <i>p</i> =0.000			
<i>Axin2 2062 C/T</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	261 (85.6)	257 (84.8)	1.00 (Reference)	
CT	44 (14.4)	46 (15.2)	0.96 (0.60-1.52)	0.86
C (Allele)	283	280		
T (Allele)	22	23		
MAF	0.07	0.07		
HWE	$\chi^2=1.84$, df=1, <i>p</i> =0.17			
<i>Axin2 1386 C/T</i>	Controls (305) N (%)	Cases (303) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	103 (33.8)	94 (31)	-----	
CT	169 (55.4)	172 (56.8)		
TT	33 (10.8)	37 (12.2)		
CT+TT	202 (66.2)	209 (69)		
C (Allele)	187.5	180		
T (Allele)	117.5	123		
MAF	0.38	0.40		
HWE	$\chi^2=8.79$, df=1, <i>p</i> =0.003			

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

Another variant lying in exon 7 of the *Axin2* gene namely 2062 C/T polymorphism was also analyzed. When further stratification was carried out on basis of histological subtype, this polymorphic variant did not pose any significant modulation for any of the histological subtype. As evident by the slight differences in the frequencies of the heterozygotes among the SQCC (19%), ADCC (12.4%) and SCLC (12.5%) subgroups in comparison to the control group (14.4%) as shown in table 5.16. None of the difference in frequencies was found to be significant.

Table 5.16: Genotypic distribution of the *Axin2* genetic variants and their association with risk of lung cancer based on histology

Genotype		SQCC			ADCC			SCLC		
<i>Axin2</i> 148 C/T	Controls (305) N (%)	Cases (131) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (72) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	81 (26.5)	47 (35.9)	1.00 (Reference)		30 (30.9)	1.00 (Reference)		20 (27.8)	1.00 (Reference)	
CT	144 (47.3)	61 (46.7)	0.79 (0.48-1.29)	0.35	49 (50.5)	0.89 (0.51-1.55)	0.68	39 (54.2)	1.15 (0.61-2.16)	0.64
TT	80 (26.2)	23 (17.5)	0.51 (0.28-0.94)	0.03	18 (18.5)	0.62 (0.31-1.22)	0.17	13 (18)	0.68 (0.30-1.51)	0.34
CT+TT	224 (73.5)	84 (64.2)	0.69 (0.43-1.06)	0.11	67 (69)	0.79 (0.47-1.34)	0.39	52 (72.2)	0.98 (0.53-1.78)	0.95
		SQCC			ADCC			SCLC		
<i>Axin2</i> 1365 G/A	Controls (305) N (%)	Cases (131), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (72) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
GG	249 (81.6)	116 (88.5)	1.00 (Reference)		81 (83.5)	1.00 (Reference)		68 (94.4)	1.00 (Reference)	
GA	51 (16.7)	13 (9.9)	0.53 (0.27-1.04)	0.06	12 (12.6)	0.57 (0.27-1.19)	0.13	4 (5.6)	0.24 (0.08-0.72)	0.01
AA	5 (1.6)	2 (1.5)	0.77(0.14-4.28)	0.77	4 (4.1)	2.74 (0.66-11.57)	0.16	----	-----	
GA+AA	56 (18.3)	15 (11.4)	0.55 (0.29-1.04)	0.06	16 (16.5)	0.73 (0.38-1.40)	0.34			
		SQCC			ADCC			SCLC		
<i>Axin2</i> 956+16 A/G	Controls (305), N (%)	Cases (131), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (72) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
AA	248 (81.3)	109 (83.2)	1.00 (Reference)		84 (86.6)	1.00 (Reference)		59 (81.9)	1.00 (Reference)	
AG	55 (18.0)	21 (16)	0.92 (0.52-1.64)	0.79	13 (13.4)	0.69 (0.35-1.36)	0.29	12 (16.7)	0.97 (0.47-1.97)	0.93
GG	2 (0.6)	1 (0.8)			...			1 (1.4)		
AG+GG	57 (18.6)	22 (16.8)	0.92 (0.52-1.61)	0.78				13 (18.1)	1.0(0.50-2.00)	0.98
		SQCC			ADCC			SCLC		
<i>Axin2</i> 432 T/C	Controls (305) N (%)	Cases (131), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (72) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
TT	289 (94.8)	127 (96.9)	1.00 (Reference)		94 (96.9)	1.00 (Reference)		71 (98.6)	1.00 (Reference)	
TC	16 (5.2)	4 (3.1)	0.69 (0.22-2.19)	0.53	3 (3.1)	0.62 (0.17-2.27)	0.47	1 (1.4)	0.35 (0.04-2.76)	0.32
		SQCC			ADCC			SCLC		
		SQCC			ADCC			SCLC		
<i>Axin2</i> 2062 C/T	Controls (305), N (%)	Cases (131), N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (97) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (72) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	261 (85.6)	106 (81)	1.00 (Reference)		85 (87.6)	1.00 (Reference)		63 (87.5)	1.00 (Reference)	
CT	44 (14.4)	25 (19)	1.16 (0.66-2.03)	0.59	12 (12.4)	0.85 (0.41-1.74)	0.66	9 (12.5)	0.79 (0.36-1.77)	0.58
		SQCC			ADCC			SCLC		

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung carcinoma

5.6.4 Effects of smoking status and its association between *Axin2* gene polymorphisms and lung cancer risk

In order to evaluate the role of smoking in modulating the association between the *Axin2* gene polymorphisms and lung cancer risk, the subjects were classified on the basis of their smoking status into smokers and non-smokers. The association analysis was carried out for all the seven polymorphic sites of *Axin2* gene and the genotypic frequencies along with the adjusted odds ratio with 95% C.I within both the subgroups have been tabulated in table 5.17 (a) for smokers and table 5.17 (b) for non-smokers.

In case of *Axin2* 148 C/T, smokers harboring both the susceptible alleles were found to show a protective effect towards lung cancer as compared to the wild genotype patients (OR= 0.50, 95% C.I: 0.29-0.85 $p=0.01$) as the frequency of smokers having both susceptible allele was higher among controls (27.3%) as compared to cases (16.7%) as shown in table 5.17 (a). However, no such association was observed in case of non-smokers. In the non-smoker category, the frequency of mutant subjects was essentially comparable among cases (21.9%) and controls (23.2%). Thus, no correlation was evident between genotype and lung cancer risk among non-smokers as shown in table 5.17 (b).

A similar finding was observed in case of smokers having the heterozygous GA genotype for *Axin2* 1365 G/A (OR= 0.51, 95% C.I: 0.28-0.91, $p=0.02$) as the frequency of heterozygotes was much lower in cases (9.6%) than controls (16.6%). Also, the non-smokers having the same genotype as above exhibited a decreased risk towards lung cancer (OR= 0.32, 95% C.I: 0.10-1.10, $p=0.05$) as the frequency of mutants was lower among cases (9.4%) in comparison to controls (17.1%) as shown in table 5.17 (b). Patients in both smoker and non-smoker group, possessing the combined variant genotype (GA+AA) reported a decreased risk (OR=0.57, 95% C.I: 0.33-0.98; OR=0.33, 95% C.I: 0.11-0.99) towards lung cancer with a significant p -value of **0.04**.

On the contrary, in case of the following genetic variants namely *Axin2* 432 T/C, 956+16 A/G, and *Axin2* 2062 C/T no significant association was observed in any of the groups classified on the basis of smoking.

In case of *Axin2* 956+16 A/G, smokers having the heterozygous (AG) genotype were at a similar frequency in the case (15.1%) and control study groups (17.5%) as shown in table 5.17 (a). Also, among non-smokers percentage of heterozygotes was not found to be significantly different in between cases (15.6%) and controls (19.5%). In case of *Axin2* 432 T/C, in the smoker category the percentage of heterozygotes was not found to be significantly different among cases (2.5%) and controls (4.5%). A similar trend was seen in case of non-smokers where an increased frequency of heterozygotes (TC) was observed in the control group (7.3%) when compared to cases (3.1%).

However, the difference did lack statistical significance. In case of *Axin2* 2062 C/T, smokers carrying the heterozygous genotype were found at a similar frequency in both cases (16.3%) and controls (15.7%). A similar trend was observed among the non- smokers carrying heterozygous (CT) genotype as their proportion was almost equivalent in cases (10.9%) and controls (11%).

Table 5.17 (a): Genotypic distribution of the *Axin2* variants based on smoking status and its association with risk of lung cancer (smokers)

Genotype	SMOKERS			
<i>Axin2</i> 148 C/T	Controls n (%) N =223	Cases n (%) N=239	AOR (95% CI)^b	p^a
CC	61 (27.3)	82 (34.3)	1.00 (Reference)	
CT	101 (45.2)	117 (49)	0.85 (0.55-1.32)	0.47
TT	61 (27.3)	40 (16.7)	0.50 (0.29-0.85)	0.01
CT+TT	162 (72.6)	157 (65.6)	0.71 (0.47-1.08)/	0.11
<i>Axin2</i> 1365 G/A	Controls n (%) N =223	Cases n (%) N=239	AOR (95% CI)^b	p^a
GG	182 (81.6)	211 (88.3)	1.00 (Reference)	
GA	37 (16.6)	23 (9.6)	0.51 (0.28-0.91)	0.02
AA	4 (1.8)	5 (2.1)	1.21 (0.31-4.75)	0.77
GA+AA	41 (18.4)	28 (11.7)	0.57 (0.33-0.98)	0.04
<i>Axin2</i> 956+16 A/G	Controls n (%) N =223	Cases n (%) N=239	AOR (95% CI)^b	p^a
AA	182 (81.6)	201 (84.1)	1.00 (Reference)	
AG	39 (17.5)	36 (15.1)	0.82 (0.49-1.38)	0.46
GG	2 (0.9)	2 (0.8)		...
AG+GG	41 (18.4)	38 (15.9)	0.82 (0.50-1.36)	0.46
<i>Axin2</i> 432 T/C	Controls n (%) N =223	Cases n (%) N=239	AOR (95% CI)^b	p^a
TT	213 (95.5)	233 (97.5)	1.00 (Reference)	
TC	10 (4.5)	6 (2.5)	0.63 (0.22-1.79)	0.38
<i>Axin2</i> 2062 C/T	Controls n (%) N =223	Cases n (%) N=239	AOR (95% CI)^b	p^a
CC	188 (84.3)	200 (83.7)	1.00 (Reference)	
CT	35 (15.7)	39 (16.3)	0.97 (0.58-1.63)	0.93

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

Table 5.17 (b): Genotypic distribution of the *Axin2* variants based on smoking status and its association with risk of lung cancer (non-smokers)

Genotype	NON SMOKERS			
<i>Axin2</i> 148 C/T	Controls, n (%) N =82	Cases n (%) N=64	AOR (95% CI)^b	p^a
CC	20 (24.4)	17 (26.6)	1.00 (Reference)	
CT	43 (52.4)	33 (51.6)	0.95 (0.41-2.17)	0.91
TT	19 (23.2)	14 (21.9)	0.93 (0.37-2.57)	0.89
CT+TT	62 (75.6)	57 (89.1)	0.94 (0.43-2.08)	0.89
<i>Axin2</i> 1365 G/A	Controls (%) N =82	Cases n (%) N=64	AOR (95% CI)^b	p^a
GG	67 (81.7)	57 (89.1)	1.00 (Reference)	
GA	14 (17.1)	6 (9.4)	0.32 (0.10-1.01)	0.05
AA	1 (1.2)	1 (1.6)	0.64 (0.03-12.22)	0.77
GA+AA	15 (18.3)	7 (11)	0.33 (0.11-0.99)	0.04
<i>Axin2</i> 956+16 A/G	Controls (%) N =82	Cases n (%) N=64	AOR (95% CI)^b	p^a
AA	66 (80.5)	54 (84.4)	1.00 (Reference)	
AG	16 (19.5)	10 (15.6)	0.73 (0.29-1.82)	0.50
GG	...			
AG+GG				
<i>Axin2</i> 432 T/C	Controls (%) N =82	Cases n (%) N=64	AOR (95% CI)^b	p^a
TT	76 (92.7)	62 (96.9)	1.00 (Reference)	
TC	6 (7.3)	2 (3.1)	0.41 (0.07-2.18)	0.30
<i>Axin2</i> 2062 C/T	Controls (%) N =82	Cases n (%) N=64	AOR (95% CI)^b	p^a
CC	73(89.0)	57 (89.1)	1.00 (Reference)	
CT	9(11)	7 (10.9)	0.72 (0.23-2.20)	0.56

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

5.6.5 Combinatorial analysis of the risk associated with different polymorphic sites of *Axin2* gene

An attempt was made to analyze the SNP to SNP interaction within *Axin2* gene so as to evaluate the combined effect of these polymorphic sites on lung cancer risk. The genotypic frequencies of these different combinations and the adjusted odds ratio and 95% confidence intervals are summarized in table 5.18. Various double combinations were analyzed for finding whether there exists any significant difference between the genotypic frequencies among cases and controls. The frequency of subjects having heterozygous genotype for genotypic combination *Axin2* 148 C/T and 1365 G/A, was higher in controls (28.6%) than cases (16.2%). Hence, both these SNP's conferred a joint protective effect (OR=0.42, 95% C.I.:0.20-0.86, ***p*=0.017**) for lung cancer risk. A similar observation was also made in the case of subjects having combined variant genotype, where in a decreased risk of lung cancer was observed (OR=0.44, 95% C.I.:0.22-0.89, ***p*=0.012**). Likewise, the patients having the heterozygous (CT+CT) genotype for *Axin2* 148 C/T and *Axin2* 2062 C/T also exhibited a protective effect (OR=0.38, 95% C.I: 0.17-0.83, ***p*=0.01**), this is due to decreased frequency of heterozygotes observed in cases (12.3%) in comparison to controls (26%) as shown in table 5.18.

Table 5.18: Genotypic distribution based on different genotypic combinations of *Axin2* variants and their association with lung cancer risk

Genotype					
<i>Axin2</i> 148 C/T & <i>Axin2</i> 1365 G/A	Cases, n (%) N =105	Controls, n (%) N = 98	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(CC+GG) ⁰	87 (82.8)	70 (71.4)	1.00 (Reference)		
(CT+GA) ¹	17 (16.2)	28 (28.6)	0.42 (0.20-0.86)	0.017	0.04
(TT+AA) ²	1 (1)	...	...		
(CT+TT+GA+AA) ³	18 (17.2)	28 (28.6)	0.44 (0.22-0.89)	0.012	0.06
<i>Axin2</i> 148 C/T & <i>Axin2</i> 2062 C/T	Cases, n (%) N =97	Controls, n (%) N =104	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(CC+CC) ⁰	75 (77.3)	71 (68.3)	1.00 (Reference)		
(CT+CT) ¹	12 (12.3)	27 (26)	0.38 (0.17-0.83)	0.01	0.03
(CT+TT+CT+TT) ³	22 (10.4)	34 (32.7)	0.57 (0.30-1.09)	0.08	0.13
<i>Axin2</i> 148 C/T & <i>Axin2</i> 432 T/C	Cases, n (%) N =97	Controls, n (%) N = 85	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(CC+TT) ⁰	94 (96.9)	75 (88.2)	1.00 (Reference)		
(CT+TC) ¹	2 (2.1)	9 (10.6)	0.19 (0.03-0.93)	0.04	0.09
(CT+TT+TC+CC) ³	3 (3.0)	10 (11.8)	0.2 (0.06-0.97)	0.04	0.1
<i>Axin2</i> 148 C/T & <i>Axin2</i> 956 A/G	Cases, n (%) N = 113	Controls, n (%) N = 102	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(CC+AA) ⁰	82 (72.6)	63 (61.8)	1.00 (Reference)		
(CT+AG) ¹	22 (19.5)	23 (22.5)	0.76 (0.38-1.55)	0.46	0.63
(TT+GG) ²	1 (0.9)	...			
(CT+TT+AG+GG) ³	31 (2.6)	39 (38.2)	0.63 (0.34-1.18)	0.13	0.16
<i>Axin2</i> 148 C/T & <i>Axin2</i> 1386 C/T	Cases, n (%) N =121	Controls, n (%) N =109	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(CC+CC) ⁰	29 (24)	24 (22)	1.00 (Reference)		
(CT+CT) ¹	87 (72)	76 (69.7)	1.03 (0.52-2.01)	0.92	0.92
(TT+TT) ²	5 (4)	9 (8.3)	0.46 (0.13-1.66)	0.24	0.24
(CT+TT+CT+TT) ³	92 (76)	85 (78)	0.96 (0.50-1.86)	0.92	0.92
<i>Axin2</i> 1365 G/A & <i>Axin2</i> 2062 C/T	Cases, n (%) N =229	Controls, n (%) N = 214	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(GG+CC) ⁰	227 (99.1)	210 (98.1)	1.00 (Reference)		
(GA+CT) ¹	2 (0.9)	4 (1.9)	0.43 (0.07-2.47)	0.34	0.51
<i>Axin2</i> 1365 G/A & <i>Axin2</i> 956 A/G	Cases, n (%) N = 235	Controls, n (%) N =209	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>
(GG+AA) ⁰	228 (97)	201 (96.2)	1.00 (Reference)		
(GA+AG) ¹	7 (3)	8 (3.8)	0.72 (0.25-2.06)	0.54	0.67
<i>Axin2</i> 2062 C/T & <i>Axin2</i> 432 T/C	Cases, n (%) N =255	Controls, n (%) N = 253	Adjusted OR (95% CI) ^b	<i>p</i> ^a	<i>FDR p</i>

(CC+TT) ⁰	252 (98.8)	249 (98.4)	1.00 (Reference)		
(CT+TC) ¹	3 (1.2)	4 (1.6)	0.72 (0.15-3.38)	0.68	0.21
<i>Axin2</i> 2062 C/T & <i>Axin2</i> 956 A/G	Cases, n (%) N = 235	Controls, n (%) N = 214	Adjusted OR (95% CI)^b	p^a	FDR p
(CC+AA) ⁰	222 (94.5)	209 (97.7)	1.00 (Reference)		
(CT+AG) ¹	13 (5.5)	5 (2.3)	2.2 (0.77-6.49)	0.13	0.76
<i>Axin2</i> 432 T/C & <i>Axin2</i> 956 A/G	Cases, n (%) N = 257	Controls, n (%) N =240	Adjusted OR (95% CI)^b	p^a	FDR p
(TT+ AA) ⁰	252 (98)	236 (98.3)	1.00 (Reference)		
(TC+AG) ¹	5 (2)	4 (1.7)	1.26 (0.32-4.88)	0.72	0.76

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status, FDR: False discovery rate. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype

5.6.6 Relationship of genetic variations in *Axin2* gene with clinicopathological parameters

Clinicopathological parameters such as stage, primary tumor extension, lymph node involvement and metastasis were taken into account and analyzed to find out the effect of genetic variations in *Axin2* gene on these factors. The effect of different genotypes of *Axin2* genetic variants on the advancement of these clinical parameters and progression of the disease was assessed using logistic regression approach. The odd of advanced state of each parameter was calculated using wild genotype as the reference group. Odds ratio obtained was adjusted for age, gender and smoking status. The patients were stratified into categories based on each parameter. On the basis of clinical stage, patients were stratified into stage III and stage IV as most of the recruited cases were diagnosed with advanced lung cancer. In case of primary tumor extension, the patients were stratified into TX+T1+T2 and T3+T4 subgroup. The genotypic distribution for staging and primary tumor extension has been summarized in table 5.19 (a).

On investigating the relationship of the different polymorphic site with the clinical stage, it was evident that none of the seven *Axin2* genetic variants contribute towards predicting the stage of tumor. In case of *Axin2* 148 C/T, the prevalence of mutant (TT) genotype carrying subjects was almost similar in both stage III (15.6%) and stage IV (18.5%) patients as shown in table 5.19 (a). Similar trend was also observed for patients having heterozygous genotype for *Axin2* 1365 G/A as their proportion among stage III was 9.7% and stage IV was 9.6%. Also, the heterozygotes for *Axin2* 956+16 T/C were found at similar frequencies among stage III (14.5%) and stage IV (15.3%) category. In case of *Axin2* 432 C/T, the percentage of heterozygotes did not differ significantly among the stage III (2.1%) and stage IV subgroup (1.6%).

In case of *Axin2* 2062 C/T, the proportion of heterozygotes was slightly higher among stage III patients (17.2%) as compared to stage IV (11.3%), but the difference was not found to be statistically significant as shown in table 5.19 (a).

The genotypic frequencies have also been computed for the each polymorphic site and compared into two groups on the basis of primary tumor extension. The analysis showed that none of the polymorphic site was associated with primary tumor extension. The genotypic frequencies of the patients having TT genotype for *Axin2* 148 C/T were not statistically different among TX+T1+T2 (21%) and T3+T4 (15.7%) subgroup. Also, in case of *Axin2* 1365 G/A the frequency of heterozygous patients was 9.7% in TX+T1+T2 and 9.5% in T3+T4 subgroup. This trend of similar frequencies of heterozygous patients was evident amongst *Axin2* 956+16 T/C (13% vs 15.2%), *Axin2* 432 C/T (4.8% vs 2.4%) and *Axin2* 2062 C/T (13% vs 14.3%) SNPs in both TX+T1+T2 and T3+T4 subgroup respectively.

For lymph node invasion, patients showing lack of lymph node involvement were grouped in N0 group and others showing positive lymph node invasion were grouped under N1+N2+N3 category. Similarly, for metastatic status the patients showing positive metastasis were included in group M1 and those exhibiting lack of metastasis were grouped in M0 category and data has been tabulated in table 5.19 (b). The patients were further stratified in accordance to the genotype for each polymorphic variant in case of each parameter. In case of *Axin2* 148, the subjects having TT genotype were over represented in N1+N2+N3 category as compared to N0 group (18.7% vs 8.3%). As a result, a significant association was observed for TT genotype in case of lymph node invasion where these TT genotype patients showed a two-fold risk for lymph node invasion when compared to CC genotype patients (OR=2.07, 95% C.I.:1.11-3.86, $p=0.02$). On the contrary, in case of *Axin2* 432 C/T, a decreased frequency of heterozygous patients were observed in N1+N2+N3 (1.8%) subgroups as compared to N0 (8.3%) subgroup. Hence, a strong protective effect for lymph node invasion was observed in these patients (OR=0.19, 95% C.I.:0.04-0.83, $p=0.02$). In case of *Axin2* 1365 G/A, the frequency of heterozygous patients was 10.4% in N0 and 6.2% in N1+N2+N3 subgroups. The trend of similar frequencies of heterozygous patients was evident amongst *Axin2* 956+16 T/C (14.6% vs 14.7%) and *Axin2* 2062 C/T (12.5% vs 14.3%) SNPs in both N0 and N1+N2+N3 subgroup respectively as shown in table 5.19 (b).

Further, analyzing the genotypic distribution on the basis of distant metastasis revealed that none of the polymorphic sites of *Axin2* gene contributed in modulating the metastatic potential of lung tumor. In case of *Axin2* 148 C/T, the patients carrying both the susceptible alleles did not show any significant difference in frequency among M0 (14.6%) and M1 (20%) subgroup as shown in table 5.19 (b). Also, a similar trend was observed in subjects carrying heterozygous genotype for *Axin2* 1365, where in the percentage was 8.9% in M0 group and 10.4% among M1 subgroups. The trend of similar frequencies of heterozygous patients was evident amongst *Axin2* 956+16 T/C (15.3% vs 14%), *Axin2* 432 C/T (3.8% vs 1.7%) and *Axin2* 2062 C/T (16% vs 11.3%) SNPs in both M0 and M1 subgroup respectively.

Table 5.19 (a): Relationship of different genotypes of *Axin2* variants with the clinico-pathological parameters (stage and primary tumor extension)

Genotype	Clinical Stage		AOR (95% CI) ^a	<i>p</i> ^b	Primary Tumor Extension		AOR (95% CI) ^a	<i>p</i> ^b
	III	IV			Tx+T1+T2	T3+T4		
<i>Axin2 148 C/T</i>	145	124			62	210		
CC	48 (33.1)	38 (30.6)	Reference		16 (25.8)	71 (33.8)	Reference	
CT	73 (50.3)	63 (50.8)	0.98 (0.55-1.75)	0.95	33 (53.2)	106 (50.4)	0.72 (0.37-1.41)	0.34
TT	24 (15.6)	23 (18.5)	1.06 (0.74-1.53)	0.71	13 (21)	33 (15.7)	0.80 (0.52-1.23)	0.31
CT+TT	97 (66.9)	86 (69.4)	1.03 (0.60-1.73)	0.89	46 (74.2)	139 (66.2)	0.69 (0.36-1.32)	0.27
<i>Axin2 1365 G/A</i>	145	124			62	210		
GG	128 (88.2)	110 (88.7)	Reference		55 (88.7)	187 (89)	Reference	
GA	14 (9.7)	12 (9.6)	1.02 (0.44-2.39)	0.94	6 (9.7)	20 (9.5)	1.08 (0.40-2.89)	0.86
AA	3 (2)	2 (1.6)	0.84 (0.32-2.17)	0.72	1 (1.6)	3 (1.4)	0.93 (0.29-2.95)	0.91
GA+AA	17 (11.8)	14 (11.3)	0.96 (0.44-2.10)	0.92	7 (11.3)	23 (11)	1.05 (0.42-2.63)	0.90
<i>Axin2 956+16 A/G</i>	145	124			62	210		
AA	123 (84.8)	104 (83.9)	Reference		54 (87)	176 (83.8)	Reference	
AG	21 (14.5)	19 (15.3)	1.07 (0.53-2.14)	0.84	8 (13)	32 (15.2)	1.23 (0.53-2.86)	0.61
GG	1 (0.6)	1 (80.6)	--		---	2 (1)	--	
AG+GG	22 (15.1)	20 (16.1)	1.09 (0.55-2.15)	0.80	8 (13)	34 (16.2)	1.32 (0.57-3.04)	0.51
<i>Axin2 432 T/C</i>	145	124			62	210		
TT	142 (97.9)	122 (98.3)	Reference		59 (95.2)	205 (97.6)	Reference	
TC	3 (2.1)	2 (1.6)	0.80 (0.12-5.08)	0.81	3 (4.8)	5 (2.4)	0.50 (0.11-2.19)	0.35
<i>Axin2 2062 C/T</i>	145	124			62	210		
CC	120 (82.8)	110 (88.7)	Reference		54 (87)	180 (85.7)	Reference	
CT	25 (17.2)	14 (11.3)	0.65 (0.32-1.35)	0.25	8 (13)	30 (14.3)	1.13 (0.48-2.64)	0.76

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*^b-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

Table 5.19 (b): Relationship of different genotypes of *Axin2* variants with the clinico-pathological parameters (lymph node invasion and metastasis)

Genotype	Lymph node Invasion		AOR (95% CI) ^a	<i>p</i> ^b	Metastasis		AOR (95% CI) ^a	<i>p</i> ^b
	N0	N1+N2+N3			No	Yes		
<i>Axin2</i> 148 C/T	48	224			157	115		
CC	19 (39.6)	68 (30.3)	Reference		54 (34.4)	33 (28.7)	Reference	
CT	25 (52)	114 (50.9)	1.34 (0.68-2.66)	0.38	80 (51)	59 (51.3)	1.13 (0.64-2.01)	0.65
TT	4 (8.3)	42 (18.7)	2.07 (1.11-3.86)	0.02	23 (14.6)	23 (20)	1.24 (0.86-1.80)	0.24
CT+TT	29 (60.4)	156 (69.7)	1.59 (0.82-3.06)	0.16	103 (65.6)	82 (71.3)	1.22 (0.72-2.10)	0.44
<i>Axin2</i> 1365 G/A	48	224			157	115		
GG	42 (87.5)	141 (63)	Reference		141 (89.8)	101 (87.8)	Reference	
GA	5 (10.4)	14 (6.2)	0.81 (0.28-2.32)	0.69	14 (8.9)	12 (10.4)	1.35 (0.58-3.14)	0.47
AA	1 (2.1)	2 (0.8)	0.81 (0.25-2.57)	0.72	2 (1.3)	2 (1.7)	1.13 (0.40-3.15)	0.81
GA+AA	6 (12.5)	16 (7.1)	0.79 (0.29-2.09)	0.63	16 (10.1)	14 (12.1)	1.33 (0.60-2.93)	0.47
<i>Axin2</i> 956+16 A/G	48	224			157	115		
AA	41 (85.4)	189 (84.3)	Reference		132 (84.1)	98 (85.2)	Reference	
AG	7 (14.6)	33 (14.7)	1.00 (0.41-2.43)	0.99	24 (15.3)	16 (14)	0.91 (0.45-1.84)	0.79
GG	--	2 (0.8)	---		1 (0.6)	1 (0.8)	---	
AG+GG	7 (14.6)	35 (15.6)	1.05 (0.43-2.53)	0.90	25 (16)	17 (14.8)	0.94 (0.47-1.87)	0.87
<i>Axin2</i> 432 T/C	48	224			157	115		
TT	44 (91.7)	220 (98.2)	Reference		151 (96.1)	113 (98.3)	Reference	
TC	4 (8.3)	4 (1.8)	0.19 (0.04-0.83)	0.02	6 (3.8)	2 (1.7)	0.41 (0.07-2.20)	0.30
<i>Axin2</i> 2062 C/T	48	224			157	115		
CC	42 (87.5)	192 (85.7)	Reference		132 (84)	102 (88.7)	Reference	
CT	6 (12.5)	32 (14.3)	1.09 (0.42-2.80)	0.85	25 (16)	13 (11.3)	0.72 (0.34-1.51)	0.39

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*^b-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

5.6.7 MDR analysis within *Axin2* gene

MDR uses a computational algorithm approach to investigate all the possible interactions between the different susceptibility contributing factors which might include different SNP's, other genes and some environmental determinants. The results obtained from MDR analysis have been tabulated in table 5.20. The interpretation and selection of the best model is done on the basis of the cross-validation consistency (CVC) and average prediction error (1-testing accuracy). The best model can be determined if it has maximum CVC and minimum prediction error or maximum testing accuracy. In the present study, the best interaction model (*Axin2* 1365) was the one having CVC 9/10 and prediction error of 0.49 with a permutation *p*-value of 0.04. Out of the three models given by the software, this interaction model gave the minimum prediction error, maximum CVC and significant *p*-value. It is clear from these results that only this SNP is evidently contributing to disease susceptibility. However, in the current study the data have not been adjusted for covariates. So, further advanced analysis can reveal some interesting findings.

Table 5.20: Multifactor Dimensionality reduction (MDR) analysis of *Axin2* variants showing association of high order interactions with lung cancer.

No. of risk factors	Best interaction model	Cross Validation Consistency (CVC)	Prediction error	<i>P</i>
1	<i>Axin2</i> 1365	9/10	0.49	0.04
2	<i>Axin2</i> 1365, <i>Axin2</i> 432	7/10	0.49	0.008
3	<i>Axin2</i> 1365, <i>Axin2</i> 2062, <i>Axin2</i> 956	8/10	0.46	0.001

The entropy dendrogram obtained from this analysis is illustrated in fig 5.21. This figure suggests the degree and type of interaction taking place in the between the SNP's which are contributing in modulating the lung cancer susceptibility. As the variants, *Axin2* 1365 and *Axin2* 432 are linked with the other SNPs with a brown colored line, it shows that they have a weak interaction with the other two SNP's namely *Axin2* 956 and 2062 C/T. However, these two variants were found to be closely related as they were lined with a red colored line which is showing synergy between the two. Also, the length of the line connecting these two SNP's is shorter and hence, it is evident that they have a strong interaction.



Fig 5.21: Multifactor dimensionality reduction (MDR) interaction dendrogram. X2: *Axin2* 1365, X3: *Axin2* 432, X4: *Axin2* 2062, X5: *Axin2* 956.

5.6.8 CART analysis within *Axin2* gene:

CART analysis was used to find out the complex interactive patterns of all the polymorphic sites in the present study. This analysis did reveal all the possible high order interactions in between these variants of *Axin2* gene. For this interactive and complex algorithm, the mutant and heterozygous genotypes were combined as a single variant genotype for the polymorphic sites. The decision tree gave seven terminal nodes as shown in fig 5.22. The splitting of the tree started from the *Axin2* 1365. Terminal node 3 (*Axin2* 148 (M)/ *Axin2* 2062 (M)/ *Axin2* 956 (W)/*Axin2* 432 (W)/ *Axin2* 1365 (W) was selected as the reference as it has lowest case rate of 36.8. All the other genotypes were compared with the genotype of this terminal node so as to obtain the odds ratio associated with them which was adjusted for age, gender and smoking. The adjusted odds ratio and 95% C.I associated with all the terminal nodes have been enlisted in table 5.21. The subjects having the genotype as mentioned in terminal node 2 and 5 were found to possess maximum risk of developing lung cancer which is approximately four fold risk.

Table 5.21: Risk estimates based on CART analysis for *Axin2* variants

Terminal Node	Genotype for each terminal node	Control(n)	Case(n)	Case Rate^a	AOR (95% CI)	<i>p</i>^b
Node 1	<i>Axin2</i> 2062 (W)/ <i>Axin2</i> 956 (W)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1365 (W)	135	185	57.8	1.97 (1.20-3.22)	0.006
Node 2	<i>Axin2</i> 148 (W)/ <i>Axin2</i> 2062 (M)/ <i>Axin2</i> 956 (W)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1365 (W)	5	16	76.1	4.61(1.54-13.79)	0.006
Node 3	<i>Axin2</i> 148 (M)/ <i>Axin2</i> 2062 (M)/ <i>Axin2</i> 956 (W)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1365 (W)	24	14	36.8	1.0 ^c	
Node 4	<i>Axin2</i> 2062 (M)/ <i>Axin2</i> 956 (M)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1365 (W)	3	9	75	0.84 (0.38-1.85)	0.66
Node 5	<i>Axin2</i> 2062 (W)/ <i>Axin2</i> 956 (M)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1365 (W)	34	26	43.3	4.32 (1.08-17.15)	0.03
Node 6	<i>Axin2</i> 432 (M)/ <i>Axin2</i> 1365 (W)	13	8	38.09	1.10 (0.56-2.15)	0.77
Node 7	<i>Axin2</i> 1365 (M)	49	34	40.9	0.88 (0.33-2.37)	0.81

case rate is percentage LC cases among all individuals in each terminal node {case / (case+ control) x100};^bp<0.05 is considered statistically significant; ^c reference group

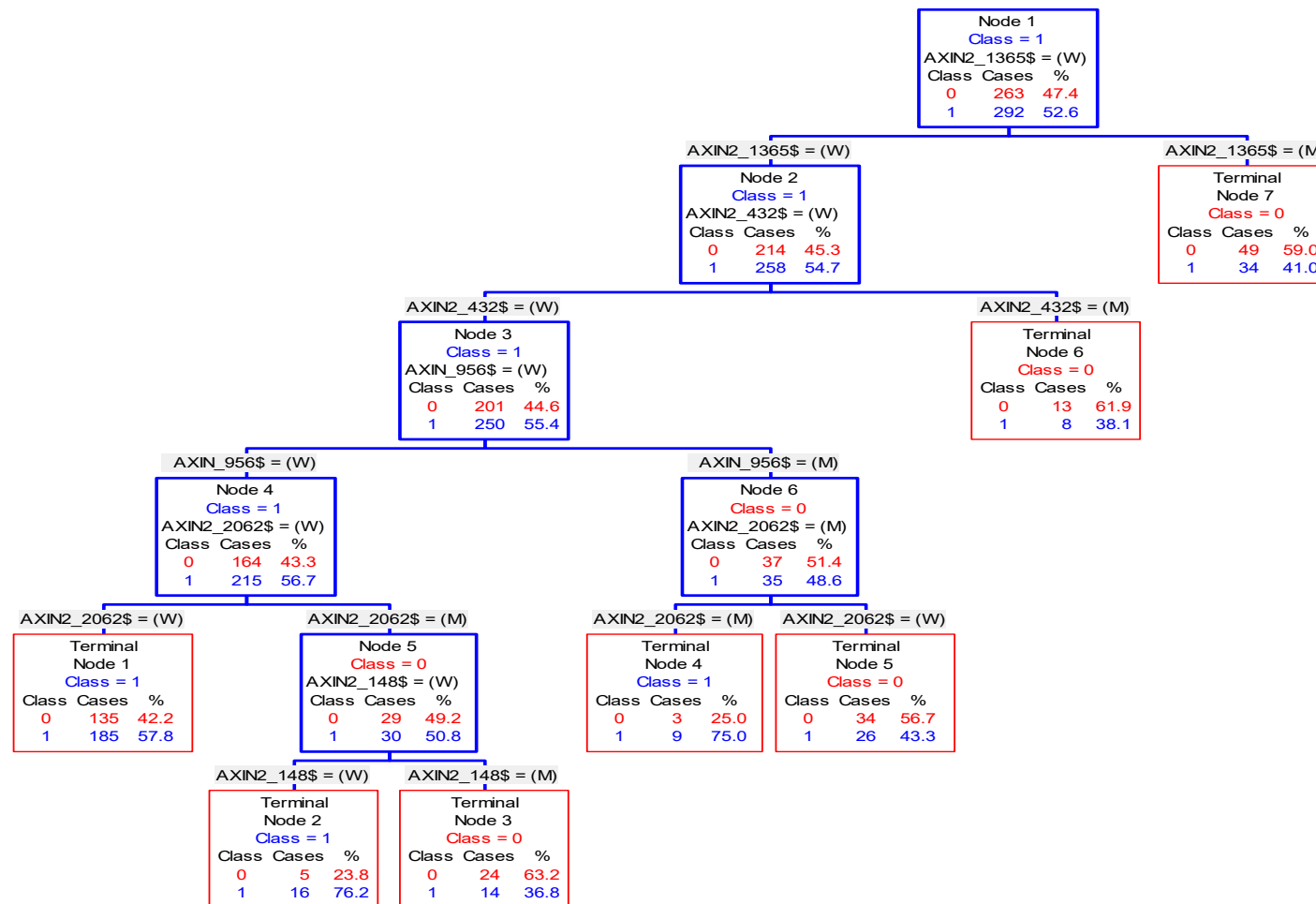


Fig 5.22: CART analysis of *Axin2* variants; W: wild type, M: variant type, Red lined color indicates terminal nodes.

5.6.9 Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of *Axin2* polymorphism

The clinical characteristics of the lung cancer patients are exactly the same as mentioned under section 5.5.10

5.6.10 Association of *Axin2* genetic variants with overall survival of lung cancer patients:

Overall survival analysis was carried for the *Axin2* polymorphic sites where a total of 212 subjects were recruited. MST was estimated for each genotypic group and compared with reference group using Kaplan Meier method and obtaining log rank *p*. Also, adjusted hazards ratio associated with the variant genotype was estimated using cox proportional hazard. The seven variants of *Axin2* gene under study, also negated the role of Wnt antagonists sequence variations in predicting the prognosis of the patients as shown in table 5.22. None of the SNPs showed any contribution in modulating the MST of the particular genotypic group.

In case of *Axin2* 148, the percentage of subjects who were either dead or alive and carry mutant genotype was 23.2% and 25.8% respectively as shown in table 5.22. The mutant (TT) genotype had a better survival time (9.4 months vs 5.5 months, log rank $p=0.39$) as compared to wild (CC) genotype which was not significant and hence, there was no association of this genotype with the survival of lung cancer patients as illustrated in fig 5.23 (A). Further, in case of *Axin2* 1365, it was observed that the percentage of dead and alive patients who were heterozygous for this SNP was 9.4%. On comparing MST of subjects having different genotypes, it was evident that wild genotype patients showed a slightly higher MST of 7.3 months as compared to 6.7 months in case of heterozygotes. However, the hazards ratio obtained from both univariate (HR=1.08, 95% C.I.: 0.54-2.13, log rank $p=0.82$) and multivariate (HR'=1.27, 95% C.I.: 0.61-2.63, $p=0.52$) analysis lacked statistical significance. The proportion of dead and alive patients heterozygous for *Axin2* 956+16 variant was 14.5% and 14.6% respectively as shown in table 5.22. The trend in this case showcased that the mutant (GG) genotype had a better MST (10 months) as compared to wild genotype (6.4 months). The difference in the MST was found to lack statistical significance and also, the hazards ratio obtained showed that this polymorphic site cannot act as an independent predictor of survival. Similarly, In case of *Axin2* 432 C/T, the percentage of patients who were either dead or alive and carried a single copy of variant allele for *Axin2* 432 was 3% and 3.2% respectively as shown in table 5.22. The subjects having the heterozygous genotype showed a better MST of 16.4 months as compared to wild type subjects (6.4 months). However, the difference in MST was not found to be statistically significant (log rank $p=0.26$) and hence the associated hazards ratio both obtained from univariate and multivariate (HR'=0.64, 95% C.I.:0.26-1.58, $p=0.34$) analysis lacked statistical significance as illustrated in fig 5.23 (B). For *Axin2* 1712+19, a decreased MST (4.3 months) was observed in case of subjects having GG genotype as compared to those having TT genotype (6.7

months). Hazards ratio in case of these subjects (HR=1.22, 95% C.I:0.84-1.77, log rank $p=0.29$) increased to HR'=1.47, 95% C.I.: 0.98-2.20, $p=0.06$ after adjusting with covariates using cox regression analysis as illustrated in fig 5.23 (B). Hence, a slight increase in death probability was observed in these patients. Further in case of *Axin2* 2062, the percentage of dead and alive lung cancer patients who were carriers was 16.6% and 13% respectively. The overall survival of heterozygous patients (6.2 months) was not found to be significantly different (log rank $p=0.69$) from that of wild type patients (7.2 months). Hence, this polymorphic site also lacked the potential of predicting the prognosis of the lung cancer patients (HR'=1.07, $p=0.71$). A similar trend was observed in case of *Axin2* 1386 polymorphic site where the percentage of dead and alive subjects carrying CT+TT genotype were 66.3% and 67.8% respectively as shown in table 5.22. Overall survival analysis exhibited that the MST of wild genotype (7.3 months) carrying subjects was highest followed by heterozygotes (6.7 months) and mutants (6.1 months). Log rank test showed that there was no significant difference among the survival times of these genotypic categories (log rank $p=0.81$). Also, the associated hazards ratio did not show any significant relationship between these genotypes and probability of death of lung cancer patients.

Table 5.22: Univariate and Multivariate analysis for Wnt antagonist (*Axin2*)

Genotype	CASES n(%) N=212	DEATH n(%, N=181	ALIVE n(%, N=31	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank <i>p</i>	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	<i>p</i>
<i>Axin2 148 C/T</i>								
CC	57 (26.9)	47 (26)	10 (32.2)	5.5		Reference	Reference	
CT	109 (51.4)	92 (50.8)	17 (54.8)	6.7	0.57	0.90 (0.571-1.42)	0.86 (0.52-1.44)	0.38
TT	46 (21.7)	42 (23.2)	8 (25.8)	9.4	0.22	0.84 (0.64-1.18)	0.76 (0.55-1.05)	0.09
CT+TT	155 (73.1)	134 (74)	21 (67.7)	7.3	0.81	0.96 (0.68-1.34)	0.91 (0.64-1.29)	0.60
<i>Axin2 1365 G/A</i>								
GG	188 (88.7)	160 (88.4)	28 (90.3)	7.3		Reference	Reference	
GA	20 (9.4)	17 (9.4)	3 (9.6)	6.7	0.82	1.08 (0.54-2.13)	1.27 (0.61-2.63)	0.52
AA	4 (1.9)	4 (2.2)	0	6.1	0.41	1.23 (0.75-2.02)	1.50 (0.89-2.50)	0.12
GA+AA	24 (11.3)	21 (11.6)	3 (9.6)	6.4	0.18	1.35 (0.81-2.26)	1.33 (0.82-2.15)	0.23
<i>Axin2 956+16 A/G</i>								
AA	179 (84.4)	153 (84.5)	26 (83.8)	6.4	0.60	Reference	Reference	
AG	31 (14.6)	26 (14.5)	5 (16.2)	7.3		0.89 (0.59-1.35)	0.90 (0.59-1.37)	0.62
GG	2 (1)	2 (1.1)	0	10			...	
AG+GG	33 (15.6)	28 (15.5)	5 (16.2)	8.1	0.49	0.87 (0.59-1.27)	0.86 (0.57-1.29)	0.47
<i>Axin2 432 T/C</i>								
TT	206 (97.2)	176 (94)	30 (96.7)	6.4		Reference	Reference	
TC	6 (2.8)	5 (3)	1 (3.2)	16.4	0.26	0.60 (0.24-1.46)	0.64 (0.26-1.58)	0.34
<i>Axin2 1712+19 G/T</i>								
TT	60 (28.3)	48 (26.5)	12 (38.7)	6.7		Reference	Reference	
GT	144 (67.9)	125 (69)	19 (61.3)	7.2	0.55	1.10 (0.79-1.54)	1.13 (0.80-1.58)	0.47
GG	8 (3.7)	8 (4.4)	0	4.3	0.29	1.22 (0.84-1.77)	1.47 (0.98-2.20)	0.06
GT+TT	152 (71.7)	133(73.5)	19 (61.3)	7.1	0.45	1.13 (0.82-1.56)	1.17 (0.84-1.64)	0.34
<i>Axin2 2062 C/T</i>								
CC	178 (84)	151(83.4)	27 (87)	7.2		Reference	Reference	
CT	34 (16)	30 (16.6)	4 (13)	6.2	0.69	1.08 (0.72-1.61)	1.07 (0.72-1.60)	0.71
<i>Axin2 1386 C/T</i>								
CC	71 (33.5)	61 (33.7)	10 (32.2)	7.3		Reference	Reference	
CT	113 (53.3)	97 (53.5)	16 (51.7)	6.7	0.88	1.02 (0.74-1.41)	1.03 (0.74-1.44)	0.81
TT	28 (13.2)	23 (12.7)	5 (16.1)	6.1	0.81	0.97 (0.76-1.23)	1.10 (0.85-1.44)	0.43
CT+TT	141 (66.5)	120 (66.3)	21(67.8)	6.4	0.98	0.99 (0.73-1.35)	0.97 (0.71-1.33)	0.89

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

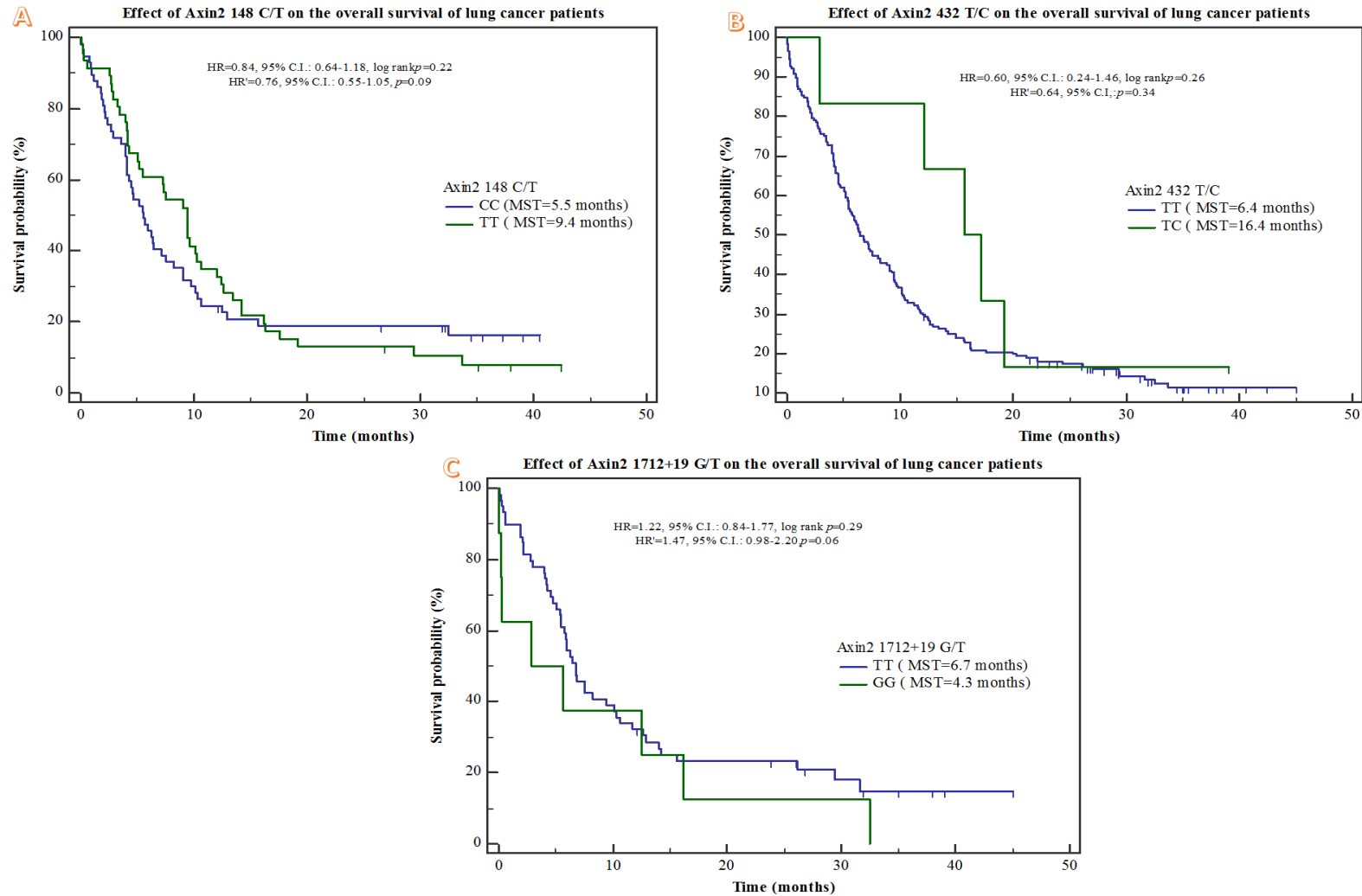


Fig 5.23: Kaplan meir survival curves for *Axin2* variants; A: *Axin2* 148 C/T (CC: reference), B: *Axin2* 432 T/C (TT: reference), C: *Axin2* 1712+19 G/T (TT: reference) for lung cancer patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time.

5.6.11 Association of Wnt antagonist (*Axin2*) polymorphism and overall survival according to histological subtype of lung cancer patients

Seven *Axin2* variants were analyzed for finding their relevance in predicting the prognosis of lung cancer patients with different histological subtype, the data for each histological subtype has been demonstrated in different tables. Results for overall survival analysis of ADCC patients is tabulated in table 5.23 (a), SQCC patients in table 5.23 (b) and SCLC patients in table 5.23(c).

For *Axin2* 148 C/T, the percentage of subjects having mutant genotype was 23.6%, 19.8% and 21.6% in case of ADCC, SQCC and SCLC subtype respectively. ADCC patients having mutant (TT) genotype showed a better survival as compared to ADCC patients having wild (CC) genotype (9.4 months vs 5.7 months) as demonstrated in fig 5.24 (A). Hazards ratio obtained from univariate analysis was found not to be statistically significant (HR=0.63, 95% C.I.: 0.30-1.29, log rank $p=0.19$) but, after applying cox regression the death ratio decreased and was found to be highly significant (HR'=0.48, 95% C.I.: 0.30-0.78, $p=0.003$) as shown in table 5.23 (a). However, in case of SQCC patients the MST of mutant genotype was the highest (9.6 months) in comparison to wild type (6.2 months) and heterozygous (5.4 months) but, it did not show any significant association with the mortality rate as hazards ratio in case of both univariate and multivariate were found to be statistically insignificant as shown in table 5.23 (b). Further, in case of SCLC patients the MST of mutants (4.1 months) and wild type (4.3 months) patients was nearly similar and no significant association was reported among this genetic variant and overall survival of SCLC and SQCC patients. In case of *Axin2* 1365, the proportion of heterozygotes in ADCC, SQCC and SCLC subtype was 11.1%, 10.5% and 5.9% respectively. ADCC patients having mutant (AA) genotype were very few in number but showed a much lower MST of 3.3 months in comparison to wild type (8 months). Log rank p -value obtained was near to significance (0.07) and adjusted hazards ratio obtained by cox regression showed a two-fold increased rate of mortality (HR'=2.07, 95% C.I.: 0.97-4.43, $p=0.06$) as shown in fig 5.24 (B). SQCC patients with mutant genotype were few in number and had the highest MST of 10.8 months in comparison to wild type (7.2 months) followed by heterozygous patients (4.7 months) as illustrated in fig 5.24 (C). However, difference in MST was not found to be statistically significant as shown in table 5.23 (b). Also, the SCLC patients having heterozygous (GA) genotype showed a shorter MST of 3.3 months as compared to wild (AA) type (7.3 months). This difference in MST pertains to an increased mortality rate (HR=2.18, 95% C.I.: 0.41-11.55, log rank $p=0.17$), but lacked statistical significance.

The next variant *Axin2* 956+16 was reported to lack the ability of acting as an independent predictor of survival in any of the three histological subtypes. The proportion of heterozygotes among ADCC, SQCC and SCLC subtype was 13.9%, 14% and 17.6%. ADCC patients having single susceptible allele for this genetic variant showed a similar MST as compared to wild type (6.7 months vs 6.3

months) shown in table 5.23 (a). However, SQCC patients having heterozygous (AG) genotype did show a better OS of 9.7 months as compared to wild (AA) type patients (6.2 months) but this difference lacked statistical significance. On the contrary, an opposite trend was observed in case of SCLC patients having heterozygous genotype as they depicted a shorter MST of 4.1 months in comparison to wild type (8 months). But, the mortality rate associated with this genotype also lacked statistical significance. In case of *Axin2 432*, the proportion of patients having heterozygous genotype was found under represented in all the three histological subtypes namely ADCC (2.8%), SQCC (3.5%) and it was nil in SCLC category. However, the MST of heterozygous (TC) subjects in ADCC subgroup had a better MST of 11 months in comparison to 6.7 months as in the case of wild (TT) type patients. A similar trend was observed in case of SQCC patients who were carrying a single copy of the variant allele for *Axin2 432*, as their MST was 15.7 months and the wild type were found to have a median survival time of 6.2 months as shown in fig 5.24 (D). But, among both the histological subtypes a statistical significance was absent and hence, this genetic variant cannot be assessed to predict the survival of lung cancer patients. Furthermore, for *Axin2 1712+19*, the frequency of heterozygotes in ADCC, SQCC and SCLC subgroup was 61.1%, 73.3% and 66.6% respectively. It was found that ADCC patients possessing mutant (GG) genotype, had a very poor MST as compared subjects with wild (TT) genotype (0.27 months vs 8.3 months, $p<0.0001$) as illustrated in fig 5.24 (E). The adjusted hazards ratio depicted four-fold higher death probability (HR=4.14, 95% C.I.: 1.19-8.97, $p=0.0003$) in these patients when compared to those with TT genotype. SQCC patients having the heterozygous genotype (GT) were found to have a MST of 6.4 months which is similar to that of wild type (5.4 months) and hence, no association was observed in this case. In case of SCLC subgroup, the subjects having heterozygous genotype had a better OS (MST=8 months) as compared to wild type (6.2 months). But, there was no evident significant association as the log rank p -value was 0.40 shown in table 5.23 (c).

The genetic variant *Axin2 2062* showed lack of association with the overall survival of lung cancer patients in any of the histological subtype. MST of ADCC patients having heterozygous (CT) genotype was 3.4 months which was shorter in comparison to wild (CC) type subjects (7.5 months) shown in table 5.23 (a). SQCC patients carrying heterozygous (CT) genotype showed the survival time of 6.2 months and wild (CC) type had MST of 6.8 months. Heterozygotes characterized with SCLC subtype of lung cancer depicted a MST of 7.3 months and wild type showed the median survival time of 8 months. The log rank test reported insignificant p -values in case of each histological subtype as shown in table 5.23. In case of *Axin2 1386*, the ADCC patients having CT+TT genotype showed the MST of 6.4 months and wild type showed the MST of 7.8 months. Also, the SQCC patients carrying CT+TT genotype had a survival time of 6.8 months and wild type had the MST of 5.5 months. But in both the cases, none of the genotypes showed a significant difference in MST among the two histological subgroups. On the contrary, SCLC patients having the heterozygous

(CT) genotype for *Axin2* 1386, showed a two-fold higher death risk (HR^a=2.33, 95% C.I.: 1.11-4.90, *p*=0.02) with reference to wild type SCLC subjects shown in table 5.23 (c).

Table 5.23 (a): Association of *Axin2* SNPs on overall survival according to tumor histology (ADCC)

Genotype	ADCC					
	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
<i>Axin2</i> 148 C/T		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>P</i>
CC	18 (25)	5.7		Reference	Reference	
CT	37 (51.3)	6.3	0.21	0.68 (0.35-1.31)	0.76 (0.37-1.58)	0.47
TT	17 (23.6)	9.4	0.19	0.63 (0.30-1.29)	0.48 (0.30-0.78)	0.003
CT+TT	54 (75)	8.03	0.21	0.70 (0.37-1.31)	0.73 (0.37-1.43)	0.36
<i>Axin2</i> 1365 G/A	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>P</i>
GG	62 (86.2)	8.0		Reference	Reference	
GA	8 (11.1)	5.7	0.35	1.44 (0.58-3.61)	1.29 (0.50-3.33)	0.59
AA	2 (29.2)	3.3	0.07	3.33 (0.27-41.09)	2.07 (0.97-4.43)	0.06
GA+AA	10 (13.8)	5.47	0.15	1.64 (0.69-3.86)	1.55 (0.65-3.69)	0.32
<i>Axin2</i> 956+16 A/G	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
AA	62 (86.1)	6.3		Reference	Reference	
AG	10 (13.9)	6.7	0.15	0.57 (0.30-1.10)	0.55 (0.23-1.29)	0.17
<i>Axin2</i> 432 T/C	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
TT	70 (97.2)	6.7		Reference	Reference	
TC	2 (2.8)	11.0	0.86	1.12 (0.25-5)	1.06 (0.25-4.49)	0.93
<i>Axin2</i> 1712+19 G/T	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
TT	23 (32)	8.3		Reference	Reference	
GT	44 (61.1)	7.5	0.76	1.08 (0.63-1.87)	1.16 (0.64-2.10)	0.62
GG	5 (6.9)	0.27	<0.0001	6.46 (0.82-50.83)	4.14 (1.19-8.97)	0.0003
GT+GG	49 (68)	6.1	0.49	1.20 (0.71-2.03)	1.28 (0.72-2.28)	0.39
<i>Axin2</i> 2062 C/T	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	64 (88.9)	7.5		Reference	Reference	
CT	8 (11.1)	3.4	0.19	1.67 (0.63-4.42)	1.53 (0.64-3.65)	0.33
<i>Axin2</i> 1386 C/T	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	26 (36.1)	7.8		Reference	Reference	
CT	38 (52.7)	6.4	0.42	1.24 (0.73-2.12)	1.30 (0.69-2.45)	0.41
TT	8 (11.1)	7.7	0.60	0.79 (0.33-1.85)	0.92 (0.54-1.55)	0.76
CT+TT	46 (63.9)	6.4	0.73	1.09 (0.64-1.84)	1.10 (0.61-2)	0.73

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.23 (b): Association of Axin2 SNPs on overall survival according to tumor histology (SQCC)

Genotype	SQCC					
	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
Axin2 148 C/T						
CC	27 (31.4)	6.2		Reference	Reference	
CT	42 (48.8)	5.4	0.76	1.08 (0.62-1.87)	1.22 (0.68-2.17)	0.49
TT	17 (19.8)	9.6	0.78	1.09 (0.56-2.10)	1.19 (0.82-1.71)	0.35
CT+TT	59 (68.6)	6.8	0.65	1.10 (0.67-1.85)	1.26 (0.73-2.16)	0.39
Axin2 1365 G/A						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
GG	75 (87.2)	7.2		Reference	Reference	
GA	9 (10.5)	4.7	0.66	1.18 (0.51-2.74)	1.45 (0.64-3.30)	0.36
AA	2 (2.3)	10.8	0.87	1.12 (0.25-4.97)	1.29 (0.58-2.84)	0.52
GA+AA	11 (12.8)	5.3	0.62	1.1 (0.56-2.51)	1.51 (0.71-3.21)	0.28
Axin2 956+16 A/G						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
AA	73 (84.9)	6.2		Reference	Reference	
AG	12 (14)	9.7	0.73	0.88 (0.46-1.69)	0.87 (0.43-1.73)	0.70
GG	1 (1.1)	10.1	0.96	1.04 (0.13-7.86)	1.13 (0.40-3.18)	0.81
AG+GG	13 (15.1)	10.1	0.66	0.86 (0.46-1.60)	0.86 (0.45-1.67)	0.67
Axin2 432 T/C						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
TT	83 (96.5)	6.2		Reference	Reference	
TC	3 (3.5)	15.7	0.74	0.72 (0.28-2.38)	0.73 (0.19-2.72)	0.64
Axin2 1712+19 G/T						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
TT	21 (24.4)	5.4		Reference	Reference	
GT	63 (73.3)	6.4	0.81	1.07 (0.60-1.88)	1.14 (0.62-2.09)	0.65
GG	2 (2.3)	22.5	0.71	0.76 (0.20-2.90)	0.81 (0.33-1.95)	0.64
GT+GG	65 (75.6)	7.1	0.83	1.06 (0.60-1.86)	1.12 (0.61-2.04)	0.70
Axin2 2062 C/T						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	H.R ^b	<i>p</i>
CC	68 (79)	6.8		Reference	Reference	
CT	18 (21)	6.2	0.94	0.98 (0.55-1.73)	0.92 (0.49-1.72)	0.79
Axin2 1386 C/T						
	Cases n (%) N=86					
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	30 (34.9)	5.5		Reference	Reference	
CT	45 (52.3)	7.5	0.51	0.84 (0.49-1.42)	0.85 (0.49-1.48)	0.58
TT	11 (12.8)	5.4	0.78	1.10 (0.52-2.34)	1.20 (0.78-1.83)	0.39
CT+TT	56 (65.1)	6.8	0.77	0.93 (0.56-1.52)	0.89 (0.53-1.51)	0.68

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.23 (c): Association of Axin2 SNPs on overall survival according to tumor histology (SCLC)

Genotype	SCLC					
<i>Axin2 148 C/T</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	11 (21.6)	4.3		Reference	Reference	
CT	29 (56.9)	8.2	0.96	0.98(0.46-2.08)	0.72(0.32-1.65)	0.45
TT	11(21.6)	4.1	0.77	1.13(0.46-2.79)	0.69(0.37-1.26)	0.23
CT+TT	40(78.4)	8.0	0.89	1.05 (0.51-2.14)	0.72(0.33-1.58)	0.42
<i>Axin2 1365 G/A</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
GG	48(94.1)	7.3		Reference	Reference	
GA	3 (5.9)	3.3	0.17	2.18(0.41-11.55)	1.23(0.34-4.38)	0.74
AA	..	..				
GA+AA						
<i>Axin2 956+16 A/G</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
AA	41 (80.4)	8.0		Reference	Reference	
AG	9 (17.6)	4.1	0.09	1.18(0.74-4.45)	1.77(0.82-3.80)	0.14
GG	1 (2)	9.8	0.84	1.21(0.13-10.66)	0.64(0.22-1.85)	0.41
AG+GG	10 (19.6)	5.7	0.15	1.62 (0.71-3.68)	1.49(0.71-3.14)	0.28
<i>Axin2 432 T/C</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
TT	50 (98)	7.2		Reference	Reference	
TC	1 (2)	..	0.06	...	...	
<i>Axin2 1712+19 G/T</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
TT	16 (31.4)	6.2		Reference	Reference	
GT	34 (66.6)	8.0	0.40	1.29(0.71-2.35)	1.06(0.47-2.47)	0.87
GG	1(2)	16.1	0.51	0.52(0.11-2.49)	0.90(0.23-3.48)	0.88
GT+GG	35(68.6)	8.03	0.31	1.35(0.75-2.45)	1.26(0.55-2.91)	0.57
<i>Axin2 2062 C/T</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	44(86.2)	8.0		Reference	Reference	
CT	7(13.7)	7.3	0.87	0.93(0.42-2.05)	0.83(0.32-2.11)	0.70
<i>Axin2 1386 C/T</i>	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	13(25.5)	9		Reference	Reference	
CT	29(56.9)	6.7	0.85	1.06(0.55-2.03)	0.88(0.41-1.88)	0.75
TT	9(17.6)	8.0	0.71	0.84(0.34-2.07)	2.33(1.11-4.90)	0.02
CT+TT	38(74.5)	6.7	0.65	0.86(0.44-1.67)	0.82(0.41-1.66)	0.60

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG;

FDR (False discovery rate) adjusted *p*-value

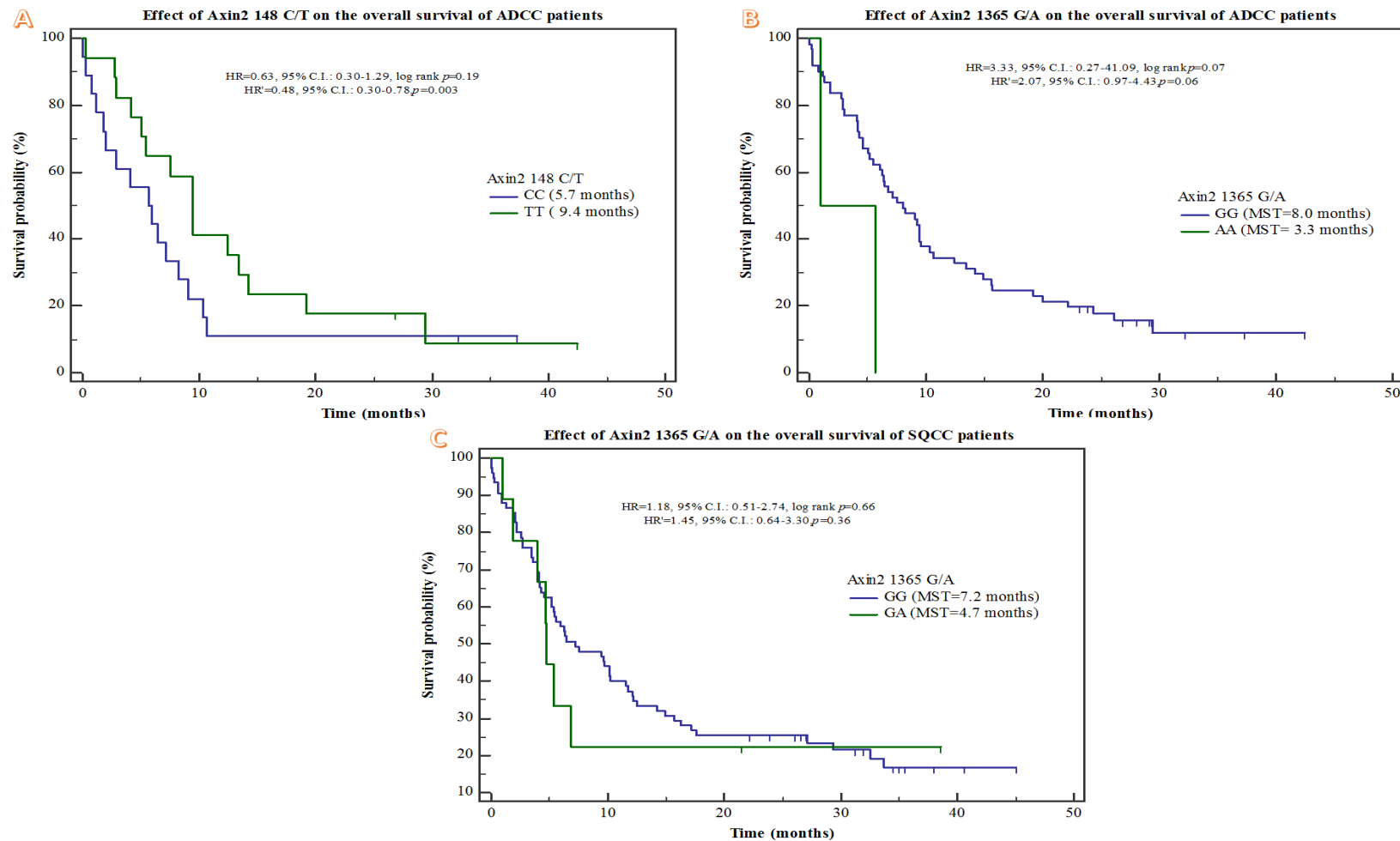


Fig 5.24: Kaplan meir survival curves for *Axin2* variants in lung cancer patients stratified according to their histology; A: *Axin2* 148 C/T (CC: reference) in ADCC patients, B: *Axin2* 1365 G/A (GG: reference) in ADCC patients, C: *Axin2* 1365 G/A (GG: reference) in SQCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time, ADCC: adenocarcinoma, SQCC: squamous cell carcinoma.

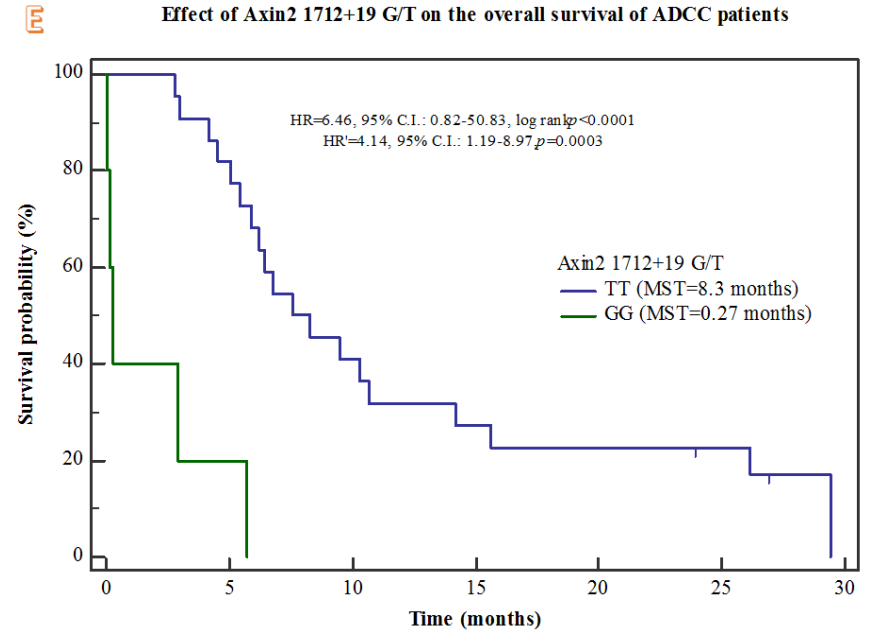
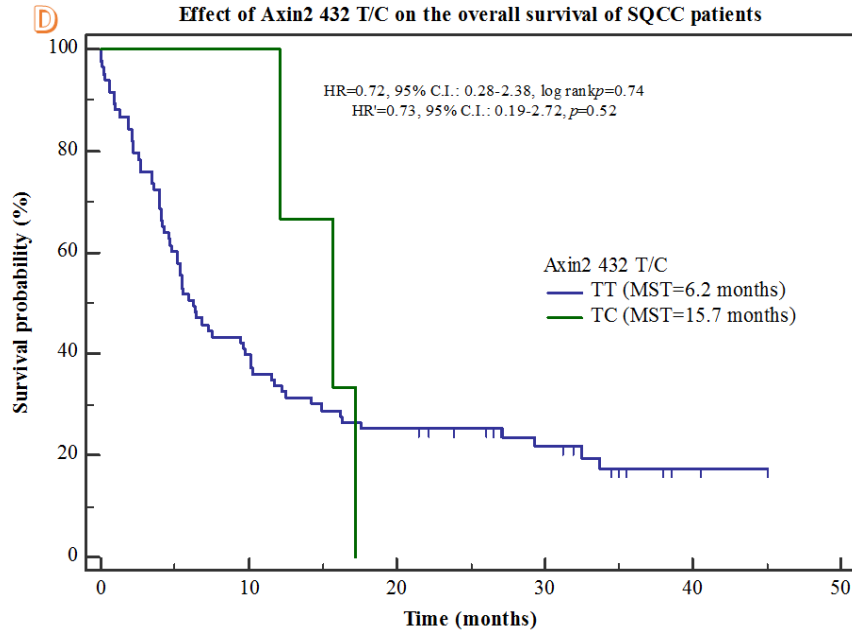


Fig 5.24: Kaplan meir survival curves for *Axin2* variants in lung cancer patients stratified according to their histology; D: *Axin2* 432 T/C (TT: reference) in SQCC patients, E: *Axin2* 1712+19 G/T (TT: reference) in ADCC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time , ADCC: adenocarcinoma, SQCC: squamous cell carcinoma.

5.6.12 Association of Wnt antagonist (*Axin2*) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response

The objective response of the patients towards chemotherapeutic treatment was assessed in relation to the genotype of the *Axin2* polymorphic sites which has been summarized in table 5.24. On the basis of objective response using the RECIST criteria, the patients were stratified as good responders which included patients showing complete response (CR) and partial response (PR) and bad responders comprising patients showing progressive disease (PD) and stable disease (SD). Overall, the numbers of good and bad responders were found to be in same proportion. However, we assessed the number of good and bad responders in each genotypic category for each polymorphic site. The association of genotype with the objective response was estimated using logistic regression approach where an adjusted odds ratio was obtained after adjusting for various confounders such as age, gender, smoking, ECOG, KPS, stage, histology and chemotherapeutic regimen. In case of *Axin2 148*, the percentage of subjects having combined variant genotype was similar in case of both good responders (68%) and bad responders (68.5%). When stratification was carried out for *Axin2 1365*, it was observed that bad responders included 19.1% of patients having combined variant genotype (GA+AA) as compared to 4% in good responder category and therefore these patients exhibited almost a 5.6 fold statistically significant poor response as compared to patients having wild genotype (AOR=5.65, 95% C.I.: 1.45-22.05, $p=0.01$) as shown in table 5.24. Further, in case of four polymorphic sites, namely *Axin2 956+16*, *Axin2 432*, *Axin2 1712+19* and *Axin2 2062* no significant association was observed between the genotype and objective response of the patients towards chemotherapy. The frequency of variant genotype carrying subjects in case of *Axin2 956+16*, *Axin2 432*, *Axin2 1712+19* and *Axin2 2062* among good and bad responders was 16% vs 13.7%, 5.3% vs 1.4%, 68% vs 56.2% and 14.7% vs 9.5% respectively. On the other hand as shown in table 5.24, in case of *Axin2 1386* there was difference in the percentage of wild (CC) (41.3% vs 19.2%), heterozygous (CT) (44% vs 65.7%) and CT+TT (58.7% vs 80.8%) genotypes among good and bad responders respectively. The patients having heterozygous (CT) genotype exhibited 3.5 fold poor response as compared to wild CC genotype patients (AOR=3.53, 95% C.I.:1.51-8.27, $p=0.003$). Similarly, the patients having combined variant (CT+TT) genotype also showed a three-fold poor response as compared to those having CC genotype (AOR=3.05, 95% C.I.: 1.37-6.76, $p=0.006$).

Table 5.24: Relationship of different *Axin2* genotypes with the chemotherapy response

Genotype	Response of chemotherapy		AOR ^a (95% CI) ^b	<i>p</i>
	Good Responders CR+PR (75) n (%)	Bad Responders SD+PD(73) n (%)		
<i>Axin2 148 C/T</i>				
CC	24 (32)	23 (31.5)	Reference	
CT	36 (48)	39 (53.4)	1.14 (0.51- 2.55)	0.74
TT	15 (20)	11 (15.1)	0.76 (0.43-1.34)	0.35
CT+TT	51 (68)	50 (68.5)	0.99 (0.46-2.09)	0.98
<i>Axin2 1365 G/A</i>				
GG	72 (96)	59 (80.9)	Reference	
GA	3 (4)	9 (12.3)	3.52 (0.82-15.12)	0.08
AA	...	5 (6.8)	...	
GA+AA	3 (4)	14 (19.1)	5.65 (1.45-22.05)	0.01
<i>Axin2 956+16 A/G</i>				
AA	61 (81.3)	63 (86.3)	Reference	
AG	12 (16)	10 (13.7)	1.06 (0.39-2.86)	0.89
GG	2 (2.7)	---	...	
GA+AA	14 (18.7)	10 (13.7)	0.93 (0.35-2.42)	0.88
<i>Axin2 432 T/C</i>				
TT	71 (94.7)	72 (98.6)	Reference	
TC	4 (5.3)	1 (1.4)	0.13 (0.01-1.43)	0.09
<i>Axin2 1712+19 G/T</i>				
TT	22 (29.3)	30 (41.1)	Reference	
GT	51 (68)	41 (56.2)	0.63 (0.29-1.37)	0.25
GG	2 (2.7)	2 (2.7)	1.07 (0.33-3.42)	0.90
GG+GT	53 (70.7)	43 (58.9)	0.65 (0.30-1.40)	0.27
<i>Axin2 2062 C/T</i>				
CC	64 (85.3)	66 (90.4)	Reference	
CT	11 (14.7)	7 (9.5)	0.70 (0.24-2.03)	0.51
<i>Axin2 1386 C/T</i>				
CC	31 (41.3)	14 (19.2)	Reference	
CT	33 (44)	48 (65.7)	3.53 (1.51-8.27)	0.003
TT	11 (14.7)	11 (15.1)	1.59 (0.90-2.81)	0.10
CT+TT	44 (58.7)	59 (80.8)	3.05 (1.37-6.76)	0.006

^a Adjusted Odds ratios, ^b 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, ECOG, KPS, stage, histology and chemotherapeutic regimen Abbreviations: CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; AOR, Adjusted Odds ratio; CI, confidence interval.

5.7 Statistical analysis of polymorphic sites of *sFRP* gene

5.7.1 Demographic characteristics

Demographic characteristics of the study population including age, gender, smoking status, pack years and clinico-pathological factors have been summarized in table 5.25. The mean age of cases and controls was found to be 57.3 ± 10.86 and 53.11 ± 10.42 respectively. However, the proportion of males and females were almost comparable in cases and controls as the percentage of males and females in both the study groups was found to be approximately same and no significant difference was observed ($p=0.956$) as shown in table 5.25. The numbers of smokers were found to be over-represented in cases (78.3%) as compared to controls (71.4%) ($p=0.06$). The mean pack years were also calculated for both cases (28.9) and controls (19.4), and there was significant difference between the pack years of both the study groups ($p=0.0003$). Hence, smoking came out to be significant factor in assessing the risk for lung cancer. Furthermore, the clinico-pathological factors including histology, staging and TNM status of the lung cancer subjects has also been documented. Out of 300 subjects, 42.7% of the cases were confirmed with SQCC, 32% were found to have ADCC and 23.7% were classified as SCLC along with 1.3% as others and 0.3% as unknown. Further, the classification was done on the basis of stage, where the percentage of subjects falling in different categories was as follows: Stage I: 1%, Stage 2: 4%, Stage 3: 46% and Stage 4: 41.7%. The proportion of subjects possessing the clinical parameters namely, primary tumour extension (Tx (3.3%), T1 (4.7%), T2 (13%), T3 (26%) and T4 (42.3%)), lymph node involvement (N0 (16.3%), N1 (12%), N2 (37%) and N3 (24%)) and metastasis (M0 (50.6%), M1 (38.7%)) along with 10.7% in unknown category are also listed in table 5.25.

5.7.2 Genotypic and allelic distribution of the four genetic variants and their association with risk of lung cancer

The genotypic frequency distribution for all the four genetic variants are given in table 5.26. The control group of all the four polymorphic sites showed some deviations from HWE and has been mentioned in the table 5.26. The possible reasons for these deviations are attributable to the fact that HWE which is based on various assumptions such as random mating, genetic drift etc. which might be violated in this sample set taken from the entire population. The reasons include population stratification, sample size and non-random mating. In case of *sFRP3* rs288326 C/T, the proportion of subjects with the wild type genotype (CC) were 84.8% and 87.8 % in both cases and controls respectively. Also, no major difference was found in the frequency of subjects with CT genotype in both the study groups ($\chi^2=1.68$, $df=2$, $p=0.42$). The minor allele (T) frequency was 0.081 and 0.10 in controls and cases respectively. With CC genotype as a reference, the association analysis for estimating the susceptibility towards lung cancer was carried out using multivariate logistic regression

after adjusting for age, gender and smoking parameter. No significant odds for developing lung cancer in patients carrying the mutant TT genotype for *sFRP3* rs288326 was observed (OR=1.47, 95% CI: 0.70-3.10, $p=0.30$). Genotypic Frequencies for *sFRP3* rs7775 C/G in cases and controls respectively is as follows CC, CG, and GG: 77.3 vs 78.8%, 18 vs 15% and 4.7 vs 5.5% ($\chi^2=0.69$, $df=2$, $p=0.70$).

Table 5.25: Demographic characteristics among cases and controls

Variables	Cases, n (%) N=300	Controls, n (%) N=311	<i>p</i> ^a
Age			
Mean ± SD	57.23 ± 10.86	53.11 ± 10.42	<0.0001
Range			
Gender			
Male	258 (86)	268 (86.2)	0.956
Female	42 (14)	43 (13.8)	
Smoking Status			
Smokers	235 (78.3)	222 (71.4)	0.059
Non-smokers	65 (21.7)	89 (28.6)	
Pack Years			
Mean ± SD	28.09 ± 34.09	19.4 ± 19.56	0.0003
Histological Types			
SQCC	128 (42.7)		
ADCC	96 (32)		
SCLC	71 (23.7)		
Others	4 (1.3)		
Unknown	1 (0.3)		
TNM Staging			
I	3 (1)		
II	12 (4)		
III	138 (46)		
IV	125 (41.7)		
Unclassified	22 (7.3)		
Tumor Size			
Tx	10 (3.3)		
T1	14 (4.7)		
T2	39 (13)		
T3	78 (26)		
T4	127 (42.3)		
Unknown	32 (10.7)		
Lymph Node			
N0	49 (16.3)		
N1	36 (12)		
N2	111 (37)		
N3	72 (24)		
Unknown	32 (10.7)		
Metastasis			
M0	152 (50.6)		
M1	116 (38.7)		
Unknown	32 (10.7)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects ^a p -values were derived from Pearson Chi – square test except age; Student t-test was used for age. All p -values are two – sided. $p < 0.05$ was considered statistically significant

MAF for both cases and controls was found to be 0.13 and 0.13 however, the association analysis showed that subjects with mutant (GG) genotype or combined variant (CG+GG) genotype were found not to have any significant predisposition towards lung cancer as shown in table 5.26. For *sFRP4* rs1802073 C/A, the distribution of the genotypic frequencies for both cases and controls was found to be significantly different as the proportion of subjects with mutant AA genotype were found to be over represented in cases as compared to controls (29.7% vs 14.5%) $\chi^2=23.10$, df=2, $p<0.0001$. The minor allele frequency (A) was found to be 0.45 and 0.57 in controls and cases respectively. In case of *sFRP4* rs1802074 G/A, the genotypic frequencies for genotype GG, GA and AA genotype were as follows in cases and controls (12 vs 18.3%, 82.2 vs 73.3%, 5.7 vs 8.4%) ($\chi^2=7.19$, d.f=2, $p=0.02$).The minor allele frequency was found to be 0.46 and 0.45 in cases and controls respectively.

Table 5.26: Genotypic distribution of the *sFRP* genetic variants and their association with risk of lung cancer

Genotype				
<i>sFRP3</i> rs288326	Controls (311)	Cases (300)	AOR (95% CI)^b	<i>p</i>^a
C/T	N (%)	N (%)		
CC	273 (87.8)	254 (84.7)	1.00 (Reference)	
CT	25 (8)	27 (9)	1.33 (0.73- 2.41)	0.34
TT	13 (4.2)	19 (6.3)	1.47 (0.70 - 3.10)	0.30
CT+TT	38 (12.2)	46 (15.3)	1.36 (0.84- 2.19)	0.19
C	285	267		
T	25	32		
MAF	0.081	0.10		
HWE	$\chi^2=67.54$ df=2, $p=0.00$			
<i>sFRP3</i> rs7775 C/G	Controls (311)	Cases (300)	AOR (95% CI)^b	<i>p</i>^a
C/G	N (%)	N (%)		
CC	245 (78.8)	232 (77.3)	----	
CG	49 (15.7)	54 (18)		
GG	17 (5.5)	14 (4.7)		
GG+CG	66 (21.2)	68 (22.7)		
C	269	259		
G	41	41		
MAF	0.13	0.13		
HWE	$\chi^2=31.59$ df=2, $p=0.000$			
<i>sFRP4</i> rs1802073	Controls (311)	Cases (300)	AOR (95% CI)^b	<i>p</i>^a
C/A	N (%)	N (%)		
CC	75 (24.1)	46 (15.3)	----	
CA	191 (61.4)	165 (55)		
AA	45 (14.5)	89 (29.7)		
CA+AA	236 (75.9)	254 (84.7)		
C	170	128		
A	140	171		
MAF	0.45	0.57		
HWE	$\chi^2=17.88$, df=2, $p=0.00$			
<i>sFRP4</i> rs1802074	Controls (311)	Cases (300)	AOR (95% CI)^b	<i>p</i>^a
G/A	N (%)	N (%)		
GG	57 (18.3)	36 (12)	-----	
GA	228 (73.3)	247 (82.3)		
AA	26 (8.4)	17 (5.7)		
GA+AA	254 (81.7)	264(88)		
G	171	159		
A	140	140		
MAF	0.46	0.45		
HWE	$\chi^2= 71.93$, df=2, $p=0.00$			

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

5.7.3 Genotypic and allelic distribution of the four genetic variants and the stratified association analysis based on histological subtypes

Further, we also evaluated the risk for developing lung cancer after stratification on basis of histological subtypes as shown in table 5.27. For *sFRP3* rs288326 C/T, the stratified analysis revealed that ADCC patients with heterozygous (CT) genotype showed a two-fold higher risk towards lung cancer with OR=2.32, 95% C.I: 1.14-4.73, ***p*=0.02** which was found to be statistically significant as the frequency of heterozygotes was 15.6% which was higher as compared to control group (8%). Similarly, for the combined CT+TT genotype a significant two-fold risk was observed (OR= 2.28 95% C.I: 1.26-4.14 ***p*=0.006**). However, the subjects having SQCC and SCLC did not show any such association. The frequency distribution in case of SQCC (7% vs 8%) and SCLC (2.8% vs 8%) showed that the frequency of heterozygotes (CT) was not statistically different among these histological group and control group respectively.

Table 5.27: Genotypic distribution of the *sFRP3* genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histology

Genotype		SQCC			ADCC			SCLC		
<i>sFRP3</i> rs288326 C/T	Controls (311) N (%)	Cases (128) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (96) N (%)	AOR (95% CI) ^b	<i>p</i> ^a	Cases (71) N (%)	AOR (95% CI) ^b	<i>p</i> ^a
CC	273 (87.8)	113 (88.3)	1.00 (Reference)		73 (76.1)	1.00 (Reference)		64 (90.1)	1.00 (Reference)	
CT	25 (8)	9 (7)	1.03 (0.45- 2.38)	0.92	15 (15.6)	2.32(1.14- 4.73)	0.0 2	2 (2.8)	0.45 (0.10- 2.01)	0.2 9
TT	13 (4.2)	6 (4.7)	0.83 (0.30- 2.30)	0.73	8 (8.3)	2.34(0.91- 5.96)	0.0 7	5 (7.1)	1.09 (0.36- 3.31)	0.8 6
CT+TT	38 (12.2)	15 (11.7)	0.94 (0.48- 1.82)	0.86	23 (23.9)	2.28(1.26- 4.14)	0.0 06	7 (9.8)	0.76 (0.31- 1.85)	0.5 5

^aTwo- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. Abbreviations: ADCC, Adenocarcinoma; SQCC, Squamous cell carcinoma; SCLC, Small cell lung ca

5.7.4 Genotypic distribution of the four genetic variants based on smoking status and its association with risk of developing lung cancer

The study subjects were classified as smokers and non-smokers on basis of their smoking habits. Table 5.28 (a) and 5.28 (b) shows the genotypic distributions for both smokers and non-smokers for all the four *sFRP* SNP's respectively. In case of *sFRP3* rs288326 C/T, no significant association was seen in smokers with combined variant (CT+TT) genotype of this polymorphic site as the frequency of subjects were similar in cases (12.8%) and controls (12.2%). However, an over representation of subjects with CT+TT genotype was seen among non-smoker cases (24.6%) as compared to controls (12.4%) and it was observed that the non-smoker subjects with the combined variant CT+TT genotype showed an approximately three-fold increased risk (OR= 2.82, 95% C.I: 1.16-6.83, $p=0.02$) of developing lung cancer.

Table 5.28 (a): Genotypic distribution of the *sFRP* variants based on smoking status and its association with risk of lung cancer (smokers)

Genotype	SMOKERS			
<i>sFRP3</i> rs288326 C/T	Controls n (%) N =222	Cases n (%) N=235	AOR (95% CI) ^b	p^a
CC	195 (87.9)	205 (87.2)	1.00 (Reference)	
CT	14 (6.3)	19 (8.1)	1.28 (0.61-2.66)	0.50
TT	13 (5.8)	11 (4.7)	0.67 (0.28-1.56)	0.35
CT+TT	27 (12.2)	30 (12.8)	0.97 (0.55-1.72)	0.93

.^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding p -values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

Table 5.28 (b): Genotypic distribution of the *sFRP* variants based on smoking status and its association with risk of lung cancer (non-smokers)

Genotype	NON-SMOKERS			
<i>sFRP3</i> rs288326 C/T	Controls n (%) N =89	Cases n (%) N=65	AOR (95% CI) ^b	p^a
CC	78 (87.6)	49 (75.4)	1.00 (Reference)	
CT	11 (12.4)	8 (12.3)	1.30 (0.47-3.60)	0.61
TT	--	8 (12.3)	--	
CT+TT	11 (12.4)	16 (24.6)	2.82 (1.16- 6.83)	0.02

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls .^bAdjusted Odds ratios, 95% confidence intervals and their corresponding p -values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status

5.7.5 Clinico-pathological parameters and their association with genotype

Various clinico-pathological parameters such as clinical staging, primary tumor extension, metastasis etc. were analysed to find out their correlation with the genotypes of various *sFRP* variants. This association analysis basically gives an insight about the role of the studied SNPs in tumor progression. The subjects were categorized into groups of two on the basis of their clinical stage, primary tumor extension and metastasis as summarized in table 5.29.

On the basis of clinical stage, the patients were stratified into stage III and stage IV. The genotypic distribution for each polymorphic site has been estimated. In case of *sFRP3* rs288326, the heterozygous patients were found at slightly higher frequency among stage III (10.1%) as compared to stage IV (5.6%) group. However, the difference was not statistically significant. The next parameter evaluated for these four polymorphic sites was primary tumor extension, the patients were stratified into two groups namely, T3+T4 and T1+T2. None of the four polymorphic sites conferred any effect on the tumor extension. In case of *sFRP3* rs288326 (8.8% vs 9.5%) and rs7775 (18% vs 19%), the percentage of heterozygotes was similar among T3+T4 and T1+T2 subgroup respectively. Further, in case of distant metastasis, the patients were divided into M0 and M1 subgroup. In case of *sFRP3* rs288326, no significant difference was observed in the frequency of heterozygous subjects among M0 (9.9%) and M1 (7.8%) subgroup.

Table 5.29: Relationship of clinico-pathological features and *sFRP3* & *sFRP4* genotypes in lung cancer patients

Genotype	Clinical stage		AOR (95% CI)	<i>p</i>	TNM stage		AOR (95% CI) ^b	<i>p</i>	Metastasis		AOR (95% CI)	<i>p</i>
	III	IV			T3+T4	T2+T1			Positive	Negative		
<i>sFRP3</i> rs288326	138	125			205	63			116	152		
CC	118 (85.5)	116 (92.8)	Reference		176 (85.9)	54 (85.7)	Reference		99 (85.3)	131 (86.1)	Reference	
CT	14 (10.1)	7 (5.6)	0.52 (0.19-1.4)	0.19	18 (8.8)	6 (9.5)	0.91(0.34-2.4)	0.86	9(7.8)	15 (9.9)	0.70 (0.28-1.74)	0.45
CT+TT	20 (14.5)	9 (7.2)	0.80 (0.34-1.62)	0.55	29 (14.1)	9 (14.3)	0.99 (0.44-2.25)	0.99	17(14.7)	21 (13.8)	0.82(0.45-1.91)	0.84

^b Adjusted Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking and histological sub-types.

5.7.6 Clinical characteristics of lung cancer patients evaluated for detecting the prognostic relevance of *sFRP* polymorphism

It is same as mentioned under section 5.5.10

5.7.7 Association of *sFRP* genetic variants with overall survival of lung cancer patients

Overall survival analysis was carried out by using kaplan meier and cox regression method for the two genetic variants of *sFRP3* gene (rs288326 and rs7775) and *sFRP4* gene (rs1802073 and rs1802074). A total of 212 subjects were included in the study and out of them 181 were reported as dead and 31 were alive. As shown in table 5.30 in case of *sFRP3* gene out of the two SNPs studied none of them showed significant contribution in predicting the survival of the patients. For the *sFRP3* rs7775, the subjects having the mutant (GG) genotype had a better survival time (9.4 months) as compared to wild (6.4 months) genotype as shown in fig 5.25. But, both univariate and multivariate analysis revealed insignificant hazards ratio. Similarly, in case of rs288326 subjects carrying mutant (TT) genotype showed the least MST (5.4 months) followed by wild type (6.7 months) and heterozygous (6.8 months). There was no significant difference in MST of any of the genotype (HR=0.75, 95% C.I.: 0.53-1.08, $p=0.13$). Similarly, both the polymorphic variations of *sFRP4* gene (rs1802073 and rs1802074) reported lack of association in relation to OS of lung cancer patients. Although, in case of rs1802073 the mutant (AA) genotype did revealed a lower MST (5.8 months) as compared to wild type (6.1 months) but the difference was not found to be significant and adjusted hazards ratio obtained in that case was also insignificant as shown in table 5.30. Lastly, in case of rs1802074, the heterozygous (GA) genotype carrying patients showed the maximum MST of 6.8 months, followed by the wild type (6.4 months) and mutants showed the least MST (5.7 months). However, none of the categories were significantly different with each other in terms of MST and log rank test evaluated the same. Also, both the univariate and multivariate analysis estimated no significant association between these genotypes and mortality rate of lung cancer patients.

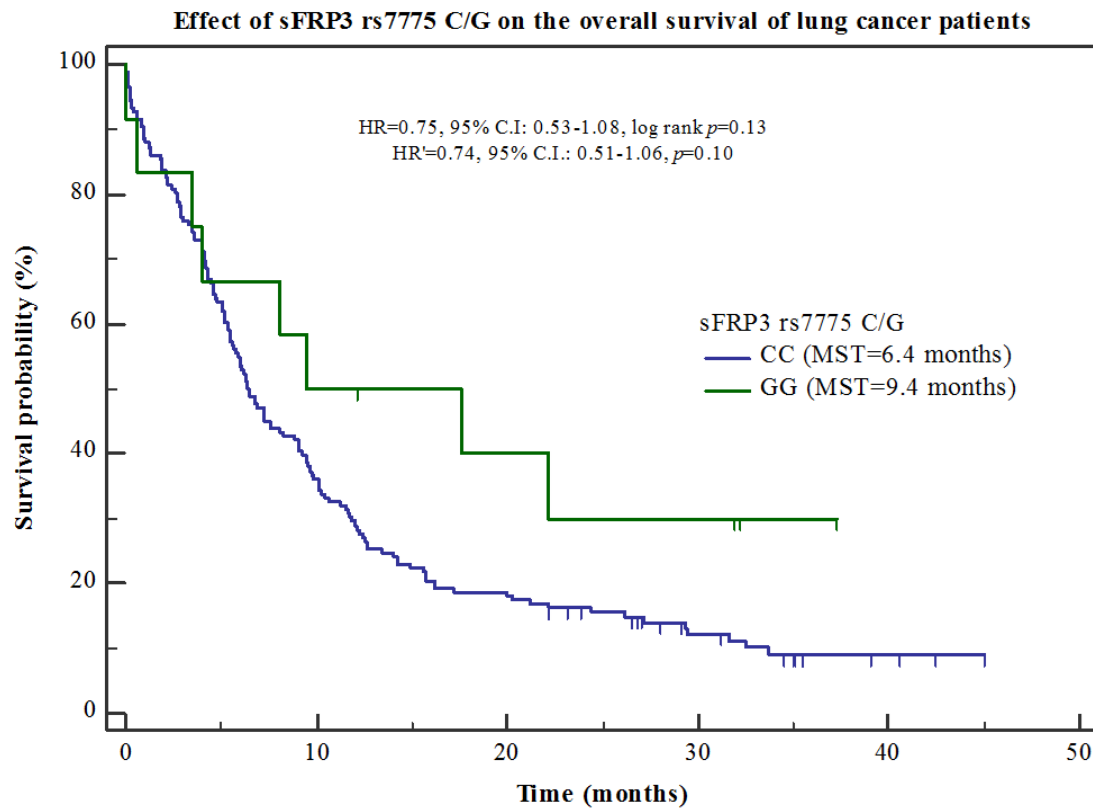


Fig 5.25: Kaplan meir survival curves for *sFRP3* rs7775 C/G (CC: reference) in lung cancer patients.
 HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time.

Table 5.30: Univariate and Multivariate analysis for Wnt antagonist (*sFRP*)

Genotype	CASES n(%) N=212	DEATH n(%) N=181	ALIVE n(%) N=31	Univariate analysis			Multivariate analysis	
				MST (months)	Log rank <i>p</i>	Unadjusted HR ^a	Adjusted HR ^b (95% CI)	<i>p</i>
<i>sFRP3</i> rs7775 C/G								
CC	167 (78.7)	147(81.2)	20(64.5)	6.4		Reference	Reference	
CG	33 (15.5)	26(14.3)	7(22.5)	7.3	0.41	0.83 (0.55-1.27)	0.84 (0.55-1.28)	0.44
GG	12 (5.6)	8 (4.4)	4(12.9)	9.4	0.13	0.75 (0.53-1.08)	0.74 (0.51-1.06)	0.10
CG+GG	45 (21.2)	34 (18.8)	11(35.5)	8.3	0.12	0.73 (0.53-1.05)	0.73 (0.50-1.07)	0.11
<i>sFRP3</i> rs288326 C/T								
CC	179 (84.4)	153 (84.5)	26(83.8)	6.7		Reference	Reference	
CT	20 (9.4)	19 (10.4)	1(3.2)	6.8	0.48	1.18 (0.71-1.97)	1.30 (0.80-2.12)	0.28
TT	13 (6.1)	9 (5)	4(12.9)	5.4	0.68	0.93 (0.66-1.30)	0.98 (0.69-1.39)	0.94
CT+TT	33 (15.6)	28 (15.5)	5(16.1)	6.1	0.74	1.06 (0.70-1.61)	1.21 (0.79-1.83)	0.37
<i>sFRP4</i> rs1802073 C/A								
CC	34 (16)	28 (15.5)	6(19.3)	6.1		Reference	Reference	
CA	118 (55.6)	99 (46.7)	19(61.3)	7.5	0.54	0.87 (0.57-1.33)	0.89 (0.57-1.38)	0.22
AA	60 (28.3)	54 (29.8)	6(19.3)	5.8	0.58	1.06 (0.84-1.33)	1.20 (0.92-1.57)	0.16
CA+AA	178 (84)	153 (84.5)	25(80.6)	6.7	0.73	0.93 (0.61-1.41)	1.04 (0.68-1.59)	0.83
<i>sFRP4</i> rs1802074 G/A								
GG	28 (13.2)	24 (13.2)	4(12.9)	6.4		Reference	Reference	
GA	174(82)	149(82.3)	25(80.6)	6.8	0.72	0.92 (0.60-1.41)	0.93 (0.58-1.48)	0.77
AA	10(4.7)	8(4.4)	2(6.5)	5.7	0.35	0.82 (0.55-1.23)	0.74 (0.45-1.21)	0.74
GA+AA	184(86.7)	157(86.7)	27(87.1)	6.8	0.57	0.88 (0.56-1.38)	0.85 (0.53-1.36)	0.51

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, histology, stage, KPS, ECOG; FDR (False discovery rate) adjusted *p*-value

5.7.8 Association of Wnt antagonist (*sFRP*) polymorphism and overall survival according to histological subtype of lung cancer patients

The four variants of this Wnt antagonist were evaluated for their potential to act as the predictor of survival of lung cancer patients stratified on the basis of histological subtype. The result of the analysis has been tabulated in different tables for different histological subtype. Table 5.31 (a) contains the overall survival data for ADCC patients, 5.31 (b) for SQCC patients and 5.31 (c) for SCLC patients.

In case of *sFRP3* rs7775, the percentage of heterozygous subjects among ADCC, SQCC and SCLC subtype was found to be 20.8, 12.8 and 9.8% respectively. ADCC patients having heterozygous (CG) genotype had slightly better MST of 7.3 months in comparison to 6.3 months in case of subjects having wild (CC) genotype as shown in table 5.31 (c). Similarly, the heterozygous subjects of SQCC subtype also showed a better survival (16.2 months) as compared to wild type subjects (6.2 months) respectively as shown in fig 5.26 (A). Out of the two variants of *sFRP3* gene, rs7775 exhibited a penchant in predicting the survival of SCLC patients only. SCLC patients having heterozygous (CG) genotype showed a poor survival time when compared with wild (CC) genotype (1.5 months vs 7.2 months, log rank $p=0.01$) as shown in fig 5.26 (B). Death rate was found to be two times higher in case of these subjects (HR'=2.03, 95% C.I.: 0.97-4.25, $p=0.05$).

Further, in case of *sFRP3* rs288326 the proportion of heterozygotes found across the ADCC, SQCC and SCLC subtype was 12.6, 7 and 3.9% respectively. ADCC subjects having heterozygous genotype had slightly poor MST of 6.2 months in comparison to 7.1 months in case of wild (CC) type subjects. SQCC patients having the mutant genotype showed the least MST of 5.4 months followed by wild type (6.4 months) and maximum in case of heterozygous (CT) subjects (7.2 months) as shown in table 5.31 (b). SCLC patients having heterozygous (CT) genotype had better MST of 14.6 months in comparison to 7.3 months in case of wild type subjects. However, there was no significant difference in MST among any of the histological subtypes and hence, this polymorphic variant could not act as independent predictor of survival for any of them.

On the other hand, *sFRP4* genetic variants were found to play a negligible role in prognosis of patients having any of the histological subcategory. In case of rs1802073, 27.7, 26.7 and 31.4% of the subjects in ADCC, SQCC and SCLC subtype respectively were found to be mutant (AA) for this polymorphic site. ADCC patients having mutant genotype showed a lower MST of 4.1 months as compared to wild type (7.5 months) and heterozygous subjects showed a MST of 9.4 months as shown in fig 5.26 (C), however it was not found to be statistically significant. On the contrary, in case of SQCC subjects, mutant (AA) genotype showed a better MST of 6.4 months in comparison to wild type (5.4 months). However, the trend was opposite in case of SCLC patients where in the mutant

genotype carrying patients had a lower MST of 5.8 months as compared to wild type (8.2 months) as shown in table 5.31 (b).

In case of rs1802074 G/A, the percentage of heterozygous subjects among ADCC, SQCC and SCLC subtype was found to be 83.2, 80.2 and 84.3 respectively. ADCC subjects having heterozygous genotype showed a better survival (7.1 months) as compared to wild type (5.1 months) as shown in table 5.31 (a). However, in the case of SQCC patients the MST of both wild and heterozygous was found to be 6.4 months followed by 8.5 months in case of mutants. Lastly, in SCLC subgroup the MST of wild, heterozygous and mutant genotype was estimated as 8, 7.2 and 12.9 months respectively as shown in fig 5.26 (D). However, no significant association was found on applying either univariate or multivariate analysis.

Table 5.31 (a): Association of *sFRP* SNPs on overall survival according to tumor histology (ADCC)

Genotype	ADCC					
<i>sFRP3</i> rs7775 C/G	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	Hr ^a	H.R ^b	<i>p</i>
CC	52 (72.2)	6.3		reference	Reference	
CG	15 (20.8)	7.1	0.85	1.06 (0.56-1.98)	0.76 (0.39-1.47)	0.42
GG	4 (5.5)	-	0.12	0.35 (0.14-0.89)	0.43 (0.19-0.99)	0.05
CG+GG	19 (26.3)	8.3	0.39	0.77(0.44-1.34)	0.65 (0.33-1.28)	0.21
<i>sFRP3</i> rs288326 C/T	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	Hr ^a	H.R ^b	<i>p</i>
CC	55 (76.3)	7.1		Reference	Reference	
CT	12 (16.6)	6.2	0.87	1.05 (0.52-2.11)	1.30 (0.62-2.71)	0.47
TT	5 (7)	6.1	0.42	0.62 (0.24-1.62)	1 (0.47-2.10)	0.99
CT+TT	17 (23.6)	6.1	0.80	0.92 (0.50-1.68)	1.25 (0.63-2.49)	0.52
<i>sFRP4</i> rs1802073 C/A	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	H.R ^b	<i>p</i>
CC	9 (12.5)	7.5		Reference	Reference	
CA	43 (59.7)	9.4	0.31	0.67 (0.27-1.62)	0.72 (0.28-1.84)	0.50
AA	20 (27.7)	4.1	0.65	1.20 (0.53-2.70)	1.20 (0.66-2.16)	0.54
CA+AA	63 (87.5)	6.6	0.55	0.80 (0.35-1.80)	0.95 (0.40-2.27)	0.91
<i>sFRP4</i> rs1802074 G/A	Cases n (%) N=72	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	H.R ^b	<i>p</i>
GG	9 (12.5)	5.1		Reference	Reference	
GA	60 (83.3)	7.1	0.41	0.73 (0.31-1.70)	0.85 (0.37-1.95)	0.71
AA	3 (4.1)	5.7	0.51	0.60 (0.15-2.34)	0.61 (0.02-14.45)	0.76
GA+AA	63 (87.5)	6.9	0.43	0.74 (0.32-1.71)	0.77 (0.33-1.75)	0.53

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.31 (b): Association of *sFRP* SNPs on overall survival according to tumor histology (SQCC)

Genotype	SQCC					
<i>sFRP3</i> rs7775 C/G	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	71 (82.5)	6.2		Reference	Reference	
CG	11 (12.8)	16.2	0.10	0.50 (0.26-0.97)	0.48 (0.20-1.14)	0.10
GG	4 (4.7)	10.7	0.64	0.76 (0.27-2.13)	0.79 (0.43-1.45)	0.73
CG+GG	15 (17.44)	16.1	0.11	0.57 (0.32-1.03)	0.52 (0.25-1.08)	0.08
<i>sFRP3</i> rs288326 C/T	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	Hr^a	HR^b	<i>p</i>
CC	75 (87.2)	6.4		Reference	Reference	
CT	6 (7)	7.2	0.24	1.62 (0.57-4.55)	1.91 (0.75-4.87)	0.17
TT	5 (5.8)	5.4	0.71	0.80 (0.28-2.31)	0.91 (0.51-1.65)	0.77
CT+TT	11 (12.8)	6.8	0.64	1.17 (0.55-2.48)	1.18 (0.56-2.50)	0.64
<i>sFRP4</i> rs1802073 C/A	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
CC	15 (17.44)	5.4		Reference	Reference	
CA	48 (55.8)	6.8	0.76	1.11 (0.56-2.18)	0.96 (0.45-2.04)	0.91
AA	23 (26.7)	6.4	0.41	1.36 (0.65-2.82)	1.09 (0.73-1.63)	0.66
CA+AA	71 (82.5)	6.8	0.69	1.14 (0.60-2.16)	1.05 (0.52-2.12)	0.88
<i>sFRP4</i> rs1802074 G/A	Cases n (%) N=86	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR^a	HR^b	<i>p</i>
GG	13 (15.1)	6.4		Reference	Reference	
GA	69 (80.2)	6.4	0.78	0.91 (0.46-1.78)	0.95 (0.46-1.95)	0.89
AA	4 (4.6)	8.5	0.88	1.08 (0.33-3.48)	0.67 (0.32-1.41)	0.30
GA+AA	73 (84.8)	6.4	0.75	0.90 (0.46-1.76)	0.88 (0.42-1.84)	0.75

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

Table 5.31 (c): Association of *sFRP* SNPs on overall survival according to tumor histology (SCLC)

Genotype	SCLC					
<i>sFRP3</i> rs7775 C/G	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	42 (82.3)	7.2		Reference	Reference	
CG	5 (9.8)	1.5	0.01	2.13 (0.93-4.87)	2.03 (0.97-4.25)	0.05
GG	4 (7.8)	15.1	0.39	0.60 (0.23-1.56)	0.82 (0.45-1.52)	0.54
CG+GG	9 (17.6)	8.0	0.78	0.90 (0.43-1.87)	0.98 (0.43-2.21)	0.96
<i>sFRP3</i> rs288326 C/T	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	Hr ^a	HR ^b	<i>p</i>
CC	46 (90.1)	7.3		Reference	Reference	
CT	2 (3.9)	14.6	0.54	0.65 (0.20-2.09)	0.70 (0.16-3.01)	0.64
TT	3 (5.8)	3.3	0.15	2.23 (0.41-12.11)	1.29 (0.69-2.39)	0.42
CT+TT	5 (9.8)	8.0	0.39	1.47 (0.49-4.39)	1.52 (0.58-3.94)	0.38
<i>sFRP4</i> rs1802073 C/A	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
CC	10 (19.6)	8.2		Reference	Reference	
CA	25 (49)	7.3	0.28	0.67 (0.29-1.52)	0.63 (0.26-1.54)	0.32
AA	16 (31.4)	5.8	0.41	0.73 (0.31-1.68)	1.46 (0.81-2.64)	0.20
CA+AA	41 (80.4)	7.1	0.40	0.75 (0.35-1.60)	1.23 (0.53-2.84)	0.62
<i>sFRP4</i> rs1802074 G/A	Cases n (%) N=51	Univariate analysis			Multivariate analysis	
		MST (months)	Logrank <i>p</i>	HR ^a	HR ^b	<i>p</i>
GG	5 (9.8)	8.0		Reference	Reference	
GA	43 (84.3)	7.2	0.60	1.30 (0.51-3.27)	1.49 (0.48-4.59)	0.48
AA	3 (5.8)	12.9	0.46	0.54 (0.10-2.70)	0.81 (0.26-2.48)	0.72
GA+AA	46 (90.1)	7.3	0.91	0.94 (0.33-2.69)	0.79 (0.24-2.53)	0.70

Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, smoking, stage, KPS, ECOG

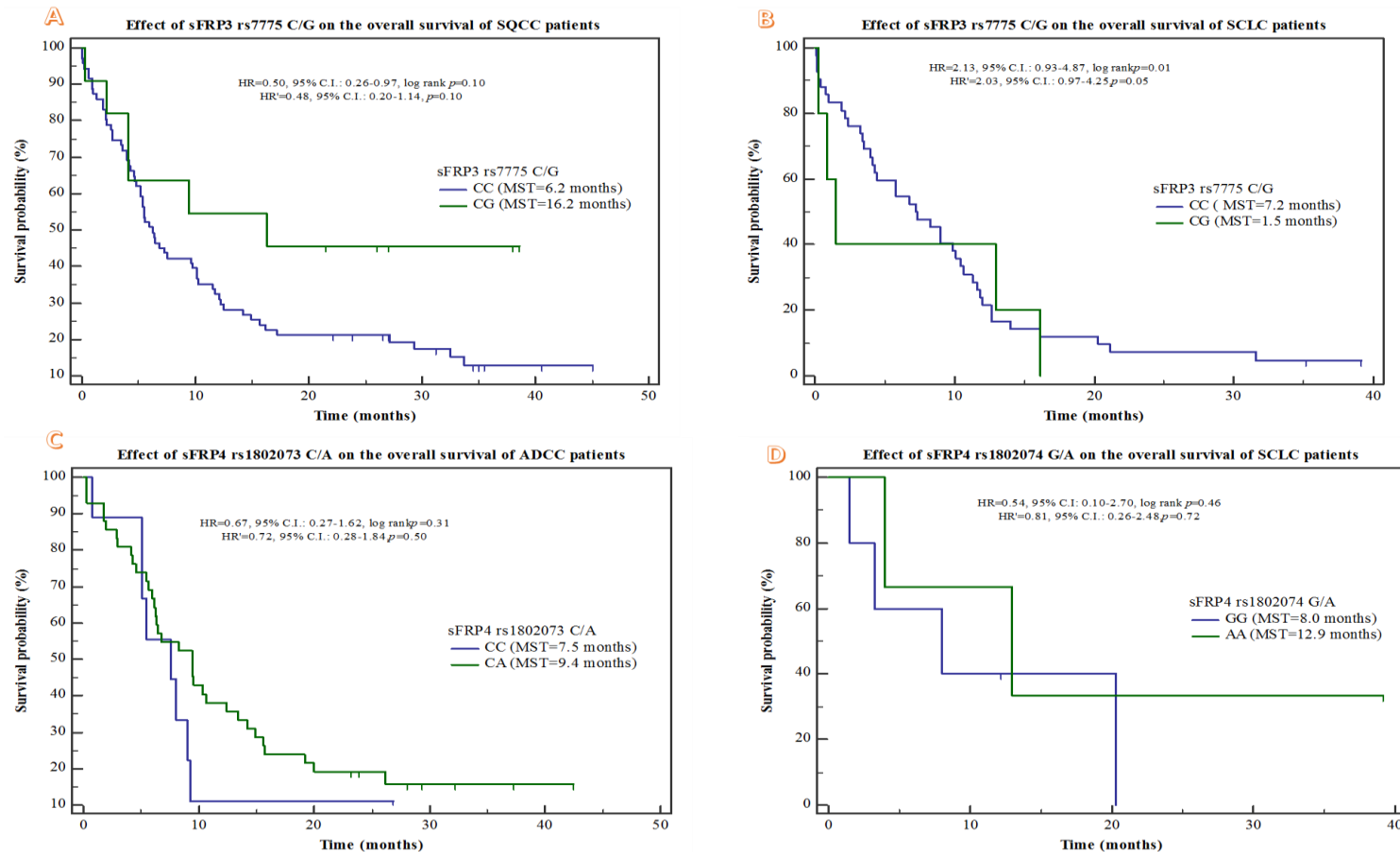


Fig 5.26: Kaplan meir survival curves for *sFRP* variants in lung cancer patients stratified according to their histology; A: *sFRP3* rs7775 C/G (CC: reference) in SQCC patients, B: *sFRP3* rs7775 C/G (CC: reference) in SCLC patients, C: *sFRP3* rs1802073 C/A (CC: reference) in ADCC patients, D: *sFRP4* rs1802074 G/A (GG: reference) in SCLC patients. HR: hazards ratio, HR': adjusted hazards ratio, MST: Median survival time , ADCC: adenocarcinoma, SQCC: squamous cell carcinoma, SCLC: Small cell lung carcinoma.

5.7.9 Association of Wnt antagonist (*sFRP*) polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response

The objective response of the patients was assessed in relation to the genotype of the *sFRP* variants as shown in table 5.42. The patients exhibiting complete and partial response were characterized as good responders and those showing stable and progressive disease were classified as bad responders. The frequency of patients carrying the variant genotype were almost similar in both group of responders, as a result, none of the genotype of *sFRP3* rs228326 C/T, *sFRP3* rs7775 C/G and *sFRP4* rs1802074 G/A could act as a significant predictor of chemotherapeutic response rate or clinical benefit. The frequency of heterozygous (CT) genotype carrying subjects in case of *sFRP3* rs228326 was 8.2% in good responders and 12.3% in bad responders. Similarly, a slightly higher frequency of heterozygotes for *sFRP3* rs7775 were reported among bad responders (19.2%) as compared to good responders (13.7%). Further, in case of *sFRP4* rs1802074 the percentage of GA+AA genotype carrying subjects were similar among good (90.4%) and bad (93.2%) responders. Only *sFRP4* rs1802073 variant was found to modulate the response towards chemotherapy. All the genotypes namely CA, AA and CA+AA genotype were found to enhance the response of these patients. As the a decreased frequency of patients with CA, AA and CA+AA genotype were reported among bad responders (53.4%, 24.7% and 78.1%) as compared to good responders (60.3%, 32.9% and 93.2%) respectively. Hence, the odds ratio obtained in CA, AA and CA+AA genotype carrying patients was 0.19 ($p=0.01$), 0.39 ($p=0.01$) and 0.19 ($p=0.006$) respectively.

5.8 Multiple gene interaction analysis (*sFRP*, *DKK* and *Axin2*)

5.8.1 Demographic characteristics

The multiple gene interaction analysis was carried for the sample population to unravel the interactive potential of these polymorphic sites in modulating lung cancer predisposition and also detecting their synergistic role in predicting clinical outcomes. However, it was a constraint to obtain the genotypic data for each sample for all the polymorphic sites. So, the size of the study population was reduced to 292 cases and 263 controls as for these subjects the genotyping data was available for all polymorphic sites. Study population consists of 292 cases and 263 controls where in the mean age was 57.38 ± 10.74 in cases and 53.23 ± 10.44 in controls. The number of males were 252 (86.3%) among cases and 241 (91%) among controls. The number of smokers were higher in cases (78%) as compared to controls (74%) as smoking is a significant risk factor for lung cancer. The patients were further stratified on the basis of their histological subtype and it was observed that 42.4% of the patients were classified into SQCC category, 32.5% fall under ADCC subgroup and 23.3% fall into SCLC category.

Table 5.32: Relationship of different *sFRP* genotypes with the chemotherapy response

	Response of chemotherapy			
Genotype	Good Responders CR+PR (73) n (%)	Bad Responders SD+PD(73) n (%)	AOR ^a (95% CI) ^b	<i>p</i>
<i>sFRP3</i> rs228326 C/T				
CC	62 (84.9)	60 (82.2)	Reference	
CT	6 (8.2)	9 (12.3)	2.15 (0.64-7.19)	0.21
TT	5 (6.9)	4 (5.5)	0.88 (0.39-1.98)	0.77
CT+TT	11 (15.1)	13 (17.8)	1.58 (0.58-4.29)	0.36
<i>sFRP3</i> rs7775 C/G				
GG	58 (79.5)	56 (76.7)	Reference	
GA	10 (13.7)	14 (19.2)	1.30 (0.49-3.42)	0.58
AA	5 (6.8)	3 (4.1)	0.72 (0.31-1.65)	0.44
GA+AA	15 (20.5)	17 (23.3)	1.06 (0.45-2.49)	0.89
<i>sFRP4</i> rs1802073 C/A				
CC	5 (6.8)	16 (21.9)	Reference	
CA	44 (60.3)	39 (53.4)	0.19 (0.05-0.67)	0.01
AA	24 (32.9)	18 (24.7)	0.39 (0.19-0.81)	0.01
CA+AA	68 (93.2)	57 (78.1)	0.19 (0.05-0.62)	0.006
<i>sFRP4</i> rs1802074 G/A				
GG	7 (9.6)	5 (6.8)	Reference	
GA	60 (82.2)	65 (89)	1.65 (0.45-6.02)	0.44
AA	6 (8.2)	3 (4.2)	0.11 (0.002-5.57)	0.27
GA+AA	66 (90.4)	68 (93.2)	1.51 (0.42-5.41)	0.52

^a Adjusted Odds ratios, ^b 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking, ECOG, KPS, stage, histology and chemotherapeutic regimen. Abbreviations: CR, complete response; PR, partial response; SD, stable disease; PD, progressive disease; AOR, Adjusted Odds ratio; CI, confidence interval.

Table 5.33: Demographic characteristics among cases and controls

Variable	Cases, n (%) N=292	Controls, n (%) N=263	<i>p</i> ^a – value
Age (years)			
Mean ± SD	57.38 ± 10.74	53.23 ± 10.44	p< 0.0001
Range			
Gender			
Male	252 (86)	241 (92)	0.06
Female	40 (14)	22 (8)	
Smoking Status			
Smokers	229 (78)	197 (74)	0.37
Non-Smokers	63 (22)	66 (26)	
Pack Years			
Mean ± SD	28.09 ± 34.09	19.4 ± 19.56	0.0003
Histological Types			
SQCC	124 (42.4)		
ADCC	95 (32.5)		
SCLC	68 (23.3)		
Others	4 (1.4)		
Unknown	1(0.3)		

Abbreviations: SD = Standard Deviation, n = total number of case patients or control subjects ^a*p*-values were derived from Pearson Chi-square test except age; Student t-test was used for age. All *p*-values are two-sided. *p*< 0.05 was considered statistically significant

5.8.2 Genotypic distribution of double combinations of polymorphisms between Wnt antagonist genes (*sFRP*, *DKK* and *Axin2*) and their association with lung cancer predisposition

Table 5.34 shows the possible double combinations between the polymorphic sites of different genes namely *sFRP*, *DKK* and *Axin2*. The association analysis was carried using unconditional multiple logistic regression for all the intergenic genotypic combinations to find out their contribution in modulating predisposition towards lung cancer. The genotypic frequencies of the all combinations have been enlisted in the table 5.42.

The genotypic combination, *sFRP4* rs1802073 C/A and *DKK2* rs419558 C/T was found to greatly influence the risk of lung cancer in north Indians. The frequency of subjects having heterozygous (CA+CT) genotype was higher among cases (56%) as compared to control (40.96%) group. Taking the wild (CC+CC) genotype as reference, logistic regression was applied which revealed that the subjects carrying heterozygous genotype were at almost 3.5 fold increased risk of developing cancer (OR=3.47, 95% C.I.: 1.73-6.16, ***p*=0.004**). Similarly, the subjects harbouring the combined variant *i.e.* heterozygous and mutant (CA+AA+CT+TT) genotype also exhibited 3.8 fold risk of developing cancer (OR=3.83, 95% C.I.: 1.97-7.42, ***p*=0.001**). Further, the double combination of *sFRP4* rs1802073 C/A and *DKK2* rs17037102 G/A exhibited evident association with lung cancer susceptibility. The proportion of heterozygous (CA+GA) subjects was found to be over represented among cases (40.2%) in comparison to controls (29.5%). The association analysis revealed that the subjects carrying the single susceptible allele for this genotypic combination were at a twofold higher risk of acquiring lung cancer (OR=2.27, 95% C.I.: 1.16-4.44, ***p*=0.016**). However, this risk increased to 2.6 fold in case of subjects having the combined variant (CA+AA+GA+AA) genotype (OR=2.66, 95% C.I.: 1.45-4.36, ***p*=0.001**).

As shown in table 5.34, an increased frequency of subjects having heterozygous (GA+CT) genotype for combination *sFRP4* rs1802074 G/A and *DKK2* rs419558 C/T was estimated among cases (73.4%) as compared to controls (45.1%). Henceforth, the heterozygous (GA+CT) genotype conferred a 3.3 fold increased risk for lung cancer in subjects in comparison to those having wild (GG+CC) genotype (OR=3.29, 95% C.I.: 1.65-6.52, ***p*=0.007**). Similarly, subjects carrying both the single and double susceptible allele for this genotypic combination showed a three-fold increased risk of acquiring lung cancer (OR=3.00, 95% C.I.: 1.52-5.94, ***p*=0.001**). The next combination of *sFRP4* rs1802074 G/A and *DKK2* rs17037102 G/A also showed a pronounced effect towards lung cancer risk. The percentage of subjects carrying heterozygous (GA+GA) genotype was 66.3% among cases and 45.2% among controls. The association analysis estimated an increased odds ratio favouring the risk of developing lung cancer in the patients having heterozygous genotype (OR=3.05, 95% C.I.: 1.49-6.25, ***p*=0.002**).

and combined variant genotype (OR=3.09, 95% C.I.: 1.53-6.26, ***p*=0.001**) for this genotypic combination. When the cumulative effect of the combination, *sFRP4* rs1802073 C/A and *DKK3* rs2295199 C/T was evaluated it was found that although the frequency of subjects carrying mutant (AA+TT) genotype was higher in cases (11.26%) than controls (3.6%) but the difference was not found to be statistically significant (OR=3.42, 95% C.I.: 0.79-14.25, *p*=0.09).

The genotypic combination of *sFRP3* rs7775 C/G and *DKK3* rs3206824 G/A was found to modulate the susceptibility of north Indians towards lung cancer. A slightly higher percentage of subjects having the combined variant (CG+GG+GA+AA) genotype was reported among cases (11%) as compared to control (10%) group. Further, these subjects were found at a two-fold higher risk of developing lung cancer as compared to those having wild (CC+GG) genotype for the same genotypic combination (OR=2.06, 95% C.I.: 1.08-3.96, ***p*=0.02**). In case of the genotypic combination, *sFRP4* rs1802074 G/A and *DKK4* rs2073664 C/T it was observed that the frequency of subjects carrying the combined variant (GA+AA+CT+TT) genotype was 86.5% in case group and 78.2% in control group. The difference in the frequency was not found to be statistically significant although a trend of increased risk of developing lung cancer was observed (OR=1.76, 95% C.I.: 0.91-3.28, *p*=0.08). The genotypic combination *sFRP4* rs1802073 C/A and *DKK2* rs447372 A/G showed a marginal the association of this genotypic combination towards lung cancer risk. The frequency of mutants (AA+GG) was represented more among cases (8.2%) as compared to controls (1.9%). Henceforth, it was observed that the subjects carrying both the susceptible alleles for this combination were at 5.8 fold higher risk of developing lung cancer (OR=5.87, 95% C.I.: 0.94-36.68, ***p*=0.05**).

When the other variant of *sFRP4* gene, rs1802074 G/A was evaluated in combination with *DKK2* rs447372 A/G it was observed that the difference in the frequency of subjects carrying combined variant (GA+AA+AG+GG) genotype was not found to be statistically significant between cases and controls (85.5% vs 75% respectively). A visible trend of increased risk of developing lung cancer was observed in these subjects (OR=1.89, 95% C.I.: 0.91-3.91, *p*=0.08).

Further, another combination which was observed to show a cumulative effect on lung cancer predisposition was *sFRP3* rs7775 C/G and *Axin2* 148 C/T. A statistically significant difference was evident in the frequencies of heterozygous (CG+CT) subjects among cases (76.3%) and controls (59.3%). Therefore, the cumulative heterozygous genotype was found to confer 2.3 fold increased risk for developing lung cancer (OR=2.32, 95% C.I.: 1.15-4.67, ***p*=0.01**) as shown in table 5.34. A similar observation was seen in case of subjects carrying combined variant (CG+GG+CT+TT) genotype for this genotypic combination as they were also found at 2.3 fold increased risk of acquiring lung cancer (OR=2.35, 95% C.I.: 1.17-4.74, ***p*=0.01**).

The genotypic combination comprising *sFRP3* rs288326 C/T and *DKK2* rs17037102 G/A also exhibited a positive association between this genotype and lung cancer susceptibility. The percentage of individuals carrying both heterozygous and mutant (CT+TT+GA+AA) genotype was higher among cases (5.5%) as compared to controls (1.1%). The association analysis revealed a 4.6 fold increased risk of developing cancer in subjects having the combined variant genotype for this combination (OR=4.68, 95% C.I.: 0.98-20.29, ***p*=0.05**) as shown in table 5.34.

In case of genotypic combination, *sFRP3* rs7775 C/G and *DKK3* rs7396187 G/C it was observed that the percentage of subjects harbouring single and double both susceptible alleles (CG+GG+GC+CC) for this combination was slightly higher among cases (11%) in comparison to controls (10.1%) pertaining to double the risk of developing lung cancer (OR=2.06, 95% C.I.: 1.08-3.96, ***p*=0.02**) as shown in table 5.34.

The genotypic combination which cumulatively showcased a strong protective effect for lung cancer risk includes *sFRP3* rs7775 C/G and *Axin2* 1365 G/A as shown in table 5.34. The percentage of the subjects carrying heterozygous (CG+GA) was higher among controls (7.7%) in comparison to cases (3.8%). Hence, a strong protective effect was evident in these subjects with OR=0.39, 95% C.I.: 0.15-0.99, ***p*=0.04**. Similarly, in case of combination of *sFRP3* rs7775 C/G and *Axin2* 956+16 A/G the frequency of heterozygous (CG+AG) subjects was lower among cases (2.1%) as compared to controls (6.7%). The association analysis showed a strong significant protective effect towards lung cancer risk in the subjects carrying heterozygous genotype for this genotypic combination (OR=0.29, 95% C.I.: 0.09-0.94, ***p*=0.039**).

Further, in case of genotypic combination *sFRP4* rs1802073 C/A and *Axin2* 1386 C/T an increased frequency of subjects carrying mutant (AA+TT) genotype was observed among cases (10.8%) in comparison to controls (5.8%). When logistic regression was applied by considering CC+CC as reference, a 1.8 fold risk of developing lung cancer was estimated in subjects carrying the mutant genotype for this combination (OR=1.81, 95% C.I.: 0.98-3.33, ***p*=0.05**) as shown in table 5.34.

Table 5.34: Genotypic distribution based on different genotypic combinations and their association with lung cancer risk

Genotype					
<i>sFRP4</i> rs1802073 C/A & <i>DKK2</i> rs419558 C/T	Cases, n (%) N =100	Controls, n (%) N =83	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	24 (24)	40 (48.19)	1.00 (Reference)		
1	56 (56)	34 (40.96)	3.47 (1.73-6.16)	0.004	0.10
3	76 (76)	43 (51.81)	3.83 (1.97-7.42)	0.001	0.01
<i>sFRP4</i> rs1802073 C/A & <i>DKK2</i> rs17037102 G/A	Cases, n (%) N =112	Controls, n (%) N =88	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	36 (32.14)	51 (57.95)	1.00 (Reference)		
1	45 (40.17)	26 (29.54)	2.27(1.16-4.44)	0.016	0.16
3	76 (67.85)	37 (42.04)	2.66(1.45-4.36)	0.001	0.01
<i>sFRP4</i> rs1802074 G/A & <i>DKK2</i> rs419558 C/T	Cases, n (%) N =109	Controls, n (%) N =91	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	25 (22.93)	42 (46.15)	1.00 (Reference)		
1	80 (73.39)	41 (45.05)	3.29(1.65-6.52)	0.007	0.11
3	84 (77.06)	49 (53.84)	3.00(1.52-5.94)	0.0016	0.01
<i>sFRP4</i> rs1802074 G/A & <i>DKK2</i> rs17037102 G/A	Cases, n (%) N =95	Controls, n (%) N = 84	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	26 (27.36)	43 (51.19)	1.00 (Reference)		
1	63 (66.31)	38 (45.23)	3.05(1.49-6.25)	0.002	0.1
3	69 (72.63)	41 (48.80)	3.09(1.53-6.26)	0.0016	0.01
<i>sFRP4</i> rs1802073 C/A & <i>DKK3</i> rs2295199 C/T	Cases, n (%) N =71	Controls, n (%) N = 83	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	24 (33.80)	40 (48.19)	1.00 (Reference)		
1	39 (54.92)	40 (48.19)	1.34(0.64-2.77)	0.42	0.81
2	8 (11.26)	3 (3.61)	3.42(0.79-14.25)	0.09	0.36
3	47 (66.19)	43 (51.81)	1.49(0.73-3.03)	0.26	0.52
<i>sFRP4</i> rs1802073 C/A & <i>DKK3</i> rs7396187 G/C	Cases, n (%) N =101	Controls, n (%) N = 111	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	32 (31.68)	39 (35.13)	1.00 (Reference)		
1	43 (42.57)	98 (88.28)	0.89 (0.47-1.68)	0.72	0.91
3	69 (68.31)	72 (64.86)	1.14 (0.63-2.07)	0.65	0.83
<i>sFRP4</i> rs1802073 C/A & <i>DKK3</i> rs3206824 G/A	Cases, n (%) N =82	Controls, n (%) N = 91	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	34 (41.46)	45 (49.45)	1.00 (Reference)		
1	46 (56.09)	46 (50.55)	1.21 (0.63-2.31)	0.55	0.87
3	48 (58.53)	46 (50.55)	1.27 (0.67-2.42)	0.45	0.66
<i>sFRP4</i> rs1802074 G/A & <i>DKK3</i> rs2295199 C/T	Cases, n (%) N =87	Controls, n (%) N = 107	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	21 (24.13)	35 (32.71)	1.00 (Reference)		
1	60 (68.96)	55 (51.40)	1.53 (0.72-3.28)	0.26	0.72
3	66 (75.86)	72 (67.29)	1.33 (0.64-2.76)	0.43	0.66
<i>sFRP4</i> rs1802074 G/A & <i>DKK3</i> rs3206824 G/A	Cases, n (%) N =93	Controls, n (%) N = 86	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	20 (21.50)	34 (39.53)	1.00 (Reference)		
1	71 (76.34)	51 (59.30)	2.32 (1.15-4.67)	0.01	0.12
2	2 (2.15)	1 (1.16)			
3	73 (78.49)	52 (60.47)	2.35 (1.17-4.74)	0.01	0.08
<i>sFRP4</i> rs1802074 G/A & <i>DKK3</i> rs7396187 G/C	Cases, n (%) N =101	Controls, n (%) N = 105	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>

0	28 (27.72)	30 (28.57)	1.00 (Reference)		
1	68 (67.32)	75 (71.43)	0.96 (0.50-1.83)	0.90	0.95
3	73 (72.27)	75 (71.43)	1.02 (0.54-1.75)	0.92	0.97
sFRP3 rs288326 C/T & DKK2 rs419558 C/T	Cases, n (%) N =173	Controls, n (%) N = 188	AOR (95% CI)^b	p^a	FDR p
0	162 (93.64)	181 (96.27)	1.00 (Reference)		
1	6 (3.46)	5 (2.65)	1.66 (0.48-5.76)	0.42	0.81
3	11 (6.35)	7 (3.73)	1.83 (0.68-4.97)	0.22	0.52
sFRP3 rs288326 C/T & DKK2 rs17037102 G/A	Cases, n (%) N =183	Controls, n (%) N = 189	AOR (95% CI)^b	p^a	FDR p
0	173 (94.53)	187 (98.94)	1.00 (Reference)		
1	5 (2.73)	2 (1.06)	2.65 (0.48-14.38)	0.25	0.72
3	10 (5.46)	2 (1.06)	4.68 (0.98-20.29)	0.05	0.26
sFRP3 rs7775 C/G T & DKK2 rs419558 C/T	Cases, n (%) N =171	Controls, n (%) N = 181	AOR (95% CI)^b	p^a	FDR p
0	153 (89.47)	164 (90.60)	1.00 (Reference)		
1	15 (8.77)	12 (6.62)	1.40 (0.61-3.18)	0.42	0.81
3	18 (10.52)	17 (9.40)	1.08 (0.53-2.27)	0.80	0.87
sFRP3 rs7775 C/G & DKK2 rs17037102 G/A	Cases, n (%) N =175	Controls, n (%) N = 170	AOR (95% CI)^b	p^a	FDR p
0	161 (92)	163 (95.88)	1.00 (Reference)		
1	10 (5.71)	5 (2.94)	1.90 (0.51-5.15)	0.26	0.72
3	14 (8)	7 (4.12)	1.70 (0.64-4.47)	0.28	0.52
sFRP3 rs288326 C/T & DKK4 rs2073664 C/T	Cases, n (%) N =142	Controls, n (%) N = 126	AOR (95% CI)^b	p^a	FDR p
0	126 (88.73)	117 (92.85)	1.00 (Reference)		
1	15 (10.56)	9 (7.15)	1.36 (0.56-3.29)	0.48	0.85
3	16 (11.26)	9 (7.15)	1.45 (0.60-3.46)	0.39	0.64
sFRP3 rs7775 C/G & DKK4 rs2073664 C/T	Cases, n (%) N =146	Controls, n (%) N = 138	AOR (95% CI)^b	p^a	FDR p
0	116 (79.45)	106 (76.81)	1.00 (Reference)		
1	23 (15.75)	22 (15.94)	0.89 (0.46-1.73)	0.74	0.91
3	30 (20.54)	32 (23.19)	0.77 (0.43-1.39)	0.39	0.64
sFRP4 rs1802073 C/A & DKK4 rs2073664 C/T	Cases, n (%) N =101	Controls, n (%) N = 131	AOR (95% CI)^b	p^a	FDR p
0	24 (23.76)	51 (38.93)	1.00 (Reference)		
1	74 (73.26)	78 (59.54)	1.16 (0.61-2.21)	0.63	0.90
2	3 (2.97)	2 (1.52)	1.91 (0.28-12.84)	0.50	0.70
3	77 (76.23)	80 (61.07)	1.19 (0.63-2.26)	0.57	0.78
sFRP4 rs1802074 G/A & DKK4 rs2073664 C/T	Cases, n (%) N =148	Controls, n (%) N = 142	AOR (95% CI)^b	p^a	FDR p
0	20 (13.51)	31 (21.83)	1.00 (Reference)		
1	111 (75)	92 (64.78)	1.82 (0.93-3.53)	0.07	0.50
3	128 (86.48)	111 (78.16)	1.76 (0.91-3.28)	0.08	0.29
sFRP3 rs288326 C/T & DKK2 rs447372 A/G	Cases, n (%) N =140	Controls, n (%) N = 117	AOR (95% CI)^b	p^a	FDR p
0	120 (85.71)	105 (89.74)	1.00 (Reference)		
1	12 (8.57)	8 (6.83)	1.49 (0.57-3.90)	0.40	0.81
3	20 (14.28)	12 (10.25)	1.58 (0.72-3.44)	0.24	0.52
sFRP3 rs7775 C/G & DKK2 rs447372 A/G	Cases, n (%) N =141	Controls, n (%) N = 121	AOR (95% CI)^b	p^a	FDR p
0	118 (83.68)	99 (81.81)	1.00 (Reference)		
1	21 (14.89)	21 (17.35)	0.84 (0.42-1.66)	0.62	0.90
2	2 (1.41)	1 (0.82)			
3	23 (16.31)	22 (18.19)	0.90 (0.46-1.73)	0.75	0.87
sFRP4 rs1802073 C/A & DKK2 rs447372 A/G	Cases, n (%) N =98	Controls, n (%) N = 106	AOR (95% CI)^b	p^a	FDR p
0	22 (22.44)	30 (28.30)	1.00 (Reference)		
1	68 (69.38)	74 (69.81)	1.43 (0.72-2.87)	0.30	0.79
2	8 (8.16)	2 (1.88)	5.87 (0.94-36.68)	0.05	0.3

3	76 (77.55)	76 (71.70)	1.57 (0.79-3.14)	0.19	0.51
<i>sFRP4</i> rs1802074 G/A & <i>DKK2</i> rs447372 A/G	Cases, n (%) N =113	Controls, n (%) N = 116	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	16 (14.15)	29 (25)	1.00 (Reference)		
1	95 (84.07)	85 (73.27)	1.87 (0.90-3.89)	0.09	0.50
2	2 (1.76)	2 (1.72)	2.69 (0.30-23.70)	0.37	0.63
3	97 (85.84)	87 (75)	1.89 (0.91-3.91)	0.08	0.29
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 148 C/T	Cases, n (%) N =99	Controls, n (%) N = 80	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	82 (82.82)	65 (81.25)	1.00 (Reference)		
1	12 (12.12)	14 (17.5)	0.81 (0.34-1.96)	0.65	0.90
2	5 (5.05)	1 (1.25)	3.65 (0.39-33.61)	0.25	0.5
3	17 (17.17)	15 (18.75)	1.00 (0.45-2.25)	0.98	0.98
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 148 C/T	Cases, n (%) N =97	Controls, n (%) N = 86	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	71 (73.19)	62 (72.09)	1.00 (Reference)		
1	23 (23.71)	19 (22.09)	1.02 (0.49-2.14)	0.93	0.95
2	3 (3.09)	5 (5.81)	0.40 (0.08-1.92)	0.25	0.5
3	26 (26.80)	24 (27.91)	0.87 (0.43-1.75)	0.70	0.86
<i>sFRP3</i> rs288326 C/T & <i>DKK3</i> rs2295199 C/T	Cases, n (%) N =178	Controls, n (%) N = 162	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	161 (90.44)	153 (94.44)	1.00 (Reference)		
1	6 (3.37)	5 (3.08)	1.06 (0.31-3.64)	0.91	0.95
3	17 (9.55)	9 (5.55)	1.76 (0.74-4.18)	0.19	0.51
<i>sFRP3</i> rs288326 C/T & <i>DKK3</i> rs3206824 G/A	Cases, n (%) N =181	Controls, n (%) N = 176	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	166 (91.71)	165 (93.75)	1.00 (Reference)		
1	9 (4.97)	9 (5.11)	?	0.099	0.50
3	15 (8.28)	11 (6.25)	1.37 (0.31-1.47)	0.44	0.66
<i>sFRP3</i> rs288326 C/T & <i>DKK3</i> rs7396187 G/C	Cases, n (%) N =191	Controls, n (%) N = 156	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	178 (93.19)	146 (93.58)	1.00 (Reference)		
1	6 (3.14)	8 (5.12)	0.60 (0.19-1.64)	0.37	0.81
3	13 (6.80)	10 (6.42)	0.98 (0.40-2.38)	0.97	0.98
<i>sFRP3</i> rs7775 C/G & <i>DKK3</i> rs2295199 C/T	Cases, n (%) N =172	Controls, n (%) N = 157	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	150 (87.20)	137 (87.26)	1.00 (Reference)		
1	11 (6.39)	11 (7.00)	0.92 (0.38-2.22)	0.86	0.95
3	22 (12.79)	20 (12.74)	0.91 (0.47-1.78)	0.79	0.87
<i>sFRP3</i> rs7775 C/G & <i>DKK3</i> rs3206824 G/A	Cases, n (%) N =173	Controls, n (%) N = 159	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	154 (89.01)	143 (89.93)	1.00 (Reference)		
1	12 (6.93)	10 (6.28)	1.15 (0.47-2.81)	0.75	0.91
3	19 (10.98)	16 (10.07)	2.06 (1.08-3.96)	0.02	0.14
<i>sFRP3</i> rs7775 C/G & <i>DKK3</i> rs7396187 G/C	Cases, n (%) N =191	Controls, n (%) N = 153	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	170 (89)	131 (85.62)	1.00 (Reference)		
1	13 (6.80)	17 (11.11)	0.54 (0.24-1.16)	0.11	0.50
3	21 (11)	22 (14.38)	0.62 (0.32-1.22)	0.17	0.51
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 148 C/T	Cases, n (%) N =105	Controls, n (%) N = 99	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	14 (13.33)	18 (18.18)	1.00 (Reference)		
1	77 (73.33)	73 (73.73)	1.32 (0.60-2.90)	0.48	0.86
2	14 (13.33)	8 (8.08)	2.01 (0.62-6.45)	0.23	0.5
3	91 (86.67)	81 (81.82)	1.40 (0.64-3.06)	0.39	0.64
<i>sFRP4</i> rs1802074 G/A & <i>Axin2</i> 148 C/T	Cases, n (%) N =133	Controls, n (%) N = 101	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	10 (7.51)	12 (11.88)	1.00 (Reference)		
1	118 (88.72)	81 (80.19)	1.89 (0.75-4.73)	0.17	0.61
2	5 (3.75)	8 (7.92)	1.49 (0.26-8.36)	0.64	0.7

3	123 (92.48)	89 (88.12)	1.85 (0.73-4.65)	0.19	0.51
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 1365 G/A	Cases, n (%) N =226	Controls, n (%) N = 191	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	219 (96.90)	186 (97.38)	1.00 (Reference)		
1	2 (0.88)	3 (1.57)	0.66 (0.09-4.44)	0.67	0.90
3	7 (3.09)	5 (2.62)	1.15 (0.34-3.82)	0.817	0.87
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 1365 G/A	Cases, n (%) N =212	Controls, n (%) N = 182	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	204 (96.22)	168 (92.30)	1.00 (Reference)		
1	6 (2.83)	10 (5.49)	0.45 (0.15-1.31)	0.14	0.58
3	8 (3.77)	14 (7.69)	0.39 (0.15-0.99)	0.04	0.23
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 1365 G/A	Cases, n (%) N =56	Controls, n (%) N = 77	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	38 (67.85)	48 (62.33)	1.00 (Reference)		
1	16 (28.57)	28 (36.36)	0.60 (0.27-1.31)	0.20	0.67
2	2 (3.57)	1 (1.29)	1.36 (0.39-4.70)	0.61	0.7
3	18 (32.14)	29 (37.66)	0.64 (0.30-1.38)	0.25	0.52
<i>sFRP4</i> rs1802074 G/A & <i>Axin2</i> 1365 G/A	Cases, n (%) N =59	Controls, n (%) N = 70	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	32 (54.23)	39 (55.71)	1.00 (Reference)		
1	26 (44.06)	30 (42.85)	0.77 (0.35-1.71)	0.52	0.87
2	1 (1.69)	1 (1.42)	1.32 (0.31-5.63)	0.70	0.7
3	27 (45.76)	31 (44.29)	0.77 (0.36-1.74)	0.57	0.78
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 432 T/C	Cases, n (%) N =241	Controls, n (%) N = 218	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	240 (99.58)	217 (99.54)	1.00 (Reference)		
1	1 (0.42)	1 (0.46)	1.25 (0.07-21.89)	0.87	0.95
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 432 T/C	Cases, n (%) N =224	Controls, n (%) N = 197	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	223 (99.55)	193 (97.96)	1.00 (Reference)		
1	1 (0.45)	4 (2.04)			
3	1 (0.45)	4 (2.04)	0.24 (0.026-2.21)	0.20	0.51
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 432 T/C	Cases, n (%) N =52	Controls, n (%) N = 67	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	44 (84.61)	57 (85.07)	1.00 (Reference)		
1	4 (7.69)	7 (10.44)	0.85 (8.21-3.37)	0.82	0.95
3	8 (15.38)	10 (14.92)	1.27 (0.43-3.76)	0.65	0.83
<i>sFRP4</i> rs1802074 G/A & <i>Axin2</i> 432 T/C	Cases, n (%) N =41	Controls, n (%) N = 57	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	36 (87.80)	47 (82.45)	1.00 (Reference)		
1	5 (12.20)	10 (17.54)	0.94 (0.26-3.39)	0.93	0.95
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 956+16 A/G	Cases, n (%) N =209	Controls, n (%) N = 190	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	206 (98.56)	187 (98.42)	1.00 (Reference)		
1	3 (1.44)	3 (1.57)	1.30 (0.24-6.81)	0.75	0.91
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 956+16 A/G	Cases, n (%) N =192	Controls, n (%) N = 178	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	188 (97.91)	166 (93.25)	1.00 (Reference)		
1	4 (2.09)	12 (6.74)	0.29 (0.09-0.94)	0.039	0.32
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 956+16 A/G	Cases, n (%) N =59	Controls, n (%) N = 77	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	39 (66.10)	50 (64.93)	1.00 (Reference)		
1	20 (33.9)	27 (35.06)	0.80 (0.37-1.72)	0.56	0.87
<i>sFRP4</i> rs1802074 G/A & <i>Axin2</i> 956+16 A/G	Cases, n (%) N =66	Controls, n (%) N = 70	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	31 (46.96)	39 (55.71)	1.00 (Reference)		
1	35 (53.03)	31 (44.28)	1.43 (0.69-2.95)	0.32	0.80
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 2062 C/T	Cases, n (%) N =208	Controls, n (%) N = 199	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>
0	205 (98.55)	194 (97.48)	1.00 (Reference)		
1	3 (1.45)	5 (2.51)	0.77 (0.14-2.92)	0.56	0.87
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 2062 C/T	Cases, n (%) N =192	Controls, n (%) N = 183	AOR (95% CI) ^b	<i>p</i> ^a	FDR <i>p</i>

0	189 (98.43)	175 (95.62)	1.00 (Reference)		
1	3 (1.57)	8 (4.37)	0.37 (0.09-1.44)	0.15	0.50
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 2062 C/T	Cases, n (%) N =65	Controls, n (%) N = 81	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	38 (58.46)	51 (62.96)	1.00 (Reference)		
1	27 (41.53)	30 (37.03)	0.99 (0.48-2.03)	0.99	0.99
<i>sFRP4</i> rs1802074 G/A & <i>Axin2</i> 2062 C/T	Cases, n (%) N =62	Controls, n (%) N = 74	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	26 (41.93)	45 (60.81)	1.00 (Reference)		
1	36 (58.06)	29 (39.18)	1.22 (0.57-2.59)	0.60	0.90
<i>sFRP3</i> rs288326 C/T & <i>Axin2</i> 1386 C/T	Cases, n (%) N =88	Controls, n (%) N = 89	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	76 (86.36)	81 (91.01)	1.00 (Reference)		
1	11 (12.5)	8 (8.99)	1.55 (0.56-4.31)	0.39	0.81
2	1 (1.13)	0 (0)			
3	12 (13.64)	8 (8.99)	1.71 (0.63-4.66)	0.28	0.52
<i>sFRP3</i> rs7775 C/G & <i>Axin2</i> 1386 C/T	Cases, n (%) N =95	Controls, n (%) N = 101	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	67 (70.52)	74 (73.26)	1.00 (Reference)		
1	26 (27.36)	26 (25.74)	1.09 (0.57-2.10)	0.77	0.91
2	2 (2.10)	1 (0.99)	1.32 (0.38-4.50)	0.65	0.7
3	28 (29.47)	27 (26.73)	1.12 (0.59-2.12)	0.71	0.86
<i>sFRP4</i> rs1802073 C/A & <i>Axin2</i> 1386 C/T	Cases, n (%) N =111	Controls, n (%) N = 120	AOR (95% CI)^b	<i>p</i>^a	<i>FDR p</i>
0	12 (10.81)	23 (19.16)	1.00 (Reference)		
1	87 (78.37)	90 (75)	1.90 (0.86-4.21)	0.10	0.50
2	12 (10.81)	7 (5.83)	1.81 (0.98-3.33)	0.05	0.30
3	99 (89.18)	97 (80.84)	2.06 (0.94-4.55)	0.069	0.29

^aTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. ^bAdjusted Odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 3: combined hetero and mutant genotype. *FDR p*: False discovery rate adjusted *p*.

5.8.3 Association of different unfavorable genotypic combinations of *DKK2*, *DKK3*, *sFRP3* and *sFRP4* gene with overall survival of lung cancer patients

In order to find out the cumulative effect of the polymorphic variants of each gene, different genotypic combination of unfavorable genotypes of the wnt antagonist polymorphic variants were analyzed for predicting the overall survival of lung cancer patients as demonstrated in table 5.35. In case of *DKK2* gene, the genotypic combination comprising the unfavourable genotypes for both rs419558 and rs447372 SNP showed a better survival time (9.6 months vs 5.5 months) as compared to the patients lacking any unfavourable genotypes for three variants. Hazards ratio obtained from univariate analysis was not found to be statistically significant (log rank p -value=0.22) but when it was adjusted for several covariates the death rate was found to decrease significantly (adjusted HR=0.53, $p=0.01$). Patients harboring all the three unfavorable genotypes for *DKK2* variants showed a significantly better prognosis (adjusted HR=0.37, $p=0.03$) during multivariate analysis. In case of *DKK3* gene, none of the combination of unfavourable genotype showed any cumulative effect on the survival of lung cancer patients. A similar trend was observed in case of *sFRP3* and *sFRP4* genetic variants where none of the combination of unfavourable genotypes was found to act as an independent predictor of survival. Several different combinations of unfavourable genotypes were also examined for SNP's of *Axin2* gene, but none of them showed any significant association with prognosis of the patients (data not shown).

Table 5.35: Association of combination of unfavorable genotypes with Overall Survival

Number of unfavourable genotypes	Cases (N)	MST (months)	^a Unadjusted HR	Log rank p	^b Adjusted HR	p	FDR p
<i>DKK2</i>							
0	51	5.5	Reference		Reference		
2 (rs419558 & rs17037102)	15	6.8	1.03 (0.55-1.92)	0.91	0.64 (0.31-1.33)	0.23	0.23
2 (rs447372 & rs17037102)	22	9	0.80 (0.48-1.35)	0.43	0.58 (0.31-1.08)	0.09	0.12
2 (rs419558 & rs447372)	47	9.6	0.76 (0.49-1.18)	0.22	0.53 (0.33-0.85)	0.01	0.04
3	9	9	0.72 (0.35-1.47)	0.41	0.37 (0.14-0.95)	0.03	0.06
<i>DKK3</i>							
0	99	5.5	Reference		Reference		
2 (rs3206824 & rs7396187)	38	9.2	0.77 (0.52-1.14)	0.22	0.81 (0.53-1.24)	0.34	0.45
2 (rs3206824 & rs2291599)	48	9.2	0.73 (0.50-1.05)	0.11	0.74 (0.49-1.11)	0.15	0.45
2 (rs7396187 & rs2291599)	48	7.3	0.89 (0.62-1.30)	0.57	0.90 (0.61-1.33)	0.62	0.62
3	33	10.1	0.75 (0.50-1.14)	0.21	0.79 (0.50-1.24)	0.31	0.45
<i>sFRP3</i>							
0	137	6.4	Reference		Reference		
2	4	6.08	0.89 (0.29-2.41)	0.76	0.76 (0.23-2.47)	0.65	0.65
<i>sFRP4</i>							
0	5	5.3	Reference		Reference		
2	155	6.7	0.75 (0.24-2.35)	0.58	0.95 (0.33-2.73)	0.93	0.93

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, histology, smoking, stage, KPS, ECOG
0: wild genotype for all SNPs of a gene, 2: variant (Hetero+mutant) genotype of any two variants, 3: variant (Hetero+mutant) genotype of all the three variants, FDR (False discovery rate) adjusted p-value.

5.8.4 High order gene-gene interactions for prognosis of lung cancer patients using survival tree analysis

Multiple interaction studies using different genes involved in the Wnt signalling pathway is a new dimension to find out the high order parametric interactions taking place within them and effect of these interactions on OS of the patients. In the present study, the recursive partitioning approach was used to create a survival tree which depicts the involvement of different SNPs in prognosis. The tree obtained by STREE software is illustrated in fig 5.27. Out of all the algorithms available, log-rank method was used to create the tree. A total of nine terminal nodes were observed. The most significant factor contributing to the prognosis of patients was *DKK4* rs2073664, as it is root node. The terminal node 7 with shortest MST (3 months) was taken as a reference. MST corresponding to each terminal node is shown in table 5.36. Also, the hazards ratio for each node is mentioned. Fig 5.28 shows the comparison of survival curves for different nodes. It depicts a significant difference between the MST of different nodes when compared to reference (log rank $p < 0.0001$). Table 5.36 summarizes the results of cox-regression analysis for each terminal node. For cox-regression analysis, the hazards ratio was adjusted for various covariates such as age, sex, smoking, histology, ECOG, KPS and stage. All the nodes showed a significantly good prognosis for lung cancer patients except terminal node 4. Node 3 exhibited the lowest death rate ($HR'=0.04$, $p=0.008$) and better MST (9 vs 3 months) when compared to reference node. A similar trend was observed in case of all other nodes showing good prognostic effect.

Table 5.36: High order gene-gene interactions using survival tree analysis

Nodes	Genotype	N(Dead/Total)	MST(months)	^aUnadjusted HR	Log rank <i>p</i>	^bAdjusted HR	<i>p</i>
Node 1	<i>DKK2</i> rs419558 (M)/ <i>DKK4</i> rs2073664 (M)	26/34	9.0	0.25(0.07-0.93)	0.0002	0.11(0.03-0.37)	0.0003
Node 2	<i>DKK2</i> rs447372(W)/ <i>DKK4</i> rs2073664(W)	45/46	5.5	0.45(0.16-1.27)	0.03	0.32(0.13-0.78)	0.01
Node 3	<i>sFRP4</i> rs1802073(W)/ <i>DKK2</i> rs419558 (W)/ <i>DKK4</i> rs2073664 (M)	5/9	9.0	0.23(0.06-0.79)	0.003	0.04(0.005-0.45)	0.008
Node 4	<i>Axin2</i> 1712+19(M)/ <i>sFRP4</i> rs1802073(M)/ <i>DKK2</i> rs419558 (W)/ <i>DKK4</i> rs2073664 (M)	35/38	4.9	0.49(0.18-1.33)	0.05	0.44(0.17-1.16)	0.10
Node 5	<i>DKK3</i> rs3206824(M)/ <i>sFRP3</i> rs7775(W)/ <i>DKK2</i> rs447372(M)/ <i>DKK4</i> rs2073664(W)	14/15	10.7	0.23(0.06-0.86)	0.001	0.05(0.01-0.29)	0.0006
Node 6	<i>Axin2</i> 1712+19 (M)/ <i>DKK3</i> rs3206824(W)/ <i>sFRP3</i> rs7775(W)/ <i>DKK2</i> rs447372(M)/ <i>DKK4</i> rs2073664(W)	23/26	6.6	0.41(0.14-1.20)	0.02	0.29(0.09-0.88)	0.029
Node 7	<i>Axin2</i> 1712+19 (W)/ <i>DKK3</i> rs3206824(W)/ <i>sFRP3</i> rs7775(W)/ <i>DKK2</i> rs447372(M)/ <i>DKK4</i> rs2073664(W)	8/8	3.0	Reference		Reference	
Node 8	<i>Axin2</i> 1712+19(W)/ <i>sFRP4</i> rs1802073(M)/ <i>DKK2</i> rs419558 (W)/ <i>DKK4</i> rs2073664 (M)	14/19	10.7	0.26(0.07-0.91)	0.0006	0.05(0.01-0.25)	0.0002
Node 9	<i>sFRP3</i> rs7775(M)/ <i>DKK2</i> rs447372(M)/ <i>DKK4</i> rs2073664(W)	11/17	10.3	0.31(0.09-0.90)	0.005	0.16(0.04-0.63)	0.008

^a Unadjusted Hazards ratio for Kaplan meier analysis, ^b hazards ratio adjusted for age, sex, histology, smoking, stage, KPS, ECOG.

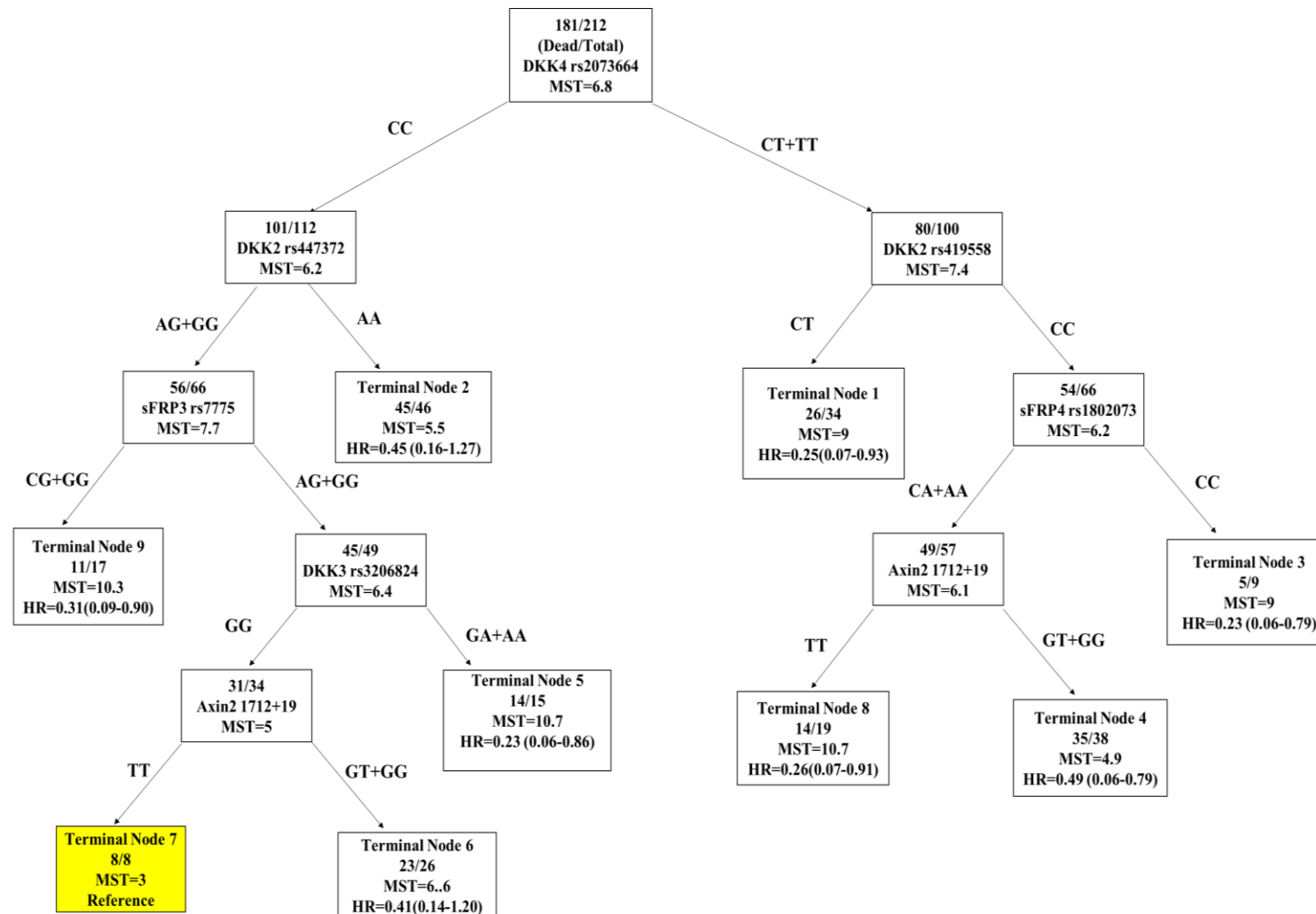


Fig 5.27: Survival Tree analysis of *sFRP*, *DKK*, and *Axin2* variants; MST: median survival time in months; HR: Hazards ratio with 95%C.I (univariate analysis); yellow color shows reference terminal node.

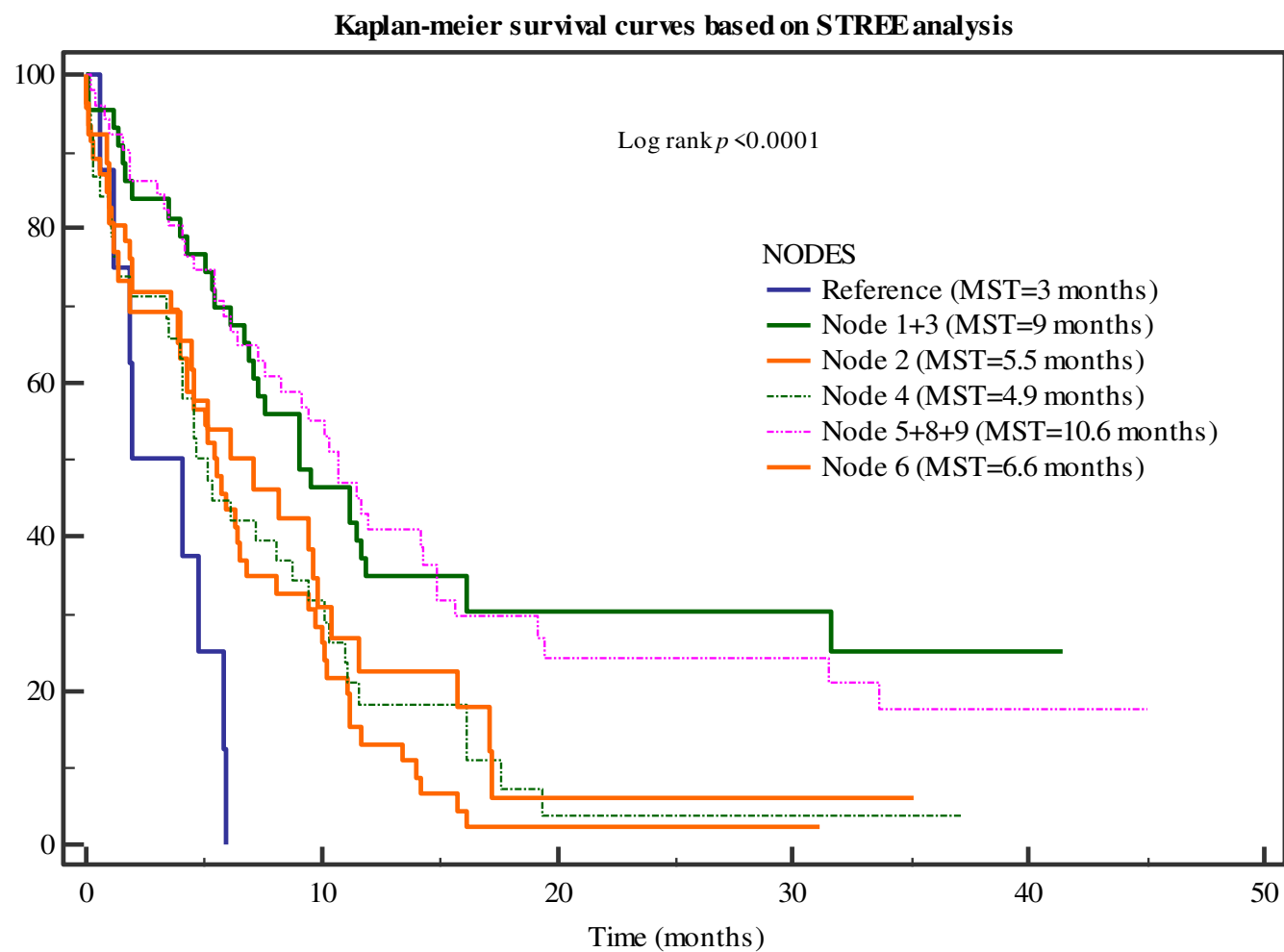


Fig 5.28: Overall survival according to different nodes identified by survival tree analysis

5.9 Methylation-Specific Polymerase Chain Reaction (MSP) Analysis for *DKK*, *sFRP* and *Axin2* promoter region

The methylation profile of *DKK2*, *DKK3*, *Axin2* and *sFRP1* gene was analyzed using MS-PCR for plasma samples of 100 lung cancer patients. The methylation specific PCR of promoter regions of these genes was carried out and methylation status was deciphered by resolving the samples by agarose gel electrophoresis. The negative control used in this case was the use of water instead of DNA as a template during PCR. The methylation status of the samples was recorded by scoring the gels for absence and presence of amplified product specific for methylated and unmethylated primer for each gene.

The representative example of agarose gel electrophoresis for MS-PCR of *DKK2*, *DKK3*, *Axin2* and *sFRP1* gene is illustrated in fig 5.29. The lanes are marked as methylated (M) and unmethylated (UM) in accordance to the sample loaded.

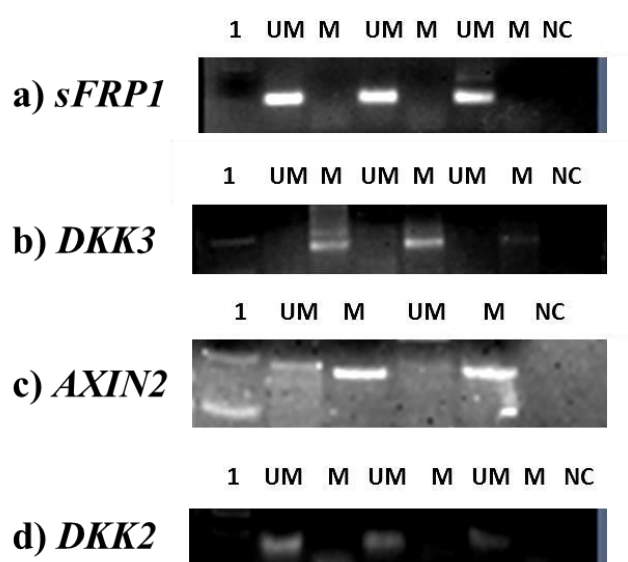


Fig 5.29: Representative example of Methylation specific PCR; a: MS-PCR of *sFRP1* gene, b: MS-PCR of *DKK3* gene, c: MS-PCR of *Axin2* gene, d: MS-PCR of *DKK2* gene. 1: marker DNA, UM: unmethylated primer, M: methylated primer, NC: negative control

The demographic and clinical characteristics of the patients which were analyzed for their methylation profile for *sFRP1*, *Axin2*, *DKK3* and *DKK2*. The distribution of the patients in accordance with these factors is listed in table 5.37. The mean age of the 100 lung cancer patients came out to be 57.87 ± 10.81 . Further, the patients were stratified according to their gender and it was observed that 92% of the sample population comprised of males and rest 8% of females. Smoking, being an established causative agent therefore it was essential to stratify the patients on the basis of their smoking habit. Out of 100 patients, 87% were reported as smokers and rest 13% as non-smokers. Further, clinical parameters were also

evaluated for these patients. The histological classification of patients reported that 27% were reported with ADCC and 38% with SQCC followed with 35% fall under SCLC category.

Table 5.37 Distribution of Demographic and Clinical parameters of lung cancer patients analyzed for their methylation profile

Variables	Total (N)	Cases, n (%)
Mean $\pm$ SD	100	57.87 $\pm$ 10.81
Range		28-86
Male	100	92 (92)
Female		8 (8)
Smokers	100	87 (87)
Non-smokers		13 (13)
Histology		
ADCC	100	27(27)
SCLC		33 (33)
SQCC		40 (40)
TNM Staging		
I + II	100	5 (5)
III		55 (55)
IV		40 (40)

Firstly, the methylation status of *sFRP1* gene promoter was analyzed. Where in the bisulphite converted DNA of the 100 lung cancer patients was subjected to hot start MS-PCR amplification. The amplified product for *sFRP1* unmethylated primer is of 135 bp and methylated is 126 bp. The presence and absence of bands in the methylated and unmethylated set of primers commences the status of the promoter region. Out of 100 subjects, 76% of the subjects carried unmethylated promoter and hence promoter region of *sFRP1* gene is not hyper methylated in lung tumors. The product size for *DKK3* (M) is 125bp and for *DKK3* (UM) is 120 bp. The percentage of samples showing hyper-methylation were 45% out of them. The *DKK2* promoter methylation was lower in case of 10% of the samples. *DKK2* is an unexplored gene which has been recently shown some interesting involvements in the process of carcinogenesis. The last gene whose promoter region was analyzed is *Axin2*. The percenatge of the *Axin2* gene hypermethylated promoter were 25% out of the total. The above data is very preliminary which shows a trend for methylation status but, certain reliable techniques are required for the analysis in order to draw a substantial conclusion about the role of epigenetic aberrations within these genes in promoting the pathogenesis of lung cancer. Further studies with good techniques such as bisulphite sequencing or RT-PCR can be used efficiently to obtain sustainable results.

5.9.1 Relationship of various clinic-pathological features with methylation profile of Wnt antagonists genes

Table 5.38 shows the effect of methylation status of promoter region of *sFRP1*, *DKK3*, *Axin2* and *DKK2* gene on the various clinic-pathological features such as tumor stage, primary tumor extension, lymph node

invasion and distant metastasis. The patients were stratified according to the clinical stage into stage III and stage IV patients and also on the basis of tumor extension into TX+T1+T2 and T3+T4 and the data has been summarized in table 5.38(a).

In order to investigate the relationship of methylation patterns of promoter region of these genes, the frequencies of patients having methylated and unmethylated promoter were calculated for each gene for both stage III and stage IV patients. The comparison in the frequencies was carried out to find out whether any significant difference lies in the frequencies among the two subgroups made on the basis of stage. Overall, none of the gene depicted any significant difference in the frequencies among the two subgroups pertaining to the lack of evident association between the methylation status of these genes and stage of the tumor. This clearly evident by the proportion of patients showing methylated promoter region in case of all the four genes. The frequencies of the patients with methylated promoter region did not show any significant difference among the stage III and stage IV patients in case of *sFRP1* (25.5% vs 22.5%), *DKK3* (41.8% vs 47.5%), *Axin2* (21.8% vs 32.5%) and *DKK2* (9.1% vs 12.5%) respectively.

For the next clinical parameter which is primary tumor extension as mentioned above, the patients were divided into two categories on the basis of tumor size. A higher proportion of lung cancer patients were reported in the T3+T4 category (69) whereas TX+T1+T2 subgroup consists of only 24 patients. No correlation was observed between the methylation status of any of the Wnt antagonists and primary tumor extension. The difference in the frequencies of the patients having methylated promoter for *sFRP1* (25% vs 26.1%), *DKK3* (45.8% vs 47.8%), *Axin2* (16.7% vs 30.4%) and *DKK2* (8.3% vs 11.6%) was found to be statistically significant among TX+T1+T2 and T3+T4 category respectively. Although, a trend of relationship was seen in case of *Axin2* (OR= 1.98, 95% C.I.: 0.57-6.50, $p=0.28$) and *DKK2* (OR=2.13, 95% C.I.: 0.37-12.08, $p=0.38$) gene as moderately increased odds ratio was observed but it was found to be statistically insignificant. For lymph node invasion, patients showing lack of lymph node involvement were grouped in N0 group and others showing positive lymph node invasion were grouped under N1+N2+N3 category. Similarly, for metastatic status the patients showing positive metastasis were included in group M1 and those exhibiting lack of metastasis were grouped in M0 category and data has been tabulated in table 5.38 (b). This data evaluated the relationship of methylation profile of the promoter region of these genes with lymph node invasion and distant metastasis. None of the genes showed any positive correlation between the methylation status and lymph node invasion for lung cancer patients. The percentage of the patients having methylated primers was not significantly among N0 and N1+N2+N3 category in case of *sFRP1* (18.2% vs 26.8%), *DKK3* (54.6% vs 46.4%) and *Axin2* (27.3% vs 26.8%) respectively. Similarly, for the clinical parameter metastasis it was observed that methylation profile of none of the genes was influencing the metastatic status of the lung cancer patients. No significant difference was estimated among the M0 and M1 subgroup in case of *sFRP1* (28.6% vs 21.6%), *DKK3* (46.4% vs 48.7%), *Axin2* (23.2% vs 32.4%) and *DKK2* (10.7% vs 10.8%).

Table 5.38 (a): Relationship of various clinic-pathological features with methylation profile of Wnt antagonist genes (stage and primary tumor extension)

Gene	Clinical Stage		^a AOR (95% CI) ^a	p^b	Primary Tumor Extension		^a AOR (95% CI) ^a	p^b
	III	IV			Tx+T1+T2	T3+T4		
<i>sFRP1</i>	55	40			24	69		
UM	41 (74.5)	31 (77.5)	Reference		18 (75)	51 (73.9)	Reference	
M	14 (25.5)	9 (22.5)	0.82 (0.30-2.19)	0.69	6 (25)	18 (26.1)	1.12 (0.37-3.43)	0.83
<i>DKK3</i>	55	40			24	69		
UM	32 (58.2)	21 (52.5)	Reference		13 (54.2)	36 (52.2)	Reference	
M	23 (41.8)	19 (47.5)	1.24 (0.54-2.84)	0.60	11 (45.8)	33 (47.8)	1.13 (0.43-2.95)	0.80
<i>Axin2</i>	55	40			24	69		
UM	43 (78.2)	27 (67.5)	Reference		20 (83.3)	48 (69.6)	Reference	
M	12 (21.8)	13 (32.5)	1.78 (0.69-4.53)	0.22	4 (16.7)	21 (30.4)	1.98 (0.57-6.50)	0.28
<i>DKK2</i>	55	40			24	69		
UM	50 (90.9)	35 (87.5)	Reference		22 (91.7)	61 (88.4)	Reference	
M	5 (9.1)	5 (12.5)	1.38 (0.34-5.57)	0.64	2 (8.3)	8 (11.6)	2.13 (0.37-12.08)	0.38

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding p^b -values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

Table 5.38(b): Relationship of various clinic-pathological features with methylation profile of Wnt antagonist genes (lymph node invasion and metastasis)

Gene	Lymph node Invasion		^a AOR (95% CI) ^a	p^b	Metastasis		^a AOR (95% CI) ^a	p^b
	N0	N1+N2+N3			No	Yes		
<i>sFRP1</i>	11	82			56	37		
UM	9 (81.8)	60 (73.2)	Reference		40 (71.4)	29 (78.4)	Reference	
M	2 (18.2)	22 (26.8)	1.20 (0.22-6.42)	0.83	16 (28.6)	8 (21.6)	0.72 (0.26-1.96)	0.52
<i>DKK3</i>	11	82			56	37		
UM	5 (45.4)	44 (53.6)	Reference		30 (53.6)	19 (51.3)	Reference	
M	6 (54.6)	38 (46.4)	0.51 (0.12-2.05)	0.34	26 (46.4)	18 (48.7)	1.14 (0.49-2.66)	0.75
<i>Axin2</i>	11	82			56	37		
UM	8 (72.7)	60 (73.2)	Reference		43 (76.8)	25 (67.6)	Reference	
M	3 (27.3)	22 (26.8)	0.85 (0.19-3.64)	0.82	13 (23.2)	12 (32.4)	1.62 (0.63-4.12)	0.31
<i>DKK2</i>	11	82			56	37		
UM	11 (100)	72 (87.8)	Reference		50 (89.3)	33 (89.2)	Reference	
M	--	10 (12.2)	----		6 (10.7)	4 (10.8)	1.13 (0.27-4.66)	0.86

^aAdjusted Odds ratios, 95% confidence intervals and their corresponding p^b -values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking

CHAPTER 6

DISCUSSION

Pathway based approach is increasingly becoming a critical way to understand the process of disease initiation and progression as it is evident that it involves multitude of signalling networks and targeting a single marker will not enable us to predict disease susceptibility and prognosis along with treatment efficacy having high accuracy [Szkandera *et al.*, 2014]. Wnt signalling cascade is a highly complex and diverse pathway which is implicated in various biological processes such as embryonic development, cell proliferation, cell adhesion, cell polarity, tumorigenesis and apoptosis [Veeck and Dahl, 2012]. Targeting the highly complex intermeshing pathways such as Wnt signalling cascade can help us to understand the pathogenesis of lung cancer as it is a highly heterogeneous and complex genetic disorder. Abnormalities in the Wnt signalling cascade have been recently linked to the process of oncogenesis in various human cancers [Mazieres *et al.*, 2005]. Aberrant Wnt signalling cascade has been recently reported to have an important implication in lung cancer and its progression [Mazieres *et al.*, 2005]. The diverse nature of lung cancer at molecular and genetic level drives the need for drawing certain new hypothesis and evaluating the role of abnormal Wnt signalling in the etiology of this disease [Kikuchi, 1999].

Wnt signalling cascade is regulated by the secreted glycoproteins DKKs which antagonise the signal cascade by binding the LRP co-receptor, in doing so it disrupts the interactive complex of Wnt ligand/Frizzled receptor/LRP. These interactions may affect the regulation of the Wnt cascade in lung tumorigenesis [Kawano and Kypta, 2003]. Axin2 forms a major component of multi protein complex comprising Dsh, APC, and GSK-3 β which regulates the level of β -catenin in the cytoplasm. β -catenin has the ability to interact with the transcription factors such as TCF/LEF and the augment the process of transcription of downstream targets of Wnt pathway such as *matrix metalloproteins*, *cox-2*, *cyclin D1*, *VEGF*, *c-myc* etc. [Stewart, 2014]. The variations in the *Axin2* gene may cause certain modulations in the protein structure and as it forms the scaffold of the destruction complex, this may affect its binding with APC and alter the process of addition of phosphate moiety to β -catenin which is the primary step for its ubiquitination [Davoodi *et al.*, 2015]. The loss of heterozygosity in the region where genomic locus of *Axin2* lies has been reported in various cancers including ovarian, hepatocellular and colorectal [Phelan *et al.*, 1998]. There have been reports in ovarian and colon cancer which have cited that sequence variations in the functional domains of *Axin2* gene, may affect its interaction with β -catenin and APC [Rubinfeld *et al.*, 1997]. The experiments involving the truncation of APC where the *Axin2* binding sites have been removed have reported its role in different human cancers [Ikeda *et al.*, 1998]. There are various domains present in *Axin2* protein and the APC binding domain has significant homology with regulators of G-protein signalling (RGS) hence, it is referred as RGS domain. This pathway is also regulated by secreted-frizzled related

proteins which basically share a structural homology at the N-terminal with the Frizzled- related receptors. This structural homology aids them in antagonising the complex signal cascade by competing with the Frizzled related receptors for binding to the Wnt ligands which transduce the signal cascade [Pohl *et al.*, 2015].

The present study focussed on the various genetic variations in the *DKK* genes, which in turn can alter the interaction between the DKK protein and LRP5/6 thus dysregulating the Wnt signalling and playing a role in modulating the lung cancer susceptibility. Also, the role of seven polymorphic sites which includes five variants located in different exons of the *Axin2* gene namely, Pro⁵⁰Ser in exon1, Ile¹⁴⁴Ile in exon 1, Pro⁴⁵⁵Pro in exon5, Pro⁴⁶²Pro in exon 5 and Leu⁶⁸⁸Leu in exon 7 and two other intronic variants *Axin2* 1712+19, 956+16 has also been evaluated in association towards lung cancer predisposition. Along with this, an attempt has also been made to evaluate the influence of the germline genetic polymorphisms in the *sFRP* genes namely *sFRP3* (rs288326 and rs7775) and *sFRP4* (rs1802073 and rs1802074) on individual's predisposition towards lung cancer. Also, genetic differences within the Wnt antagonist genes such as *DKK*, *sFRP* and *Axin2* have been examined to find their role in predicting the survival and prognosis of North Indian lung cancer patients and their correlation with various clinico-pathological features. This is probably the first attempt to explore the relevance of these polymorphic sites to act as an independent predictor of survival. To our knowledge this is the first study to utilize such an approach to link SNPs in Wnt pathway genes with predisposition to lung cancer in North Indian population.

6.1 Wnt antagonist (*DKK*) polymorphism and lung cancer predisposition

When effect of *DKK4* (rs2073664) polymorphisms on lung cancer was investigated in this study, we detected that *DKK4* (rs2073664) polymorphism was not associated with lung cancer development and thus exhibited a protective effect. Our data is consistent with a study done in Turkish lung cancer patients [Yilmaz *et al.*, 2015] and renal cancer patients of North America [Hirata *et al.*, 2009]. Similarly, another polymorphic site of the *DKK4* gene (rs3763511) was analyzed in breast cancer females of Saudi Arabian and their data also suggested that this SNP in the *DKK4* gene did not show any association with overall breast cancer [Alanazi *et al.*, 2013].

The investigations in the present study showed there was no significant difference in the frequency of the subjects with variant genotype for both the *DKK3* gene's including rs3206824 (Arg³³⁵Gly) and *DKK3* rs2291599 (intronic variation). However, for the *DKK3* rs7396187 (G/C) genotype a protective effect for lung cancer was observed (OR=0.63, ***p*=0.01**) and this protective effect was more pronounced in ADCC subjects (***p*=0.004**). Our results were concordant with the study carried by Kobayashi *et al.*, [2002] where they screened lung cancer patients for codon 335 mutation in the *DKK3* gene rs3206824 (Arg³³⁵Gly) and found no significant difference between cases and controls. Another study carried out in North Indian patients suffering from gallbladder cancer also reported a

lack of any significant association between *DKK3* rs3206824 and cancer risk [Yadav *et al.*, 2016]. On the contrary, a case–control study carried out in Turkish lung cancer and renal carcinoma patients have reported that the *DKK3* rs3206824 (Arg³³⁵Gly) was found to confer a protective effect [Yilmaz *et al.*, 2015; Hirata *et al.*, 2009]. However, these two studies did not find any association for the other *DKK3* variants. Investigations carried in breast cancer females to find out the association between the polymorphic site of *DKK3* gene (rs6485350) and *DKK4* (rs3763511) gene showed a decrease in frequency of the patients with variant genotype for both the polymorphic sites [Alanazi *et al.*, 2013].

Screening of colon cancer patients for *DKK3* rs3206824 and *DKK2* rs17037102 variants did not show any positive correlation [Páez *et al.*, 2014]. Some other reports showed association of various *DKK3* polymorphisms with prostate cancer and factors related to it [Kim *et al.*, 2015]. The functional role of *DKK3* protein as a tumor suppressor has been greatly supported by various reports [Hsieh *et al.*, 2004; Kurose *et al.*, 2004]. Unlike other members of the DKK family, *DKK3* does not have the affinity for the Wnt receptor complex, however its function yet remains unclear [Brott and Sokol, 2002] but recently it has been suggested that *DKK3* is a negative regulator of β -catenin [Lee *et al.*, 2009]. Moreover, lung cancer was the first primary tumor where the dysregulated expression pattern of *DKK3* was asserted [Nozaki *et al.*, 2001]. Recent evidences have shown the role of *DKK3* protein as a key factor in lung cancer pathogenesis [Veeck and Dahl, 2012]. This can be clearly understood by a study which states that *DKK3* protein is antagonising β -catenin driven Wnt pathway in lung cancer cells, which in turn leads to the lowering of expression of Wnt target genes such as *c-myc* and *cyclin D1* specifically in NSCLC cell lines [Veeck and Dahl, 2012]. Also, the variation in levels of *DKK3* m-RNA has been confirmed in case of human renal carcinoma cells [Kurose *et al.*, 2004].

In case of *DKK2* genetic variants, data from the current study showed that individuals carrying one susceptible allele for either Arg¹⁴⁶Glu variant (OR=1.69, ***p*=0.008**) or rs419558 (3-UTR region) (OR=1.77, ***p*=0.002**) were found to have a positive association towards lung cancer. It is evident that genotypic frequency of the control sample set in case of rs447372 variant were inconsistent with HWE. However, these deviations are attributable to the fact that HWE which is based on various assumptions such as random mating, genetic drift *etc.* which might be violated in this sample set taken from the entire population. The reasons include population stratification, sample size and non-random mating. So, the trends observed during this study for *DKK2* are evidently showing its implication in the process of lung carcinogenesis. Variations in the 3'-UTR region have important implication in transcriptional regulation as it is involved in the stability of the m-RNA and its translation [Alanazi *et al.*, 2013]. *DKK2* Arg¹⁴⁶Glu was also evaluated in Turkish lung cancer patients to find out whether it confers any effect on lung cancer susceptibility, however no significant effect was observed [Yilmaz *et al.*, 2015]. This discrepancy might be attributed to the ethnic differences in the populations which are also influenced by other factors [Alanazi *et al.*, 2013] and also the sample size was very less in that study. A study conducted in renal carcinoma patients found that none of the polymorphic sites of

DKK2 gene which was studied had any effect towards renal cancer development. However, the *DKK2* Arg¹⁴⁶Glu variant was found to have a positive association with the survival of renal cancer patients [Hirata *et al.*, 2009]. Yadav *et al.*, [2016] screened *DKK2* rs17037102 in gallbladder cancer patients of North India and reported a significant increased risk for gallbladder cancer, these results are in line with the findings from the present study where rs17037102 was positively associated with risk of lung cancer.

Out of the twenty one possible combinations among all the *DKK* variants, our study reveals for the first time that combined variant genotypes of *DKK3* rs3206824 and *DKK2* rs419558 showed a 2-fold (OR=2.32, *p*=0.008) increased risk in lung cancer patients. Similar SNP-SNP analysis of the *DKK* variants carried out by Hirata *et al.*, [2009], showed no significant correlation among the different genotypic combination, however they did report a strong association towards decreased risk of renal carcinoma in patients with variant genotype of *DKK3* rs3206824 and *sFRP4* rs1802074 [Hirata *et al.*, 2009]. A very few reports are available on these variants of *DKK2* to compare our findings with. *DKK2* protein has been categorized as a one of the important antagonist of Wnt pathway and binds to LRP/Wnt receptor complex [Mao and Niehrs, 2003], also recent studies have reported about the aberrant activation of Wnt pathway which might play an imminent role in the process of lung tumorigenesis [Mazieres *et al.*, 2005]. Therefore, the involvement of *DKK2* is inevitable in etiology of this disease. In addition our data was in accordance to that of Hirata *et al.*, [2009] and Yilmaz *et al.*, [2015] where both reported that polymorphisms within the *DKK* gene were not related to clinico-pathological features. Since Wnt signalling is a highly diverse and complex pathway with different proteins involved and many functional dimensions are still unexplored, so it may not be possible to correlate clinico-pathological factors of the disease with known functions of the Wnt pathway [Beachy *et al.*, 2004]. Future epidemiological and functional studies are warranted to confirm these findings.

Investigation for the interactive potential of *DKK4*, *DKK3* and *DKK2* genes was also carried out to uncover the cumulative role in modulating lung cancer susceptibility. MDR and CART analysis was carried out for these seven variants of *DKK4*, *DKK3* and *DKK2* gene. MDR approach elaborated the effect of these polymorphic sites. The best model which was found to have the CVC of 10/10, prediction error of 0.45 consisted of two variants of *DKK2* gene, namely rs419558 and rs17037102. Similarly, CART analysis also supported the contribution of *DKK2* rs419558 and *DKK2* rs17037102 in predicting the risk for lung cancer development as genotypic combination *DKK2*_{rs419558} (M) showed the highest risk for lung cancer. CART analysis also showcased different genotypic combinations which have evidently categorized the sample population into high and low risk subgroups. These genotypic combinations also reveal the interactive ability of these various SNP's of *DKK2* and *DKK3* gene in cumulatively affecting lung cancer predisposition.

6.2 Wnt antagonist (*Axin2*) polymorphism and lung cancer predisposition

In the present study, a protective effect was observed in subjects which carried the variant (TT) genotype for *Axin2* 148 C/T ($p=0.01$). Many studies have evidently reported the same findings in case of various cancers. A study carried in Turkish population in lung cancer patients also gave the same insight about this polymorphic site where in the TT genotype conferred a decreased risk of lung cancer [Gunes *et al.*, 2009]. Similar results have also been reported in Chinese [Liu *et al.*, 2014] and Japanese lung cancer patients [Kanzaki *et al.*, 2014]. Also, reports are available in case of other type of cancers including gall bladder [Yadav *et al.*, 2016], breast [Aristizabal-pachon *et al.*, 2015], prostate [Ma *et al.*, 2014], astrocytoma [Gunes *et al.*, 2010] and ovarian [Mostowska *et al.*, 2014]. These suggests that, being a potent tumor suppressor gene the variations in the genetic sequence might compensate on its function of regulating Wnt cascade as it is an indispensable part of the control regime of Wnt signalling [Willert *et al.*, 1999]. Another variant lying in exon 5, 1365 G/A, that has also revealed a protective effect in subjects who were carrying a single copy of the susceptible allele. However, no other study has reported such an association in case of lung cancer, perhaps only one study has evaluated this SNP in lung cancer where no association was reported [Gunes *et al.*, 2009]. The reasons in this difference might pertain to very low sample size and also difference in ethnicity. Although this variant is also studied in case of Turkish astrocytoma patients where no association was observed for this polymorphic site in regard of disease susceptibility. *Axin2* 1386 and 1712+19 showed a certain deviation from HWE. The reasons behind this slight deviation might be population stratification, random mating and sample bias for these two variants. In case of two polymorphic sites namely *Axin2* 956+16 A/G, and 2062 C/T, no effect was observed on lung cancer risk in the present study. However, these two variants are correlated with the tooth agenesis in Brazilian population [Mostowska *et al.*, 2006]. Genetic variant 956+16 A/G lies in intron2 and has been reported to influence the splicing process at exon2, although this interaction is weak [Callahan *et al.*, 2009]. Similarly, 2062 C/T, which is although a synonymous substitution of leucine with leucine at 688 codon has been proposed to effect the splicing process by disrupting the enhancer sequences present just before the splice site in exon7 [Callahan *et al.*, 2009]. In line with our findings, these genetic variants didn't confer any modulation in Turkish lung cancer patients [Gunes *et al.*, 2009].

Furthermore in the present study, lung cancer risk was also evaluated in various subgroups made on the basis of histological subtype. Being a diverse type of carcinoma, these subtypes of lung cancer have been predicted to have different disease etiology and pathogenesis [Sodhi *et al.*, 2015]. In our study, the *Axin2* 148 C/T was found to harbor a protective effect in SQCC patients ($p=0.03$). These results are concordant with those observed in Turkish lung cancer patients [Gunes *et al.*, 2009]. In case of 1365 G/A, the Turkish population didn't exhibit any significant association in any of the subgroups [Gunes *et al.*, 2009], our study also revealed similar results for SQCC patients, but a slight association was observed in SCLC patients. For all other five variants, the result from the study

carried on Turkish lung cancer population suggests that they do not contribute in modulating the risk for the disease particularly in any histological subgroup [Gunes *et al.*, 2009]. Our findings are in line with this study in all the three subgroups.

Smoking has always been considered as one of the major factors contributing towards the risk of developing lung cancer. This factor may interact with other genetic variabilities differently to modulate the individual's predisposition towards the oncogenesis of lung tissue. The current study evaluated the risk associated with these polymorphic sites in smokers and non-smokers separately. A previous study done in Turkey, has reported a lower risk of lung cancer in case of smokers having mutant (TT) genotype of *Axin2* 148 C/T. Our findings are also in accordance to the results in this study. This previous study has examined the risk for all the seven polymorphic sites under study in smokers and non-smokers. No effect has been observed any of the category in case of *Axin2* 432 T/C, 956+16 A/G and 2062 C/T [Gunes *et al.*, 2009]. Our study also revealed no significant correlation in smokers and non-smokers for above mentioned sites

Recent study done in Chinese NSCLC patients has examined Linkage Disequilibrium and risk associated with the different haplotypes of *Axin2* (rs2240308, rs22040307) and *MMP7* (rs4791169) gene [Han *et al.*, 2016]. A similar kind of study was also done in ovarian cancer patients of Poland, where haplotype analysis was done for four different SNP's of *Axin2* gene [Mostowska *et al.*, 2014]. The study done in North Indian gall bladder cancer patients also evaluated the haplotypes and their association with cancer risk for two SNP's namely rs2204038 and rs4791171. *Axin2* T_{rs4791171} A_{rs2240308} showed a decreased risk of gall bladder cancer in females [Yadav *et al.*, 2016]. None of the polymorphic site showed any significant association with any of the clinical parameters.

MDR approach was also applied to find the various non-parametric interactions taking place in between these polymorphic variants. None of the previous studies have analyzed such high order interactions for all the seven polymorphic sites of *Axin2* gene. This analysis revealed the best interaction model (*Axin2* 1365) with maximum CVC (9/10) and minimum prediction error (0.49) along with significant permutation *p*-value. Very few reports are available where, MDR was employed to find out these interactions, one such recent study was done in Chinese population where *Axin2* gene (rs2240308, rs22040307) and *MMP7* (rs4791169) and smoking was taken as a parameter to find out the best interaction model contributing towards lung cancer risk [Han *et al.*, 2016]. Another study also used this exhaustive MDR method in ovarian cancer to find out the interactions between variants of *Axin2*, *β-catenin* and *APC* [Mostowska *et al.*, 2014]. So far, this is the first study where high order regression analysis was carried out to obtain the decision tree and odds associated with combinations of various genotypes using CART analysis. This was the first attempt to look into the SNP-SNP interaction taking place between the *Axin2* genetic variants and their contribution in modulating lung cancer predisposition as *Axin2* is a potent tumor suppressor gene. This was evident

from the results of the present study that *Axin2* 1365 is the major role playing SNP's in the interactive models obtained both from MDR and CART and their association towards lung cancer predisposition.

6.3 Wnt antagonist (*sFRP*) polymorphism and lung cancer predisposition

In the present study the two *sFRP3* genetic variants did not show any pronounced association towards lung cancer susceptibility. To best of our knowledge, there have been no reports regarding the role of these two polymorphic variants in modulating lung cancer susceptibility. However, a study performed on German colorectal patients reported a significantly increased risk of colorectal cancer associated with the variant genotype of *sFRP3* rs7775. But the same study revealed the lack of association of *sFRP3* rs288326 variant genotype with risk of colorectal carcinoma, which is in accordance with results observed from the present study [Shanmugam *et al.*, 2007]. On the contrary, another study on colorectal neoplasia reported a lack of association of *sFRP3* rs7775 and *sFRP3* rs288326 variant genotype with risk for colorectal neoplasia [Berndt *et al.*, 2009]. There have been contradictory reports on the role of these genetic variants in modulating the susceptibility. These polymorphic sites have been extensively studied for their role in osteoarthritis [Loughlin *et al.*, 2004] and osteolysis [Mothersill *et al.*, 1999]. The role of *sFRP3* has also been examined in renal cancer cells [Hirata *et al.*, 2010] and androgen dependent prostate cancer cell lines [Zi *et al.*, 2005]. These studies explained that *sFRP3* expression plays an important role in cell division and metastasis of renal and prostate cancer cells.

The other two genetic variants that were evaluated in the present study included the *sFRP4* genetic variants located in exon 6 namely *sFRP4* rs1802073 C/A and *sFRP4* rs1802074 G/A. Being characterized as a tumor suppressor gene, *sFRP4* has been investigated for its altered expression levels in various tumors such as endometrial, prostate, ovarian, cervical, bladder, colorectal, mesothelium, oesophageal etc. [Pohl *et al.*, 2015]. *sFRP4* rs1802073 has been studied in colorectal cancer where a significant association was observed between the variant genotype and increased risk of colorectal cancer where it was inferred that this genetic variant leads to dysregulation of Wnt signaling pathway which results in the altered expression of the downstream gene targets that are involved in cell proliferation and apoptosis [Frank *et al.*, 2010]. Our results also exhibited an increased risk of lung cancer in association with the variant genotype of *sFRP4* rs1802073. This gives us an insight that this non-synonymous SNP which is being classified as “probably damaging” by the Polyphen software might have an influence on the functional properties of the *sFRP4* protein. This SNP has also been evaluated to determine its association with renal carcinoma; however Hirata *et al.*, [2009] observed no significant association regarding *sFRP4* variants and risk of renal cancer. There have been very few studies focussing on the modulation of susceptibility towards lung cancer due to *sFRP4* polymorphisms. However, a study done in gallbladder cancer patients of north India also revealed a positive significant association of *sFRP4* rs1802073 with cancer risk. Also, the substitution

taking place in this polymorphism is reported to be involved in transcriptional regulation of the *sFRP4* gene according to the “Golden Path” prediction tool used in this study [Yadav *et al.*, 2016]. These two polymorphic variants have also been studied for their role in bone mineralization [Boudin *et al.*, 2012] and *sFRP4* rs1802073 has been evaluated for its role in craniotubular hyperostosis [Boudin *et al.*, 2013]. *sFRP4* rs1802074 G/A is least studied polymorphic site, there are not many reports available to compare the results obtained from the present study. The findings from the study done on Turkish population state that rs1802074 GG and GA genotype confers a decreased risk of lung cancer when compared with AA genotype [Yilmaz *et al.*, 2015]. In the present study, this polymorphic site GA+AA genotype showed a marginal role in modifying the susceptibility towards lung cancer, however the stratified analysis did reveal certain high risk subgroups.

Due to variation in the pathogenesis of different histological subtypes of lung cancer, they exhibit slightly different clinical manifestations; also they are mediated through distinct alterations in expression of various genes and proteins. Therefore, it can be concluded that an individual's predisposition towards a certain type of lung cancer might be influenced by different genetic variants. In the present study, *sFRP3* rs288326 did show a major risk associated with ADCC. Experiments done in murine models have confirmed that Wnt signalling pathway plays a major role in NSCLC pathogenesis [Stewart, 2014]. Evidences are also available which clearly demonstrate that aberrant activation of Wnt signalling pathway might lead in lung adenocarcinoma progression [Coscio *et al.*, 2014]. Further, due to less sample size in the stratified categories, the findings from the present study do need further validation. In order to figure the role of smoking which is a major risk factor for lung cancer, we analysed the prevalence of the variant genotypes in smokers and non-smokers, as it is important to detect the gene-environment interactions in order to demarcate the subgroups who are at higher risk of acquiring lung cancer due to environmental factors [Li *et al.*, 2010]. In the present study, non-smokers were found to be at a higher risk as compared to smokers if they had a variant genotype for a polymorphic variant. In case of *sFRP3* rs288326, non-smoker subjects were found to be at a higher risk of developing lung cancer. The molecular mechanism for development of adenocarcinoma which is the only subtype of lung cancer that is weakly related to smoking and the process of carcinogenesis in non-smokers both are highly unclear [Li *et al.*, 2010]. The findings from the present study can be an aid in predicting the molecular changes occurring during the process of tumorigenesis. Non-parametric methods such as MDR were also employed to discover the interaction of the four SNP's of *sFRP3* and *sFRP4* gene. This analysis unraveled the lack of any significant interactive potential of these variants to predict lung cancer risk. The best model obtained from this analysis comprised of single variant namely, *sFRP4* rs1802073 with CVC of 10/10, prediction error=0.43 and $p=0.15$. Also, from the findings of CART analysis it was observed that rs1802073 came out to be most significant determinant of lung cancer risk out of all the SNP's in all the patients. However, a detailed analysis of smokers and non-smokers revealed a lack of association of any

genotypic combination with lung cancer risk in smokers. Although, in non-smokers, some genotypic combination such as *sFRP3*_{rs288326} (M)/ *sFRP4*_{rs1802074} (M)/ *sFRP4*_{rs1802073} (M) was found to confer a high risk in these subjects (OR=10.25, *p*=0.01).

6.4 Gene-gene and gene-environment interactions

Increasingly growing knowledge about Wnt field is giving us new insights about its broad and effective spectrum in cancer therapeutics. The extend of involvement of Wnt pathway cannot be entangled to a single dimension as it has highly complex and intermeshing network which controls numerous cellular functions by impeding a central role in the process of carcinogenesis [Luo *et al.*, 2007]. The present study has made an attempt to unfold the hidden interactive contributions conferred by the panel of genetic variants of these two pathways in lung cancer predisposition. This multigene approach can help in overcoming the shortcomings of traditional single gene studies where the predictive power of the study is restricted. The pathway based study can help in providing a comprehensive conclusion about the associations of several polymorphic sites with cancer risk and prognosis. In the current study, a systematic approach was used to interpret the combined and interactive ability of several SNPs within Wnt antagonist genes to predict lung cancer risk in North Indian population. This is probably the first attempt to evaluate the collaborative effect of the sequence variations among Wnt antagonists on lung cancer risk and the findings are new and have not been found in large-scale investigations. This is a highly unexplored arena which can adds a significant dimension to the lung cancer predisposition.

The double combinations analysed in this study to find their cumulative effect in lung cancer predisposition revealed that lung cancer patients having variant or combined variant genotype for these combinations namely: *sFRP4* rs1802073 C/A and *DKK2* rs419558 C/T, *sFRP4* rs1802074 G/A and *DKK2* rs419558 C/T, *sFRP4* rs1802074 G/A and *DKK2* rs17037102 G/A, *sFRP3* rs7775 C/G and *DKK3* rs3206824 G/A, *sFRP4* rs1802073 C/A and *DKK2* rs447372 A/G, *sFRP3* rs7775 C/G and *Axin2* 148 C/T, *sFRP3* rs7775 C/G and *DKK3* rs7396187 G/C can be classified as high risk subgroups. Certain combinations such as *sFRP3* rs7775 C/G and *Axin2* 1365 G/A, *sFRP3* rs7775 C/G and *Axin2* 956+16 A/G were found to be categorized as low risk subgroups. There have been very few studies which have assessed the cumulative power of these SNP's together. However, the Turkish lung cancer patients having GG of *DKK3* rs3206824 G/A and AG+GG of *sFRP4* rs1802074 G/A were classified as a low risk subgroup (OR=0.12, *p*=0.004). A similar finding was also reported by Hirata et al, in renal cancer patients where GA+AA genotype for *DKK3* rs3206824 G/A and AG+GG of *sFRP4* rs1802074 G/A showed a lower risk for renal cancer with OR=0.19, *p*<0.0001. No other reports are available to compare our findings with.

Several sequence variations of the Wnt antagonist genes have been evaluated for influencing the cancer risk in different ethnicities. Two widely studied polymorphic sites of *sFRP3* gene namely,

rs7775 and rs288326 have been studied in German colorectal cancer patients where only *sFRP3* rs7775 was found to confer an increased risk [Shanmugam *et al.*, 2007]. Another report demonstrated a lack of association of these two variants with colorectal neoplasia [Berndt *et al.*, 2009]. Polymorphic variant of *sFRP4* rs1802073 was found to be positively correlated with increased risk for rectal cancer and early stage colorectal cancer in German population [Frank *et al.*, 2010]. The findings of a study done in Turkish lung cancer patients denied the functional significance of rs1802074 in influencing the susceptibility [Yilmaz *et al.*, 2015]. Another study has investigated the role of *DKK2* (rs17037102), *DKK3* (rs3206824), *DKK3* intron4 G/C (rs7396187), *DKK4* (rs2073664), and *sFRP4* (rs1802074) variants in Turkish lung cancer subjects and none of the sequence variation conferred a significant effect on lung cancer risk. Only the gene combination namely: *DKK3* rs3206824 and *sFRP4* rs1802074 conferred a significantly decreased risk of lung cancer [Yilmaz *et al.*, 2015]. *Axin2* variants have been investigated in Turkish lung cancer patients where the rs2204038 TT genotype conferred a decreased risk of lung cancer [Gunes *et al.*, 2009]. Similar findings have also been observed in Chinese [Liu *et al.*, 2014] and Japanese LC patients [Kanzaki *et al.*, 2006]. Reports are available in case of other type of cancers including breast [Aristizabal-pachon *et al.*, 2015], prostate [Ma *et al.*, 2014], astrocytoma [Gunes *et al.*, 2010] and ovarian [Mostowska *et al.*, 2014]. Previously some studies have taken into account a comprehensive panel of Wnt signalling pathway sequence variations and carried out the association analysis using traditional logistic regression method. A study done in renal cancer patients has analyzed the several variants of *DKK2* (rs447372, rs419558 and rs17037102), *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK4* (rs2073664), *sFRP4* (rs1802073 and rs1802074), *DAAM2* (rs6937133 and rs2504106) and *SMAD7* (rs12953717). Although *sFRP4* rs1802073 was not found to be correlated to renal cancer risk but being a non-synonymous substitution (Pro³²⁰Thr), it has been characterized by Polyphen software as a “probably damaging” change which can alter the functionalities of this extracellular glycoprotein [Hirata *et al.*, 2009]. In our study, *sFRP4* rs1802073 came out to be a significant contributor and was found to be a major component of the combinations and interaction models conferring to increasing the risk for lung cancer. *DKK2* rs17037102 is found to be associated with survival of renal cancer patients [Hirata *et al.*, 2009] on the contrary in our study it came out to be a significant predictor of LC susceptibility both in MDR and CART analysis. Another study which evaluated polymorphic sites of Wnt cascade was done in NSCLC patients to find out the relapse free survival after being treated for early stage NSCLC [Cosico *et al.*, 2014]. A report on colon cancer investigated a range of SNPs of *DKK2*, *DKK3*, *sFRP4* and *Axin2* genes in predicting their clinical influence in tumor recurrence [Paez *et al.*, 2014]. Recent findings from a breast cancer study done in Saudi Arabia also significantly established the role of germline variations of Wnt signalling pathway where 15 different SNPs with in *CTNNB1*, *DKK2*, *DKK3*, *sFRP3*, *sFRP4* and *Axin2* gene of Wnt cascade were examined to find their association with breast cancer risk [Alanazi *et al.*, 2013].

The above mentioned studies have drawn the conclusions on the basis of individual gene association analysis, till now there are no reports demonstrating the collaborative effect of Wnt antagonists to modulate lung cancer susceptibility. However, a few studies have looked into the interactive potential of several genes in various cancers such as bladder cancer [Chen *et al.*, 2007; Wu *et al.*, 2006], gall bladder cancer [Srivastava *et al.*, 2012], gastric adenocarcinoma [Liu *et al.*, 2010] and tobacco related multiple primary neoplasms [Kotnis *et al.*, 2012] *etc.* These studies have employed the high order statistics which has high power in forecasting the multiple-faceted role of these SNPs along with their interaction with the environmental parameters such as smoking. MDR and CART tools were used and some interesting conclusions about the contributions of all the germline variations under study became evident. These non-parametric methods have gained much more significance recently as they give us a clear picture about the various factors which might play an inevitable role in impeding cancer predisposition. Also, unfold the cumulative role of low penetrance alleles in essentially in anticipating the relevance of their interactions in predicting the disease occurrence [Kotnis *et al.*, 2012].

Interesting findings of the present study comprised the evident role of double loci genotypic combinations where *sFRP4* rs1802073 and *DKK2* variants (rs17037102 and rs419558) along with *sFRP4* rs1802074 and *DKK3* rs3206824, both of these contribute towards increasing the risk for lung cancer. This is the first study which investigated the germline variations of a paramount signalling cascade and their accumulative effect on lung cancer risk. Interaction models within Wnt pathway showed the evident partnership of *sFRP4* rs1802073 (non-synonymous substitution) with *DKK2* rs419558 which is located in the 3'-UTR region. A study conducted in Chinese population has observed that the *Axin2* gene polymorphic site (rs2240308, rs22040307) and *MMP7* (rs4791169) and smoking was taken as a parameter to find out the best interaction model contributing towards lung cancer risk [Han *et al.*, 2016]. Another study also used this exhaustive MDR method in ovarian cancer to find out the interactions between variants of *Axin2*, *β-catenin* and *APC* [Mostowska *et al.*, 2014]. Yadav *et al.*, [2016] also employed MDR tool recently to find the best interaction model to predict gall bladder cancer predisposition in north Indians on different SNP's of Wnt pathway including *sFRP4* rs1802073, *DKK2* rs17037102, *DKK3* rs3206824, *APC* rs4595552, *APC* rs11954856, *Axin2* rs4791171, *Axin2* rs2204038, *β-catenin* rs4135385 and *GLI-1* rs222826.

Considering the CART output in case of Wnt antagonist, *DKK2* came out to be the most significant factor in determining the lung cancer risk. Traditional MDR is a rigorously used method to detect these non- linear interactions, especially in the present case where these SNP's are being analyzed for the first time in respect for lung cancer. The interaction of these SNP's has never analyzed before. The repetition of the single SNP in one or more models indicate the strong contribution of this SNP in the lung cancer susceptibility paving a way to investigate further all the SNP's with larger sample sizes in different populations. Such an analysis is the foundation of the further studies in this direction and

also further studies will validate our findings with the help of advanced data mining techniques. However, smaller sample size and lack of functional studies is a limitation to this study prospect. Also, lesser number of subjects were found in the high risk subgroups but then future endeavours in this regard can give a new dimension about lung carcinogenesis. The therapeutic potential of these pathways can also be explored on the basis of the findings of the current study. Further studies with larger sample size and more accurate statistical methods are required to warranty the above findings.

6.5 Wnt antagonists and their role in overall survival and objective response of lung cancer patients receiving doublet platinum based chemotherapy

Wnt cascade has been increasingly gaining importance due its influence on myriad of cellular and biological processes. This evolutionary conserved pathway which is has an established role in embryonic development is also critically involved in maintaining both normal cells and cancer stem cells which have an indispensable ability of self-renewability and differentiation [Willert *et al.*, 2006]. Constitutive activation of this ligand induced pathway is found to be associated with different human cancers such as breast, prostate, colorectal, hepatocellular, melanoma etc [Alanazi *et al.*, 2013]. The aberrant expression of Wnt cascade has been implicated in lung tumors due to dysregulation of various Wnt inhibitors. *DKK*, an antagonist of Wnt signaling, is expressed in NSCLC cell lines [Nozaki *et al.*, 2001], while sFRPs were found to be downregulated [Tang *et al.*, 2006]. Evidences supporting the clinical significance of Wnt pathway in predicting the prognosis of NSCLC patients are very few. These reports used immunostaining techniques for analysing the level of β -catenin in NSCLC tissues which depicts the activation of Wnt pathway. However, the inferences about the prognostic implications of this signalling pathway have been conflicting [Shapiro *et al.*, 2013].

Findings from the current study suggested a lack of any significant association of the variants of *DKK4*, *DKK3*, *DKK2*, *sFRP3*, *sFRP4* and *Axin2* gene with overall survival of lung cancer patients. However, certain trends were observed in case of *DKK4* rs2073664 where in the heterozygous CT genotype showed a slightly better survival as compared to wild (CC) type. In case of *Axin2* 1712+19, the subjects having GG genotype showed a trend of poor prognosis as compared to those with TT genotype.

Lung cancer, is a complex and intricate solid tumor, so different histological subtypes possess diverse pathological pathways for initiation and progression. So, an attempt was made to find the influence of these variants on the prognosis of patients having a particular histological subtype. When patients were stratified on the basis of their histological subtypes several interesting clinical outcomes were obtained. *DKK4* rs2073664 emerged as an independent predictor of survival in case of ADCC patients. This variant was prominently expressed as a good prognostic tool for heterozygous (CT) subjects and on the contrary, the death rate was three fold higher for patients with variant genotype (TT) (HR'=3.30, 95% C.I.: 1.02-10.60, $p=0.04$). *DKK4* is an established inhibitor of Wnt signalling

which is ligand induced, therefore it can be considered for having antitumor potential. Antitumor properties of DKK4 pertain to its ability of activating glycogen synthase kinase- β (GSK-3 β) and the subsequent degradation of cytoplasmic β –catenin [Neihrs, 2006]. The polymorphic site rs2073664 lies in exon 4 of the *DKK4* gene and results in synonymous valine substitution at 169th position. However, the functional relevance of this sequence variation is yet unclear [Yilmaz *et al.*, 2015]. Very few reports are available in regard of this SNP where its effect on prognosis of cancer patients was investigated. A study done in renal cancer patients negated the role of this polymorphic site in predicting their survival after radical nephrectomy [Hirata *et al.*, 2009].

Amongst the three genetic variations explored within the *DKK3* gene, only rs3206824 showed an increased death risk in SCLC patients after adjusting for various covariates. This outcome was not in line with the study done in renal cell carcinoma patients undergoing radical nephrectomy, where all the three SNPs did not contribute in predicting the prognosis of the patients [Hirata *et al.*, 2009]. DKK3 protein has been elucidated as key player in tumorigenesis process and is indicated to possess tumor suppressive abilities. This property of inhibiting the cancer cell growth has been confirmed in various cancer cell line models such as lung, hepatic [Osada *et al.*, 2008; Tsuji *et al.*, 2001], gastric [Yu *et al.*, 2009], glioma [Mizobuchi *et al.*, 2008], prostate [Kawano *et al.*, 2006]. Although, its role as one of the major canonical Wnt regulator is yet to be elucidated. Lung cancer was the first primary tumor where the dysregulated expression pattern of DKK3 was confirmed [Veeck and Dahl, 2012].

Regarding *DKK2* gene, which has a strong affinity for Wnt receptor complex only the variant rs419558 was found to have a good prognostic implication in ADCC patients (HR=0.50, ***p*=0.03**). This polymorphic site lies in the 3'-UTR region which is responsible for stability of mRNA [Alanazi *et al.*, 2013]. In contrast to this finding, Hirata *et al.* suggested that other variant of *DKK2* gene, rs17037102 can acts as a predictor of survival of renal cancer patients, where in the patients harboring the combined variant (GA+AA) had a decreased risk of death as compared to those having GG genotype [Hirata *et al.*, 2009]. One of the interesting finding of the present study obtained after investigating the combined role of the unfavorable genotypes of the three *DKK2* variants is a highly decreased rate of death (HR=0.37, ***p*=0.03**) in North Indian lung cancer patients. No reports are available till date about the cumulative effect of these SNPs on the survival and prognosis of cancer patients. Hence, it is plausible that these single point mutations if investigated in other ethnicities can act as biomarker for lung cancer patients.

Another outcome which was evident from the present study was the association of *sFRP3* rs7775 variant with the survival time of SCLC patients. Subjects having CG genotype had a significantly higher risk of death (adjusted HR=2.03, ***p*=0.05**) in comparison to those with CC genotype. The polymorphic site rs7775 at codon 324 is located at the C-terminal end of *sFRP3* gene which is known extracellular glycoprotein and binds to Wnt ligands in order to negatively regulate the Wnt cascade.

sFRP proteins interact with Wnt ligands through N-terminal frizzled cysteine-rich domain [Frank *et al.*, 2006; Lin *et al.*, 1997]. However, a study has shown that the direct binding between sFRP1 and its ligand Wnt was strongly influenced by heparin, and it should be noted that the C-terminus shows heparin-binding activity. Certain studies have suggested the fact that in order to antagonise the Wnt signal, sFRP protein requires both N- and C-terminal although it is able to bind the Wnt ligand when truncated [Frank *et al.*, 2006; Lin *et al.*, 1997; Chong *et al.*, 2002].

Axin2 is the major scaffold protein of the degradation complex which facilitates the GSK-3 β mediated phosphorylation of β -catenin in order to avoid its internalization into the nucleus and consequently inhibiting the activation of transcription factors TCF/LEF responsible for regulating the expression of downstream target genes of Wnt cascade involved in the process of oncogenesis [Phelan *et al.*, 1998; Dong *et al.*, 2001]. Out of all the seven variants under study, only three variants namely 148 C/T, 1712+19 G/T and 1386 C/T were impeding to have a certain implication as a prognostic marker. The two exonic variants (148 and 1386) showed contrasting results, where the ADCC subjects having mutant (TT) genotype for 148 showed a better prognosis and those with mutant (TT) genotype for *Axin2* 1386 showed a poor prognosis in SCLC subtype. The intronic *Axin2* 1712+19 variant on the other hand, indicated a highly increased death risk in ADCC patients with mutant (GG) genotype.

Prognostic role of Wnt pathway have been investigated in various cancers including breast cancer [Kazi *et al.*, 2016], pancreatic adenocarcinoma [Nakamoto *et al.*, 2014], ovarian cancer [Bodnar *et al.*, 2014], NSCLC [Shapiro *et al.*, 2013] and haematological malignancies [Ashihara *et al.*, 2015]. One study done in Korean population examined the different sequence variations of Wnt-5A, beta- catenin, Wnt16 and Fzd4 *etc.* and reported a lack of any significant association with survival of lung cancer patients [Yoo *et al.*, 2016]. Another study done in early NSCLC patients, have explored around 40 SNPs of twenty genes of Wnt pathway where some of the genetic variations in Wnt16 and beta- catenin showed a significant association with survival of these patients before and after surgical treatment [Cosico *et al.*, 2014]. Recently, another study focussed on the association of different genetic variants of Wnt pathway (*sFRP2* rs4643790, *sFRP4* rs1802073, *DKK2* rs17037102, *DKK3* rs3206824, *APC* rs4595552, *APC* rs11954856, *Axin2* rs4791171, *Axin2* rs2204038, *β -catenin* rs4135385 and *GLI-1* rs222826) with survival and prognosis of gallbladder cancer patients. Their findings did not show any association of *sFRP2* rs4643790, *DKK2* rs17037102, *DKK3* rs3206824, *WISP3* rs1230345, *APC* rs459552, *APC* rs11954856, *Axin2* rs2240308, and *β -catenin* rs4135385 polymorphisms and survival time in any group [Yadav *et al.*, 2016].

An important highlight of the present study is use of high order interaction tools to find out the subgroups showing the best prognosis and higher OS due to the presence of the different genotypic combinations which cannot be inferred by traditional statistical tools. This was probably the first

attempt to analyze the interaction of Wnt antagonists in predicting the survival of lung cancer patients depicted by a survival tree where *DKK4* rs2073664 came out as most significant factor in modulating the survival of lung cancer patients. Node 3 (*sFRP4* rs1802073 (W)/ *DKK2* rs419558 (W)/*DKK4* rs2073664 (M)) exhibited the lowest death rate (HR'=0.04, ***p*=0.008**). Another study done in North Indian gall bladder patients explored the interactive effect of *sFRP2* rs4643790, *sFRP4* rs1802073, *DKK2* rs17037102, *DKK3* rs3206824, *APC* rs4595552, *APC* rs11954856, *Axin2* rs4791171, *Axin2* rs2204038, *β-catenin* rs4135385 and *GLI-1* rs222826 polymorphic sites using recursive partitioning approach to obtain a decision tree to predict the survival. *APC* rs11954856, *GLI-1* rs2228226, *Axin2* rs4791171 were found to play a role in predicting the survival of these patients on the basis of survival tree [Yadav *et al.*, 2016]. Previous studies done in NSCLC patients have also explored the interactive potential of different SNP's in predicting the survival using STREE analysis [Pu *et al.*, 2012; Stewart *et al.*, 2014].

The present study also evaluated the potential of the genetic variants of Wnt antagonist in predicting the objective response of lung cancer patients undergoing doublet platinum based chemotherapy. Recently, efforts have been made in order to find the plausible role of Wnt signalling and its antagonists in altering the response of the patients towards chemotherapy. There is a recent report where the antitumor efficacy of Wnt antagonists in combination with various chemotherapies was evaluated in a large set of patient-derived xenograft models. Wnt antagonists used in this study were vantictumab (anti-FZD) and ipafricept (FZD8-Fc). Sequential dosing of Wnt antagonists, followed by taxane treatment, produced superior antitumor efficacy that was coupled with mitotic cell death in tumors [Fischer *et al.*, 2017]. Therefore, the essential role of Wnt antagonists in influencing the response towards taxol based drugs cannot be denied. Different processes of mitosis such as microtubule dynamics, spindle formation, and centrosome division are found to be regulated by various proteins of Wnt signalling cascade such as adenomatosis polyposis coli (APC), and axis inhibition protein 2 (*Axin2*) [Niehrs and Acebron, 2012]. Some similar indications were evident in the findings of the present study where the genetic differences such as *DKK2* rs419558 C/T, *Axin2* 1365 G/A and *Axin2* 1386 C/T were found to negatively influence the objective response of the patients carrying variant genotype as compared to those carrying wild type genotype for all the three genetic variations. However, *sFRP4* rs1802073 C/A was found to positively influence the response towards chemotherapy. None of the study till date has evaluated such clinical implication of the genetic differences of Wnt signalling pathway.

In conclusion, this study revealed the association of the sequence differences within the Wnt antagonist's gene with the survival of North Indian lung cancer patients. The present findings illuminate the scope of *DKK2* variants to act as prognostic marker in lung cancer. The role of *DKK4* variant rs2073664 is plausible in ADCC progression and initiation as it came across as an independent predictor of survival in this case. However, a major limitation of this study was small sample size in

stratified categories but these findings guide a new track for uncovering the host differences playing role in progression of lung cancer which has immense metastatic potential. Further, studies are needed to validate these facts and pave a way for exploring the Wnt antagonists which have a potential to act as a prognostic biomarker.

6.6 Methylation profile of Wnt antagonist genes (*DKK2*, *DKK3*, *Axin2* and *sFRP1*) and its relationship with various clinic pathological parameters

In the age of advancing diagnostic techniques, the non-invasive methods for cancer screening are gaining much importance. Screening of epigenetic alterations in various tumor suppressor and oncogenes via cell free DNA present in the blood plasma has been of great importance in research. Cell free DNA is the circulating extracellular DNA found in either plasma or serum. It is now a days becoming an important representative of the affected tissue and is widely used for screening biological markers, finding gene mutations, methylation of genes and also analyzing prognostic markers [Jung *et al.*, 2010]. The present study also focused on finding the epigenetic alterations in various Wnt antagonist genes such as *sFRP1*, *DKK3*, *Axin2* and *DKK2* in the plasma DNA samples of North Indian lung cancer patients. This study was conducted to find out the methylation status of the promoter region of these genes and correlate with the clinico-pathological parameters.

Findings from the present study revealed that the rate of methylation of promoters in plasma of lung cancer patients for *sFRP1*, *DKK3*, *Axin2* and *DKK2* was found to be 24%, 45%, 25% and 10% respectively. Further, the association of their methylation profile was when examined with the clinical parameters, no significance correlation was observed for staging, primary tumor extension, lymph node invasion and metastasis. Studies conducted previously for investigating the epigenetic alteration of *sFRP1* gene in different malignancies have shown distinct frequencies of methylation. A study conducted by Zhang *et al.*, [2009] in NSCLC patients, reported that 32% of the tissue samples and 28% of the plasma samples were hyper methylated for sFRP1 promoter. They also assessed the relationship of the methylation status with clinico-pathological features and found that it is associated with lymph node metastasis and disease progression which is contradictory to the findings in the present study where no such association was observed [Zhang *et al.*, 2009]. In another study done on lung cancer patients, it was observed that 55% of the primary lung tumors were hyper methylated for *sFRP1* [Fukii *et al.*, 2005]. There have been other reports where the hyper methylation of *sFRP1* was evident in NSCLC tissues [Zhang *et al.*, 2009; Zhang *et al.*, 2011] and NSCLC cell lines [Fukui *et al.*, 2005; Shi *et al.*, 2007; Licchesi *et al.*, 2008; Taguchi *et al.*, 2016]. Methylation profile of this gene has also been investigated in other types of human cancers. 61% of the breast cancer tissue samples were found to be hyper methylated for *sFRP1* promoter region when compared to normal tissue [Veeck *et al.*, 2006]. Similar trend was observed in hepatocellular carcinoma [Shih *et al.*, 2006], colorectal

cancer [Caldwell *et al.*, 2004, Suzuki *et al.*, 2004], prostate cancer [Lodygin *et al.*, 2005], bladder cancer [Marsit *et al.*, 2005] and ovarian cancer [Takada *et al.*, 2004].

Investigations regarding the epigenetic alterations in *DKK3* gene have been done extensively in the previous studies. Frequency of methylated plasma samples for *DKK3* gene was 45% in the current study. A study conducted in breast cancer patients evaluated the epigenetic status of *DKK3* in tissue samples and 61% of the samples were found to be hyper methylated. Also, this study investigated the possible association of methylation status with clinico-pathological features such as tumor size, lymph node status, histological grade and histological type but none of the parameter showed any significant correlation with methylation status [Veeck *et al.*, 2009]. The epigenetic alteration in *DKK3* gene has been widely reported in many human malignancies including lung tumors [Yue *et al.*, 2008]. Several cancer cell lines such as colorectal, gastric and hepatocellular have been assessed for the epigenetic alterations in *DKK3* promoter region and it has been reported that 56%, 63% and less proportion of cell lines were found to hyper methylated respectively [Sato *et al.*, 2007]. In another report, 67.6% and 52.3% gastric and colorectal tissue samples were found to be epigenetically altered respectively [Yu *et al.*, 2009]. The study done in leukemia patients reported that 32% of the patients showed hyper methylation in *DKK3* promoter region. They also examined the difference between methylated and unmethylated patients in terms of clinical features and found no significant association [Roman-Gomez *et al.*, 2004]. The percentage of methylated *DKK3* promoter in prostate [Lodygin *et al.*, 2005] and papillary thyroid cancer [Yin *et al.*, 2013] tissue samples was found to be 68% and 38.8%. The epigenetic status of *DKK3* has been found altered in various human malignancies [Hayashi *et al.*, 2012] such as renal carcinoma [Urakami *et al.*, 2006], bladder cancer [Urakami *et al.*, 2006], breast cancer [Xiang *et al.*, 2013], hepatocellular cancer [Yang *et al.*, 2010] and cervical cancer [van der Meide *et al.*, 2011].

Axin2 is a potent tumor suppressor gene and its down regulation by methylation of promoter region has been studied [Li *et al.*, 2015]. 112 patients were investigated and out of them 7% tumors reported hyper methylation in *Axin2* gene. *Axin2* methylation level was found not be associated with any of the clinical parameters except gender where 87.5% of high methylation was reported among women [Naghibalhossaini *et al.*, 2012]. A study conducted in NSCLC patients revealed that epigenetic alteration in *Axin2* gene in association with p53 lead to the accumulation of β -catenin and 52% of the tumor samples were found to be methylated for *Axin2* promoter region [Tseng *et al.*, 2008]. The studies done in colorectal cancer patients have established the epigenetic silencing of *Axin2* gene along with microsatellites [Koinuma *et al.*, 2006].

DKK2 is an established Wnt antagonist and its role in regulating the nuclear concentration of β -catenin is indispensable. The epigenetic alteration of *DKK2* was examined in renal carcinoma cell lines and hyper methylation was found to influence the level of *DKK2* expression [Hirata *et al.*, 2009].

A study conducted in ovarian cancer patients reported that *DKK2* was methylated in most of the cases of ovarian cancer in comparison to unmethylated form present in benign and normal tissue of ovary [Zhu *et al.*, 2012]. Also, it was reported that in colorectal cell lines and gastric cell line, 89% and 94% of the them were found to be epigenetically modified for *DKK2* respectively [Sato *et al.*, 2007]. A recent study done in gastric cancer patients also reported that *DKK2* was methylated in these patients and it controls accumulation of β -catenin [Wang *et al.*, 2017]. In a similar report regarding colorectal cancer patients, *DKK2* was found to be methylated in almost all tissue samples [Wang *et al.*, 2017].

CHAPTER 7

CONCLUSION

Salient findings from the present study

1. *DKK4* (rs2073664) did not confer any effect on lung cancer susceptibility
2. Two variants of *DKK3* namely rs3206824 and rs2291599 did not show any association
3. The subjects with heterozygous genotype of *DKK2* rs17037102 and rs419558 were found to have a higher risk of developing lung cancer. The variant genotype of the genotypic combination, *DKK3* rs3206824 and *DKK2* rs419558 showed a two-fold increased risk of developing lung cancer.
4. Three out of seven studied polymorphic sites showed a strong protective effect in subjects having mutant genotype for *Axin2* 148 and heterozygous genotype for 1365 towards lung cancer risk.
5. The patients with the genotypic combination of terminal node (*DKK2*_{rs17037102} (M)/ *DKK2*_{rs419558} (M)) showed nearly a two-fold risk of acquiring lung cancer.
6. It was observed that the unfavorable genotypes of the three *DKK2* variants collectively (rs447372, rs419558 and rs17037102) exhibited a highly decreased rate of death. Survival tree analysis depicted *DKK4* rs2073664 as the major contributor in predicting the survival of the lung cancer patients.
7. The percentage of methylated promoter region for *sFRP1*, *DKK2*, *DKK3* and *Axin2* gene in lung cancer patients was 23%, 10%, 45% and 25% respectively. However, the methylation status was not found to be associated with any of the clinic pathological parameters.

In conclusion, this is the first study to highlight the genetic variations of *DKK3*, *DKK2*, *DKK4*, *Axin2*, *sFRP3* and *sFRP4* gene and their effect on lung cancer risk in North Indians with a larger sample size and also their association with clinico-pathological factors. An attempt was made to find out the high risk subgroups, by stratifying the subjects based on smoking status also. Our study is a primary effort to introduce the concept of polygenic approach to get an insight about the various polymorphic variants in determining cancer susceptibility. These certain interesting findings along with studies in gene expression with larger sample size and stratified based on-set of the disease might help in the establishing the influence of genetic variants in the Wnt signaling pathway genes.

Future prospects of the study

The molecular abnormalities involved in the process of lung carcinogenesis are highly complex and involves the intermeshing cell signalling and proliferation pathways such as Wnt pathway. DKK, sFRP, and Axin2 family of proteins are involved in the regulation of the Wnt signalling cascade. However, the genetic variabilities within the genes coding these proteins may confer a role in altering its functionalities; hence confer a hand in the aberrant activation of Wnt pathway which has been recently known to play a significant role in process of cancer initiation and progression. The study of these single nucleotide variations present in these proteins can greatly help in looking for novel therapeutic modalities, as they can serve as a target. These kind of studies help in giving us the insight about lung carcinogenesis and these tumor suppressor proteins are strong candidature for a prognostic marker also.

BIBLIOGRAPHY

- Abrahamsson, A. E., Geron, I., Gotlib, J., Dao, K. H. T., Barroga, C. F., Newton, I. G., ... & Jones, C. (2009). Glycogen synthase kinase 3 β missplicing contributes to leukemia stem cell generation. *Proceedings of the National Academy of Sciences*, 106(10), 3925-3929.
- Agnihotri, S., Wolf, A., & Guha, A. (2010). Targeting metabolic remodeling in glioblastoma multiforme. *Oncotarget*, 1(7), 567.
- Ai, L., Tao, Q., Zhong, S., Fields, C. R., Kim, W. J., Lee, M. W., ...& Robertson, K. D. (2006). Inactivation of Wnt inhibitory factor-1 (WIF1) expression by epigenetic silencing is a common event in breast cancer. *Carcinogenesis*, 27(7), 1341-1348.
- Akiri, G., Cherian, M. M., Vijayakumar, S., Liu, G., Bafico, A., & Aaronson, S. A. (2009). Wnt pathway aberrations including autocrine Wnt activation occur at high frequency in human non-small-cell lung carcinoma. *Oncogene*, 28(21), 2163-2172.
- Akiyama, T. (2000). Wnt/ β -catenin signaling. *Cytokine & Growth Factor Reviews*, 11(4), 273-282.
- Alanazi, M. S., Parine, N. R., Shaik, J. P., Alabdulkarim, H. A., Ajaj, S. A., & Khan, Z. (2013). Association of single nucleotide polymorphisms in Wnt signaling pathway genes with breast cancer in Saudi patients. *PloS One*, 8(3), e59555.
- Alberg, A. J., Brock, M. V., Ford, J. G., Samet, J. M., & Spivack, S. D. (2013). Epidemiology of lung cancer: Diagnosis and management of lung cancer: American College of Chest Physicians evidence-based clinical practice guidelines. *CHEST Journal*, 143(5_suppl), e1S-e29S
- Amos, C. I., Xu, W., & Spitz, M. R. (1999). Is there a genetic basis for lung cancer susceptibility?. In *Chemoprevention of Cancer* (pp. 3-12). Springer, Berlin, Heidelberg.
- Angers, S., & Moon, R. T. (2009). Proximal events in Wnt signal transduction. *Nature reviews Molecular Cell Biology*, 10(7), 468-477.
- Ankathil, R. (2010). Tobacco, genetic susceptibility and lung cancer. *Tobacco Use Insights*, 3, 1.
- Aristizabal-Pachon, A. F., Carvalho, T. I., Carrara, H. H., Andrade, J., & Takahashi, C. S. (2015). AXIN2 polymorphisms, the beta-catenin destruction complex expression profile and breast cancer susceptibility. *Asian Pacific Journal of Cancer Prevention*, 16(16), 7277-7284.
- Asami, S., Manabe, H., Miyake, J., Tsurudome, Y., Hirano, T., Yamaguchi, R., & Kasai, H. (1997). Cigarette smoking induces an increase in oxidative DNA damage, 8-hydroxydeoxyguanosine, in a central site of the human lung. *Carcinogenesis*, 18(9), 1763-1766.

- Ashihara, E., Takada, T., & Maekawa, T. (2015). Targeting the canonical Wnt/ β -catenin pathway in hematological malignancies. *Cancer Science*, 106(6), 665-671.
- Ashton-Rickardt, P. G., Dunlop, M. G., Nakamura, Y., Morris, R. G., Purdie, C. A., Steel, C. M., ... & Wyllie, A. H. (1989). High frequency of APC loss in sporadic colorectal carcinoma due to breaks clustered in 5q21-22. *Oncogene*, 4(10), 1169-1174.
- Bafico, A., Liu, G., Goldin, L., Harris, V., & Aaronson, S. A. (2004). An autocrine mechanism for constitutive Wnt pathway activation in human cancer cells. *Cancer Cell*, 6(5), 497-506.
- Barker, N., & Clevers, H. (2006). Mining the Wnt pathway for cancer therapeutics. *Nature Reviews Drug Discovery*, 5(12), 997-1014.
- Barker, N., Van Es, J. H., Kuipers, J., Kujala, P., Van Den Born, M., Cozijnsen, M., ... & Clevers, H. (2007). Identification of stem cells in small intestine and colon by marker gene Lgr5. *Nature*, 449(7165), 1003-1007.
- Bartlett, J. M., & Stirling, D. (Eds.). (2003). *PCR Protocols* (Vol. 226). Totowa, NJ: Humana Press.
- Beachy, P. A., Karhadkar, S. S., & Berman, D. M. (2004). Tissue repair and stem cell renewal in carcinogenesis. *Nature*, 432(7015), 324-331.
- Begg, C. B., & Zhang, Z. F. (1994). Statistical analysis of molecular epidemiology studies employing case-series. *Cancer Epidemiology and Prevention Biomarkers*, 3(2), 173-175.
- Behera, D. (2007). New approach to the treatment of lung cancer: the molecular targeted therapy. *Indian Journal of Chest Diseases and Allied Sciences*, 49(3), 149.
- Behera, D., & Balamugesh, T. (2004). Lung cancer in India. *Indian Journal of Chest Diseases and Allied Sciences*, 46, 269-282.
- Behrens, J., Jerchow, B. A., Würtele, M., Grimm, J., Asbrand, C., Wirtz, R., ... & Birchmeier, W. (1998). Functional interaction of an axin homolog, conductin, with β -catenin, APC, and GSK3 β . *Science*, 280(5363), 596-599.
- Beir, V. I. (1999). Health effects of exposure to radon. Committee on health risks of exposure to radon, Board on radiation effects research, Commission on life sciences, National Research Council.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 289-300.

- Berndt, S. I., Huang, W. Y., Yeager, M., Weissfeld, J. L., Chanock, S. J., & Hayes, R. B. (2009). Genetic variants in frizzled-related protein (FRZB) and the risk of colorectal neoplasia. *Cancer Causes & Control*, 20(4), 487-490.
- Bhanot, P., Brink, M., Samos, C. H., Hsieh, J. C., Wang, Y., Macke, J. P., ...& Nusse, R. (1996). A new member of the frizzled family from *Drosophila* functions as a Wingless receptor. *Nature*, 382(6588), 225-230.
- Björklund, P., Åkerström, G., & Westin, G. (2007). An LRP5 receptor with internal deletion in hyperparathyroid tumors with implications for deregulated WNT/ β -catenin signaling. *PLoS Medicine*, 4(11), e328.
- Björklund, P., Svedlund, J., Olsson, A. K., Åkerström, G., & Westin, G. (2009). The internally truncated LRP5 receptor presents a therapeutic target in breast cancer. *PloS One*, 4(1), e4243.
- Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J. R., Ellis, A., Gorman, P., & Sheer, D. (1987). Localization of the gene for familial adenomatous polyposis on chromosome 5. *Nature*, 328(6131), 614-616.
- Bodnar, L., Stanczak, A., Cierniak, S., Smoter, M., Cichowicz, M., Kozlowski, W., ...& Lamparska-Przybysz, M. (2014). Wnt/ β -catenin pathway as a potential prognostic and predictive marker in patients with advanced ovarian cancer. *Journal of Ovarian Research*, 7(1), 16.
- Boffetta, P. A. O. L. O., & Trichopoulos, D. I. M. I. T. R. I. O. S. (2002). Cancer of the lung, larynx, and pleura. *Textbook of Cancer Epidemiology*, 2, 349-67.
- Borczuk, A. C., Gorenstein, L., Walter, K. L., Assaad, A. A., Wang, L., & Powell, C. A. (2003). Non-small-cell lung cancer molecular signatures recapitulate lung developmental pathways. *The American Journal of Pathology*, 163(5), 1949-1960.
- Boudin, E., Piters, E., Fijalkowski, I., Stevenheydens, G., Steenackers, E., Kuismin, O., ...& Van Hul, W. (2013). Mutations in sFRP1 or sFRP4 are not a common cause of craniotubular hyperostosis. *Bone*, 52(1), 292-295.
- Boudin, E., Piters, E., Nielsen, T. L., Andersen, M., Roef, G., Taes, Y., ...& Van Hul, W. (2012). Single nucleotide polymorphisms in sFRP4 are associated with bone and body composition related parameters in Danish but not in Belgian men. *Molecular Genetics and Metabolism*, 106(3), 366-374.
- Brookes, A. J. (1999). The essence of SNPs. *Gene*, 234(2), 177-186.
- Brott, B. K., & Sokol, S. Y. (2002). Regulation of Wnt/LRP signaling by distinct domains of Dickkopf proteins. *Molecular and Cellular Biology*, 22(17), 6100-6110.

- Cabanero, M., Sangha, R., Sheffield, B. S., Sukhai, M., Pakkal, M., Kamel-Reid, S., ... & Tsao, M. S. (2017). Management of EGFR-mutated non-small-cell lung cancer: practical implications from a clinical and pathology perspective. *Current Oncology*, 24(2), 111.
- Cabrera, C. V., Alonso, M. C., Johnston, P., Phillips, R. G., & Lawrence, P. A. (1987). Phenocopies induced with antisense RNA identify the wingless gene. *Cell*, 50(4), 659-663.
- Caldwell, G. M., Jones, C., Gensberg, K., Jan, S., Hardy, R. G., Byrd, P., ...& Morton, D. G. (2004). The Wnt antagonist sFRP1 in colorectal tumorigenesis. *Cancer Research*, 64(3), 883-888.
- Callahan, N., Modesto, A., Meira, R., Seymen, F., Patir, A., & Vieira, A. R. (2009). Axis inhibition protein 2 (AXIN2) polymorphisms and tooth agenesis. *Archives of Oral Biology*, 54(1), 45-49.
- Calles, A., Sholl, L. M., Rodig, S. J., Pelton, A. K., Hornick, J. L., Butaney, M., ... & Jänne, P. A. (2015). Immunohistochemical loss of LKB1 is a biomarker for more aggressive biology in KRAS-mutant lung adenocarcinoma. *Clinical Cancer Research*.doi: 10.1158/1078-0432.CCR-14-3112
- Cancer Genome Atlas Research Network. (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature*, 511(7511), 543.
- Carlton, V. E., Ireland, J. S., Useche, F., & Faham, M. (2006). Functional single nucleotide polymorphism-based association studies. *Human Genomics*, 2(6), 391.
- Carter, B. W., Glisson, B. S., Truong, M. T., & Erasmus, J. J. (2014). Small cell lung carcinoma: staging, imaging, and treatment considerations. *Radiographics*, 34(6), 1707-1721.

CHAPTER 8

- Chen, M., Kamat, A. M., Huang, M., Grossman, H. B., Dinney, C. P., Lerner, S. P., ...& Gu, J. (2007). High-order interactions among genetic polymorphisms in nucleotide excision repair pathway genes and smoking in modulating bladder cancer risk. *Carcinogenesis*, 28(10), 2160-2165.
- Chen, Z., Fillmore, C. M., Hammerman, P. S., Kim, C. F., & Wong, K. K. (2014). Non-small-cell lung cancers: a heterogeneous set of diseases. *Nature Reviews Cancer*, 14(8), 535-546.
- Chia, I. V., & Costantini, F. (2005). Mouse axin and axin2/conductin proteins are functionally equivalent in vivo. *Molecular and Cellular Biology*, 25(11), 4371-4376.

- Chim, C. S., Chan, W. W., Pang, A., & Kwong, Y. L. (2006). Preferential methylation of Wnt inhibitory factor-1 in acute promyelocytic leukemia: an independent poor prognostic factor. *Leukemia*, 20(5), 907-907.
- Choi, H. J., Park, H., Lee, H. W., & Kwon, Y. G. (2012). The Wnt pathway and the roles for its antagonists, DKKS, in angiogenesis. *IUBMB Life*, 64(9), 724-731.
- Chong, J. M., Üren, A., Rubin, J. S., & Speicher, D. W. (2002). Disulfide bond assignments of secreted Frizzled-related protein-1 provide insights about Frizzled homology and netrin modules. *Journal of Biological Chemistry*, 277(7), 5134-5144.
- Christiani, D.C. (2006). Genetic Susceptibility to Lung Cancer. *Journal of Clinical Oncology*, 24(11), 1651–1652.
- Chu, P. G., & Weiss, L. M. (2002). Expression of cytokeratin 5/6 in epithelial neoplasms: an immunohistochemical study of 509 cases. *Modern Pathology*, 15(1), 6-10.
- Clements, W. K., Kim, A. D., Ong, K. G., Moore, J. C., Lawson, N. D., & Traver, D. (2011). A somitic Wnt16/Notch pathway specifies haematopoietic stem cells. *Nature*, 474(7350), 220-224.
- Clevers, H., & Nusse, R. (2012). Wnt/ β -catenin signaling and disease. *Cell*, 149(6), 1192-1205.
- Clevers, H., Loh, K. M., & Nusse, R. (2014). Stem Cell Signaling: An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control. *Science*, 346(6205), 54.
- Coghlin, C., & Murray, G. I. (2010). Current and emerging concepts in tumour metastasis. *The Journal of Pathology*, 222(1), 1-15.
- Colby, T. V. (1995). Tumors of the lower respiratory tract. *Atlas of Tumor Pathology*, 309-317.
- Coscio, A., Chang, D. W., Roth, J. A., Ye, Y., Gu, J., Yang, P., & Wu, X. (2014). Genetic variants of the Wnt signaling pathway as predictors of recurrence and survival in early-stage non-small cell lung cancer patients. *Carcinogenesis*, 35(6), 1284-1291.
- Cuilliere-Dartigues, P., El-Bhiri, J., Krimi, A., Buhard, O., Fontanges, P., Flejou, J. F., ...& Duval, A. (2006). TCF-4 isoforms absent in TCF-4 mutated MSI-H colorectal cancer cells colocalize with nuclear CtBP and repress TCF-4-mediated transcription. *Oncogene*, 25(32), 4441-4448.
- Danesi, R., De Braud, F., Fogli, S., De Pas, T. M., Di Paolo, A., Curigliano, G., & Del Tacca, M. (2003). Pharmacogenetics of anticancer drug sensitivity in non-small cell lung cancer. *Pharmacological Reviews*, 55(1), 57-103.
- Davidson, G., Wu, W., Shen, J., Bilic, J., Fenger, U., Stannek, P., ...& Niehrs, C. (2005). Casein kinase 1 γ couples Wnt receptor activation to cytoplasmic signal transduction. *Nature*, 438(7069), 867-872.

- Davoodi, B., Onsory, K., & Heydari Nasrabadi, M. (2015). Genetic polymorphism of the Axin2 gene in Wnt pathway in patients with ovarian cancer. *Journal of Ardabil University of Medical Sciences*, 15(2), 220-227.
- de Sousa e Melo, F., & Vermeulen, L. (2016). Wnt signaling in cancer stem cell biology. *Cancers*, 8(7), 60.
- Deng, F., Zhou, K., Cui, W., Liu, D., & Ma, Y. (2015). Clinicopathological significance of wnt/ β -catenin signaling pathway in esophageal squamous cell carcinoma. *International Journal of Clinical and Experimental Pathology*, 8(3), 3045.
- Desjardins, P., & Conklin, D. (2010). NanoDrop microvolume quantitation of nucleic acids. *Journal of Visualized Experiments: JoVE*, (45).
- Dong, X., Seelan, R. S., Qian, C., Mai, M., & Liu, W. (2001). Genomic structure, chromosome mapping and expression analysis of the human AXIN2 gene. *Cytogenetic and Genome Research*, 93(1-2), 26-28.
- Eastman, Q., & Grosschedl, R. (1999). Regulation of LEF-1/TCF transcription factors by Wnt and other signals. *Current Opinion in Cell Biology*, 11(2), 233-240.
- Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American Statistical Association*, 83(402), 414-425.
- Eisenhauer, E., Therasse, P., Bogaerts, J., Schwartz, L. H., Sargent, D., Ford, R., ...& Rubinstein, L. (2009). New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *European Journal of Cancer*, 45(2), 228-247.
- Ekström, E. J., Sherwood, V., & Andersson, T. (2011). Methylation and loss of Secreted Frizzled-Related Protein 3 enhances melanoma cell migration and invasion. *PloS One*, 6(4), e18674.
- Fang, X., Gu, P., Zhou, C., Liang, A., Ren, S., Liu, F., ...& Zhang, Z. (2014). β -Catenin overexpression is associated with gefitinib resistance in non-small cell lung cancer cells. *Pulmonary Pharmacology & Therapeutics*, 28(1), 41-48.
- Fearon, E. R. (2011). Molecular genetics of colorectal cancer. *Annual Review of Pathology: Mechanisms of Disease*, 6, 479-507.
- Fiedler, M., Mendoza-Topaz, C., Rutherford, T. J., Mieszczanek, J., & Bienz, M. (2011). Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating β -catenin. *Proceedings of the National Academy of Sciences*, 108(5), 1937-1942.

- Fischer, M. M., Cancilla, B., Yeung, V. P., Cattaruzza, F., Chartier, C., Murriel, C. L., ...& O'Young, G. (2017). WNT antagonists exhibit unique combinatorial antitumor activity with taxanes by potentiating mitotic cell death. *Science Advances*, 3(6), e1700090.
- Fong, K. M., Sekido, Y., Gazdar, A. F., & Minna, J. D. (2003). Lung cancer 9: Molecular biology of lung cancer: clinical implications. *Thorax*, 58(10), 892-900.
- Forgacs, E., Zöchbauer-Müller, S., Oláh, E., & Minna, J. D. (2001). Molecular genetic abnormalities in the pathogenesis of human lung cancer. *Pathology Oncology Research*, 7(1), 6-13.
- Frank, B., Hoffmeister, M., Klopp, N., Illig, T., Chang-Claude, J., & Brenner, H. (2010). Single nucleotide polymorphisms in Wnt signaling and cell death pathway genes and susceptibility to colorectal cancer. *Carcinogenesis*, 31(8), 1381-1386.
- Frank, B., Shanmugam, K. S., Beckmann, L., Hemminki, K., Brenner, H., Hoffmeister, M., ...& Burwinkel, B. (2006). Death receptor 4 variants and colorectal cancer risk. *Cancer Epidemiology and Prevention Biomarkers*, 15(10), 2002-2005.
- Fukui, T., Kondo, M., Ito, G., Maeda, O., Sato, N., Yoshioka, H., ...& Sekido, Y. (2005). Transcriptional silencing of secreted frizzled related protein 1 (SFRP1) by promoter hypermethylation in non-small-cell lung cancer. *Oncogene*, 24(41), 6323-6327.
- Fukuse, T., Hirata, T., Naiki, H., Hitomi, S., & Wada, H. (2000). Prognostic significance of cyclin E overexpression in resected non-small cell lung cancer. *Cancer Research*, 60(2), 242-244.
- Funai, K., Yokose, T., Ishii, G. I., Araki, K., Yoshida, J., Nishimura, M., & Ochiai, A. (2003). Clinicopathologic characteristics of peripheral squamous cell carcinoma of the lung. *The American Journal of Surgical Pathology*, 27(7), 978-984.
- Gadgeel. (2010, August 3). Introduction to Small Cell Lung Cancer: Prevalence, Initial Symptoms, Work-Up, and Staging. Retrieved October 31, 2017, from
- Gandini, S., Botteri, E., Iodice, S., Boniol, M., Lowenfels, A. B., Maisonneuve, P., & Boyle, P. (2008). Tobacco smoking and cancer: A meta-analysis. *International journal of cancer*, 122(1), 155-164.
- Gehart, H., & Clevers, H. (2015). Repairing organs: lessons from intestine and liver. *Trends in Genetics*, 31(6), 344-351.
- Giles, R. H., van Es, J. H., & Clevers, H. (2003). Caught up in a Wnt storm: Wnt signaling in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1653(1), 1-24.

- Glinka, A., Wu, W., Delius, H., Monaghan, A. P., Blumenstock, C., & Niehrs, C. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature*, 391(6665), 357-362.
- Globocan (2012). Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer. GLOBOCAN. 2013; 2012 v1. 0. Available from: <http://globocan.iarc.fr>.
- Goldman, D., & Merril, C. R. (1982). Silver staining of DNA in polyacrylamide gels: linearity and effect of fragment size. *Electrophoresis*, 3(1), 24-26.
- Gong, J., Jiang, Y., Hao, N., Zhu, B., & Li, Y. (2015). Quantitative assessment of the association between AXIN2 rs2240308 polymorphism and cancer risk. *Scientific Reports*, 5, 10111
- Gonzalez, D. M., & Medici, D. (2014). Signaling mechanisms of the epithelial-mesenchymal transition. *Science Signaling*, 7(344), re8.
- Gordon, M. D., & Nusse, R. (2006). Wnt signaling: multiple pathways, multiple receptors, and multiple transcription factors. *Journal of Biological Chemistry*, 281(32), 22429-22433.
- Graham, T. A., Weaver, C., Mao, F., Kimelman, D., & Xu, W. (2000). Crystal structure of a β -catenin/Tcf complex. *Cell*, 103(6), 885-896.
- Graziano, S. L., Gamble, G. P., Newman, N. B., Abbott, L. Z., Rooney, M., Mookherjee, S., & Poiesz, B. J. (1999). Prognostic significance of K-ras codon 12 mutations in patients with resected stage I and II non-small-cell lung cancer. *Journal of Clinical Oncology*, 17(2), 668-668.
- Groden, J., Thliveris, A., Samowitz, W., Carlson, M., Gelbert, L., Albertsen, H., ...& Sargeant, L. (1991). Identification and characterization of the familial adenomatous polyposis coli gene. *Cell*, 66(3), 589-600.
- Gunes, E. G., Pinarbasi, E., & Pinarbasi, H. (2010). AXIN2 polymorphism and its association with astrocytoma in a Turkish population. *Molecular Medicine Reports*, 3(4), 705-709.
- Gunes, E. G., Pinarbasi, E., Pinarbasi, H., & Silig, Y. (2009). Strong association between lung cancer and the AXIN2 polymorphism. *Molecular Medicine Reports*, 2(6), 1029-1035.
- Gupta, U., Mir, S. S., Garg, N., Agarwal, S. K., Pande, S., & Mittal, B. (2016). Association study of inflammatory genes with rheumatic heart disease in North Indian population: A multi-analytical approach. *Immunology letters*, 174, 53-62.
- Han, S., Lv, L., Wang, X., Wang, X., & Zhao, H. (2016). Association of AXIN2 and MMP7 polymorphisms with non-small cell lung cancer in Chinese Han population. *International Journal of Clinical and Experimental Pathology*, 9(2), 2253-2258.
- Hanahan, D., & Weinberg, R. A. (2000). The hallmarks of cancer. *Cell*, 100(1), 57-70.

- Harada, N., Tamai, Y., Ishikawa, T. O., Sauer, B., Takaku, K., Oshima, M., & Taketo, M. M. (1999). Intestinal polyposis in mice with a dominant stable mutation of the β -catenin gene. *The EMBO Journal*, 18(21), 5931-5942.
- Hart, M., Concordet, J. P., Lassot, I., Albert, I., Del los Santos, R., Durand, H., ... & Polakis, P. (1999). The F-box protein β -TrCP associates with phosphorylated β -catenin and regulates its activity in the cell. *Current Biology*, 9(4), 207-211.
- Hayashi, T., Asano, H., Toyooka, S., Tsukuda, K., Soh, J., Shien, T., ...& Nasu, Y. (2012). DNA methylation status of REIC/Dkk-3 gene in human malignancies. *Journal of Cancer Research and Clinical Oncology*, 138(5), 799-809.
- He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., Da Costa, L. T., ... & Kinzler, K. W. (1998). Identification of c-MYC as a target of the APC pathway. *Science*, 281(5382), 1509-1512.
- Heuberger, J., & Birchmeier, W. (2010). Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. *Cold Spring Harbor Perspectives in Biology*, 2(2), a002915.
- Hirata, H., Hinoda, Y., Nakajima, K., Kawamoto, K., Kikuno, N., Kawakami, K., ... & Ishii, N. (2009) .Wnt antagonist gene DKK2 is epigenetically silenced and inhibits renal cancer progression through apoptotic and cell cycle pathways. *Clinical Cancer Research*, 15(18), 5678-5687.
- Hirata, H., Hinoda, Y., Nakajima, K., Kikuno, N., Yamamura, S., Kawakami, K., ...& Dahiya, R. (2009). Wnt antagonist gene polymorphisms and renal cancer. *Cancer*, 115(19), 4488-4503.
- Hirata, H., Hinoda, Y., Ueno, K., Majid, S., Saini, S., & Dahiya, R. (2010). Role of secreted frizzled-related protein 3 in human renal cell carcinoma. *Cancer Research*, 70(5), 1896-1905.
- Hoffmann, D., Hoffmann, I., & El-Bayoumy, K. (2001). The less harmful cigarette: a controversial issue. A tribute to Ernst L. Wynder. *Chemical Research in Toxicology*, 14(7), 767-790.
- Hommura, F., Furuuchi, K., Yamazaki, K., Ogura, S., Kinoshita, I., Shimizu, M., ...& Dosaka-Akita, H. (2002). Increased expression of β -catenin predicts better prognosis in nonsmall cell lung carcinomas. *Cancer*, 94(3), 752-758.
- Horne, B. D., Anderson, J. L., Carlquist, J. F., Muhlestein, J. B., Renlund, D. G., Bair, T. L., ... & Camp, N. J. (2005). Generating genetic risk scores from intermediate phenotypes for use in association studies of clinically significant endpoints. *Annals of Human Genetics*, 69(2), 176-186.

- Hsieh, J. C., Kodjabachian, L., Rebbert, M. L., Rattner, A., Smallwood, P. M., Samos, C. H., ...& Nathans, J. (1999). A new secreted protein that binds to Wnt proteins and inhibits their activities. *Nature*, 398(6726), 431-436.
- Hsieh, S. Y., Hsieh, P. S., Chiu, C. T., & Chen, W. Y. (2004). Dickkopf-3/REIC functions as a suppressor gene of tumor growth. *Oncogene*, 23(57), 9183-9189.
- Huang, H., & He, X. (2008). Wnt/ β -catenin signaling: new (and old) players and new insights. *Current Opinion in Cell Biology*, 20(2), 119-125.
- Huber, A. H., Nelson, W. J., & Weis, W. I. (1997). Three-dimensional structure of the armadillo repeat region of β -catenin. *Cell*, 90(5), 871-882.
- Huber, O., Korn, R., McLaughlin, J., Ohsugi, M., Herrmann, B. G., & Kemler, R. (1996). Nuclear localization of β -catenin by interaction with transcription factor LEF-1. *Mechanisms of Development*, 59(1), 3-10.
- Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C., & Birchmeier, W. (2000). Requirement for β -catenin in anterior-posterior axis formation in mice. *The Journal of Cell Biology*, 148(3), 567-578.
- Hyde, L., & Hyde, C. I. (1974). Clinical manifestations of lung cancer. *Chest*, 65(3), 299-306.
- IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization, & International Agency for Research on Cancer. (2004). *Tobacco Smoke and Involuntary Smoking* (Vol. 83). Iarc.
- Ikeda, S., Kishida, S., Yamamoto, H., Murai, H., Koyama, S., & Kikuchi, A. (1998). Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 β and β -catenin and promotes GSK-3 β -dependent phosphorylation of β -catenin. *The EMBO journal*, 17(5), 1371-1384.
- Inomata, M., Ochiai, A., Akimoto, S., Kitano, S., & Hirohashi, S. (1996). Alteration of β -catenin expression in colonic epithelial cells of familial adenomatous polyposis patients. *Cancer Research*, 56(9), 2213-2217.
- Järup, L., & Pershagen, G. (1991). Arsenic exposure, smoking, and lung cancer in smelter workers- a case-control study. *American Journal of Epidemiology*, 134(6), 545-551.
- Jindal, S. K., & Behera, D. (1990). Clinical spectrum of primary lung cancer-review of Chandigarh experience of 10 years. *Lung India*, 8(2), 94.
- Johansson, B. G. (1972). Agarose gel electrophoresis. *Scandinavian Journal of Clinical and Laboratory Investigation*, 29(sup124), 7-19.

- Jones, S. E., & Jomary, C. (2002). Secreted Frizzled-related proteins: searching for relationships and patterns. *Bioessays*, 24(9), 811-820.
- Jung, K., Fleischhacker, M., & Rabien, A. (2010). Cell-free DNA in the blood as a solid tumor biomarker—a critical appraisal of the literature. *Clinica Chimica Acta*, 411(21), 1611-1624.
- Kadowaki, T., Wilder, E., Klingensmith, J., Zachary, K., & Perrimon, N. (1996). The segment polarity gene porcupine encodes a putative multitransmembrane protein involved in Wingless processing. *Genes & Development*, 10(24), 3116-3128.
- Kaiser, U., Schilli, M., Haag, U., Neumann, K., Kreipe, H., Kogan, E., & Havemann, K. (1996). Expression of bcl-2—protein in small cell lung cancer. *Lung Cancer*, 15(1), 31-40.
- Kanzaki, H., Ouchida, M., Hanafusa, H., Yano, M., Suzuki, H., Aoe, M., ...& Shimizu, K. (2006). Single nucleotide polymorphism of the AXIN2 gene is preferentially associated with human lung cancer risk in a Japanese population. *International Journal of Molecular Medicine*, 18(2), 279-284.
- Karimi, M. Y., Kapoor, V., Sharma, S. C., & Das, S. N. (2013). Genetic polymorphisms in FAS (CD95) and FAS ligand (CD178) promoters and risk of tobacco-related oral carcinoma: gene–gene interactions in high-risk Indians. *Cancer investigation*, 31(1), 1-6.
- Katoh, M., Kirikoshi, H., Terasaki, H., & Shiokawa, K. (2001). WNT2B2 mRNA, up-regulated in primary gastric cancer, is a positive regulator of the WNT– β -catenin–TCF signaling pathway. *Biochemical and Biophysical Research Communications*, 289(5), 1093-1098.
- Kaur, N., Chettiar, S., Rathod, S., Rath, P., Muzumdar, D., Shaikh, M. L., & Shiras, A. (2013). Wnt3a mediated activation of Wnt/ β -catenin signaling promotes tumor progression in glioblastoma. *Molecular and Cellular Neuroscience*, 54, 44-57.
- Kawano, Y., & Kypta, R. (2003). Secreted antagonists of the Wnt signalling pathway. *Journal of Cell Science*, 116(13), 2627-2634.
- Kawano, Y., Kitaoka, M., Hamada, Y., Walker, M. M., Waxman, J., & Kypta, R. M. (2006). Regulation of prostate cell growth and morphogenesis by Dickkopf-3. *Oncogene*, 25(49), 6528-6537.
- Kazi, M. M., Trivedi, T. I., Kobawala, T. P., & Ghosh, N. R. (2016). The Potential of Wnt signaling pathway in cancer: A focus on breast cancer. *Cancer Translational Medicine*, 2(2), 55.
- Kelly, O. G., Pinson, K. I., & Skarnes, W. C. (2004). The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development*, 131(12), 2803-2815.

- Keum, J. S., Kong, G., Yang, S. C., Shin, D. H., Park, S. S., Lee, J. H., & Lee, J. D. (1999). Cyclin D1 overexpression is an indicator of poor prognosis in resectable non-small cell lung cancer. *British Journal of Cancer*, 81(1), 127-132.
- Key Statistics for Lung Cancer. (2017, January 2). Retrieved October 31, 2017, from <https://www.cancer.org/cancer/non-small-cell-lung-cancer/about/key-statistics.html>
- Khuder, S. A. (2001). Effect of cigarette smoking on major histological types of lung cancer: a meta-analysis. *Lung cancer*, 31(2), 139-148.
- Kikuchi, A. (1999). Modulation of Wnt signaling by Axin and Axil. *Cytokine & Growth Factor Reviews*, 10(3), 255-265.
- Kim, M. S., Lee, H. N., Kim, H. J., & Myung, S. C. (2015). Single nucleotide polymorphisms in DKK3 gene are associated with prostate cancer risk and progression. *International Brazillian Journal of Urology*, 41(5), 869-897.
- Kim, Y. M., & Kahn, M. (2014). The role of the Wnt signaling pathway in cancer stem cells: prospects for drug development. *Research and Reports in Biochemistry*, 4, 1.
- Kirikoshi, H., Sekihara, H., & Katoh, M. (2001). Expression of WNT14 and WNT14B mRNAs in human cancer, up-regulation of WNT14 by IFN γ and up-regulation of WNT14B by β -estradiol. *International Journal of Oncology*, 19(6), 1221-1225.
- Kitamura, H., Yazawa, T., Okudela, K., Shimoyamada, H., & Sato, H. (2008). Molecular and genetic pathogenesis of lung cancer: differences between small-cell and non-small-cell carcinomas. *The Open Pathology Journal*, 2(1).
- Klaus, A., & Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer*, 8(5), 387-398.
- Klaus, A., & Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer*, 8(5), 387-398.
- Kobayashi, K., Ouchida, M., Tsuji, T., Hanafusa, H., Miyazaki, M., Namba, M., ...& Shimizu, K. (2002). Reduced expression of the REIC/Dkk-3 gene by promoter-hypermethylation in human tumor cells. *Gene*, 282(1), 151-158.
- Koinuma, K., Yamashita, Y., Liu, W., Hatanaka, H., Kurashina, K., Wada, T., ...& Miyakura, Y. (2006). Epigenetic silencing of AXIN2 in colorectal carcinoma with microsatellite instability. *Oncogene*, 25(1), 139-146.
- Kondo, T., Yamada, K., Noda, K., Nakayama, H., & Kameda, Y. (2002). Radiologic–prognostic correlation in patients with small pulmonary adenocarcinomas. *Lung Cancer*, 36(1), 49-57.

- Kongkham, P. N., Northcott, P. A., Croul, S. E., Smith, C. A., Taylor, M. D., & Rutka, J. T. (2010). The SFRP family of WNT inhibitors function as novel tumor suppressor genes epigenetically silenced in medulloblastoma. *Oncogene*, 29(20), 3017-3024.
- Koo, B. K., Spit, M., Jordens, I., Low, T. Y., Stange, D. E., van de Wetering, M., ...& Clevers, H. (2012). Tumour suppressor RNF43 is a stem-cell E3 ligase that induces endocytosis of Wnt receptors. *Nature*, 488(7413), 665-669.
- Kool, M., Koster, J., Bunt, J., Hasselt, N. E., Lakeman, A., Van Sluis, P., ...& Mršić, A. (2008). Integrated genomics identifies five medulloblastoma subtypes with distinct genetic profiles, pathway signatures and clinicopathological features. *PloS One*, 3(8), e3088.
- Korinek, V., Barker, N., Moerer, P., van Donselaar, E., Huls, G., Peters, P. J., & Clevers, H. (1998). Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nature Genetics*, 19(4), 379-383.
- Kotnis, A., Namkung, J., Kannan, S., Jayakrupakar, N., Park, T., Sarin, R., & Mulherkar, R. (2012). Multiple pathway-based genetic variations associated with tobacco related multiple primary neoplasms. *PloS One*, 7(1), e30013.
- Kren, L., Hermanova, M., Goncharuk, V. N., Kaur, P., Ross, J. S., Pavlovský, Z., & Dvorak, K. (2003). Downregulation of plasma membrane expression/cytoplasmic accumulation of beta-catenin predicts shortened survival in non-small cell lung cancer. A clinicopathologic study of 100 cases. *Ceskoslovenska Patologie*, 39(1), 17-20.
- Krupnik, V. E., Sharp, J. D., Jiang, C., Robison, K., Chickering, T. W., Amaravadi, L., ...& Chang, B. (1999). Functional and structural diversity of the human Dickkopf gene family. *Gene*, 238(2), 301-313.
- Kumar, R., Prakash, S., Kushwah, A. S., & Vijayan, V. K. (2010). Breath carbon monoxide concentration in cigarette and bidi smokers in India. *Indian Journal of Chest and Allied diseases*, 52(1), 19-24.
- Kuphal, S., Lodermeier, S., Bataille, F., Schuierer, M., Hoang, B. H., & Bosserhoff, A. K. (2006). Expression of Dickkopf genes is strongly reduced in malignant melanoma. *Oncogene*, 25(36), 5027-5036.
- Kurose, K., Sakaguchi, M., Nasu, Y., Ebara, S., Kaku, H., Kariyama, R., ...& Kumon, H. (2004). Decreased expression of REIC/Dkk-3 in human renal clear cell carcinoma. *The Journal of Urology*, 171(3), 1314-1318.

- Kusserow, A., Pang, K., Sturm, C., Hroudá, M., Lentfer, J., Schmidt, H. A., ...& Holstein, T. W. (2005). Unexpected complexity of the Wnt gene family in a sea anemone. *Nature*, 433(7022), 156-160.
- Laird, P. W., Jackson-Grusby, L., Fazeli, A., Dickinson, S. L., Jung, W. E., Li, E., ... & Jaenisch, R. (1995). Suppression of intestinal neoplasia by DNA hypomethylation. *Cell*, 81(2), 197-205.
- Lamb, R., Ablett, M. P., Spence, K., Landberg, G., Sims, A. H., & Clarke, R. B. (2013). Wnt pathway activity in breast cancer sub-types and stem-like cells. *PloS One*, 8(7), e67811.
- Lammi, L., Arte, S., Somer, M., Järvinen, H., Lahermo, P., Thesleff, I., ...& Nieminen, P. (2004). Mutations in AXIN2 cause familial tooth agenesis and predispose to colorectal cancer. *The American Journal of Human Genetics*, 74(5), 1043-1050.
- Lamouille, S., Xu, J., & Derynck, R. (2014). Molecular mechanisms of epithelial–mesenchymal transition. *Nature Reviews Molecular cell biology*, 15(3), 178-196.
- Lantz, D. A., Dougherty, T. H., & Lucca, M. J. (1989). Primary angiosarcoma of the heart causing cardiac rupture. *American Heart Journal*, 118(1), 186-188.
- Latimer, K. M., & Mott, T. F. (2015). Lung cancer: diagnosis, treatment principles, and screening. *American Family Physician*, 91(4).
- Lee, E. J., Jo, M., Rho, S. B., Park, K., Yoo, Y. N., Park, J., ...& Lee, J. H. (2009). Dkk3, downregulated in cervical cancer, functions as a negative regulator of β -catenin. *International Journal of Cancer*, 124(2), 287-297.
- Lee, J. L., Chang, C. J., Chueh, L. L., & Lin, C. T. (2006). Secreted frizzled related protein 2 (sFRP2) decreases susceptibility to UV-induced apoptosis in primary culture of canine mammary gland tumors by NF- κ B activation or JNK suppression. *Breast Cancer Research and Treatment*, 100(1), 49-58.
- Lee, S., Kang, J., Cho, M., Seo, E., Choi, H., Kim, E., ...& Park, Y. H. (2009). Profiling of transcripts and proteins modulated by K-ras oncogene in the lung tissues of K-ras transgenic mice by omics approaches. *International Journal of Oncology*, 34(1), 161-172.
- Lemjabbar-Alaoui, H., Dasari, V., Sidhu, S. S., Mengistab, A., Finkbeiner, W., Gallup, M., & Basbaum, C. (2006). Wnt and Hedgehog are critical mediators of cigarette smoke-induced lung cancer. *PLoS One*, 1(1), e93.
- Lepourcelet, M., Chen, Y. N. P., France, D. S., Wang, H., Crews, P., Petersen, F., ...& Shivdasani, R. A. (2004). Small-molecule antagonists of the oncogenic Tcf/ β -catenin protein complex. *Cancer Cell*, 5(1), 91-102.

- Leppert, M., Dobbs, M., Scambler, P., O'connell, P., Nakamura, Y., Stauffer, D., & Lathrop, M. (1987). The gene for familial polyposis coli maps to the long arm of chromosome 5. *Science*, 1411-1413.
- Levanon, D., Goldstein, R. E., Bernstein, Y., Tang, H., Goldenberg, D., Stifani, S., ...& Groner, Y. (1998). Transcriptional repression by AML1 and LEF-1 is mediated by the TLE/Groucho corepressors. *Proceedings of the National Academy of Sciences*, 95(20), 11590-11595.
- Li, S., Wang, C., Liu, X., Hua, S., & Liu, X. (2015). The roles of AXIN2 in tumorigenesis and epigenetic regulation. *Familial Cancer*, 14(2), 325-331.
- Li, X., Zhang, Y., Kang, H., Liu, W., Liu, P., Zhang, J., ...& Wu, D. (2005). Sclerostin binds to LRP5/6 and antagonizes canonical Wnt signaling. *Journal of Biological Chemistry*, 280(20), 19883-19887.
- Li, Y., Sheu, C. C., Ye, Y., De Andrade, M., Wang, L., Chang, S. C., ...& Cunningham, J. M. (2010). Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *The Lancet Oncology*, 11(4), 321-330.
- Licchesi, J. D., Westra, W. H., Hooker, C. M., Machida, E. O., Baylin, S. B., & Herman, J. G. (2008). Epigenetic alteration of Wnt pathway antagonists in progressive glandular neoplasia of the lung. *Carcinogenesis*, 29(5), 895-904.
- Lien, W. H., & Fuchs, E. (2014). Wnt some lose some: transcriptional governance of stem cells by Wnt/ β -catenin signaling. *Genes & Development*, 28(14), 1517-1532.
- Lin, K., Wang, S., Julius, M. A., Kitajewski, J., Moos, M., & Luyten, F. P. (1997). The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling. *Proceedings of the National Academy of Sciences*, 94(21), 11196-11200.
- Lindsley, R. C., Gill, J. G., Kyba, M., Murphy, T. L., & Murphy, K. M. (2006). Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. *Development*, 133(19), 3787-3796.
- Liu, D., Li, L., Yang, Y., Liu, W., & Wu, J. (2014). The Axin2 rs2240308 polymorphism and susceptibility to lung cancer in a Chinese population. *Tumor Biology*, 35(11), 10987-10991.
- Liu, F., Killian, J. K., Yang, M., Walker, R. L., Hong, J. A., Zhang, M., ...& Rao, M. (2010). Epigenomic alterations and gene expression profiles in respiratory epithelia exposed to cigarette smoke condensate. *Oncogene*, 29(25), 3650-3664.
- Liu, L., Wu, C., Wang, Y., Zhong, R., Wang, F., Zhang, X., ...& Yuan, J. (2010). Association of candidate genetic variations with gastric cardia adenocarcinoma in Chinese population: a multiple interaction analysis. *Carcinogenesis*, 32(3), 336-342.

- Liu, W., Dong, X., Mai, M., Seelan, R. S., Taniguchi, K., Krishnadath, K. K., ...& Roche, P. C. (2000). Mutations in AXIN2 cause colorectal cancer with defective mismatch repair by activating β -catenin/TCF signalling. *Nature Genetics*, 26(2), 146-147.
- Lodygin, D., Epanchintsev, A., Menssen, A., Diebold, J., & Hermeking, H. (2005). Functional epigenomics identifies genes frequently silenced in prostate cancer. *Cancer Research*, 65(10), 4218-4227.
- Logan, C. Y., & Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annual Review of Cell and Developmental Biology*, 20, 781-810.
- Loh, K. M., van Amerongen, R., & Nusse, R. (2016). Generating cellular diversity and spatial form: Wnt signaling and the evolution of multicellular animals. *Developmental Cell*, 38(6), 643-655.
- Longman, D., Arfuso, F., Viola, H. M., Hool, L. C., & Dharmarajan, A. M. (2012). The role of the cysteine-rich domain and netrin-like domain of secreted frizzled-related protein 4 in angiogenesis inhibition in vitro. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics*, 20(1), 1-6.
- Longo, K. A., Kennell, J. A., Ochocinska, M. J., Ross, S. E., Wright, W. S., & MacDougald, O. A. (2002). Wnt signaling protects 3T3-L1 preadipocytes from apoptosis through induction of insulin-like growth factors. *Journal of Biological Chemistry*, 277(41), 38239-38244.
- Loprieno, N. (1975). International Agency for Research on Cancer (IARC) monographs on the evaluation of carcinogenic risk of chemicals to man: "Relevance of data on mutagenicity". *Mutation Research/Environmental Mutagenesis and Related Subjects*, 31(3), 201.
- Loughlin, J., Dowling, B., Chapman, K., Marcelline, L., Mustafa, Z., Southam, L., ...& Corr, M. (2004). Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. *Proceedings of the National Academy of Sciences of the United States of America*, 101(26), 9757-9762.
- Lung cancer - Genetics Home Reference. (2017.). Retrieved October 31, 2017, from <https://ghr.nlm.nih.gov/condition/lung-cancer>
- Lung Cancer. (2017). Retrieved October 31, 2017, from <https://www.cancer.org/cancer/lung-cancer.html>
- Luo, J., Chen, J., Deng, Z. L., Luo, X., Song, W. X., Sharff, K. A., ... & He, T. C. (2007). Wnt signaling and human diseases: what are the therapeutic implications?. *Laboratory Investigation*, 87(2), 97-103.

- Ma, C., Liu, C., Huang, P., Kaku, H., Chen, J., Guo, K., ...& Shimizu, K. (2014). Significant association between the Axin2 rs2240308 single nucleotide polymorphism and the incidence of prostate cancer. *Oncology Letters*, 8(2), 789-794.
- MacDonald, B. T., Tamai, K., & He, X. (2009). Wnt/ β -catenin signaling: components, mechanisms, and diseases. *Developmental Cell*, 17(1), 9-26.
- Mahimkar, M. B., Samant, T. A., Kannan, S., Tulsulkar, J., Pai, P. S., & Anantharaman, D. (2012). Polymorphisms in GSTM1 and XPD genes predict clinical outcome in advanced oral cancer patients treated with postoperative radiotherapy. *Molecular carcinogenesis*, 51(S1), E94-103.
- Major cancer types by deaths worldwide (2015). Statistic. Retrieved October 31, 2017, from <https://www.statista.com/statistics/288580/number-of-cancer-deaths-worldwide-by-type/>
- Malhotra, S., & Kincade, P. W. (2009). Wnt-related molecules and signaling pathway equilibrium in hematopoiesis. *Cell Stem Cell*, 4(1), 27-36.
- Malson, J. L., Sims, K., Murty, R., & Pickworth, W. B. (2001). Comparison of the nicotine content of tobacco used in bidis and conventional cigarettes. *Tobacco Control*, 10(2), 181-183.
- Mann, B., Gelos, M., Siedow, A., Hanski, M. L., Gratchev, A., Ilyas, M., ...& Hanski, C. (1999). Target genes of β -catenin–T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. *Proceedings of the National Academy of Sciences*, 96(4), 1603-1608.
- Mao, B., & Niehrs, C. (2003). Kremen2 modulates Dickkopf2 activity during Wnt/LRP6 signaling. *Gene*, 302(1), 179-183.
- Mao, B., Wu, W., Davidson, G., Marhold, J., Li, M., Mechler, B. M., ...& Glinka, A. (2002). Kremen proteins are Dickkopf receptors that regulate Wnt/ β -catenin signalling. *Nature*, 417(6889), 664-667.
- Mao, J., Wang, J., Liu, B., Pan, W., Farr, G. H., Flynn, C., ...& Wu, D. (2001). Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway. *Molecular Cell*, 7(4), 801-809.
- Marshall, A. L., & Christiani, D. C. (2013). Genetic susceptibility to lung cancer—light at the end of the tunnel?. *Carcinogenesis*, 34(3), 487-502.
- Marsit, C. J., Karagas, M. R., Andrew, A., Liu, M., Danaee, H., Schned, A. R., ...& Kelsey, K. T. (2005). Epigenetic inactivation of SFRP genes and TP53 alteration act jointly as markers of invasive bladder cancer. *Cancer Research*, 65(16), 7081-7085.
- Martello, G., Zacchigna, L., Inui, M., Montagner, M., Adorno, M., Mamidi, A., ...& Cordenonsi, M. (2007). MicroRNA control of Nodal signalling. *Nature*, 449(7159), 183-188.

- Massion, P. P., & Carbone, D. P. (2003). The molecular basis of lung cancer: molecular abnormalities and therapeutic implications. *Respiratory Research*, 4(1), 12.
- Mazieres, J., He, B., You, L., Xu, Z., & Jablons, D. M. (2005). Wnt signaling in lung cancer. *Cancer Letters*, 222(1), 1-10.
- McCrea, P. D., Turck, C. W., & Gumbiner, B. (1991). A homolog of the armadillo protein in *Drosophila* (plakoglobin) associated with E-cadherin. *Science*, 254(5036), 1359-1361.
- McMahon, A. P., & Moon, R. T. (1989). Ectopic expression of the proto-oncogene int-1 in *Xenopus* embryos leads to duplication of the embryonic axis. *Cell*, 58(6), 1075-1084.
- Merrill, B. J. (2012). Wnt pathway regulation of embryonic stem cell self-renewal. *Cold Spring Harbor Perspectives in Biology*, 4(9), a007971.
- Mishra, S., Joseph, R. A., Gupta, P. C., Pezzack, B., Ram, F., Sinha, D. N., & Jha, P. (2016). Trends in bidi and cigarette smoking in India from 1998 to 2015, by age, gender and education. *BMJ Global Health*, 1(1), e000005.
- Mizobuchi, Y., Matsuzaki, K., Kuwayama, K., Kitazato, K., Mure, H., Kageji, T., & Nagahiro, S. (2008). REIC/Dkk-3 induces cell death in human malignant glioma. *Neuro-oncology*, 10(3), 244-253.
- Moser, A. R., Pitot, H. C., & Dove, W. F. (1990). A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science*, 247(4940), 322-324.
- Mosimann, C., Hausmann, G., & Basler, K. (2006). Parafibromin/Hyrax activates Wnt/Wg target gene transcription by direct association with β -catenin/Armadillo. *Cell*, 125(2), 327-341.
- Mostowska, A., Biedziak, B., & Jagodzinski, P. P. (2006). Axis inhibition protein 2 (AXIN2) polymorphisms may be a risk factor for selective tooth agenesis. *Journal of Human Genetics*, 51(3), 262-266.
- Mostowska, A., Pawlik, P., Sajdak, S., Markowska, J., Pawłowska, M., Lianeri, M., & Jagodzinski, P. P. (2014). An analysis of polymorphisms within the Wnt signaling pathway in relation to ovarian cancer risk in a Polish population. *Molecular Diagnosis & Therapy*, 18(1), 85-91.
- Mothersill, C., Seymour, C. B., & O'Brien, A. (1991). Induction of c-myc oncoprotein and of cellular proliferation by radiation in normal human urothelial cultures. *Anticancer Research*, 11(4), 1609-1612.
- Mucci, L. A., Wedren, S., Tamimi, R. M., Trichopoulos, D., & Adami, H. O. (2001). The role of gene-environment interaction in the aetiology of human cancer: examples from cancers of the large bowel, lung and breast. *Journal of Internal Medicine*, 249(6), 477-493.

- Müller-Hermelink, H. K., Engel, P., Kuo, T. T., Ströbel, P., Marx, A., Harris, N. L., & Harris, C. C. (2004). Pathology and Genetics, Tumours of the lung, pleura, thymus and heart. *World Health Organization Classification of Tumors, IARC Press, Lyon, France*, 146-147.
- Mullighan, C. G., Zhang, J., Kasper, L. H., Lerach, S., Payne-Turner, D., Phillips, L. A., ...& Buetow, K.H. (2011). CREBBP mutations in relapsed acute lymphoblastic leukaemia. *Nature*, 471(7337), 235-239.
- Munemitsu, S., Albert, I., Souza, B., Rubinfeld, B., & Polakis, P. (1995). Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (APC) tumor-suppressor protein. *Proceedings of the National Academy of Sciences*, 92(7), 3046-3050.
- Nagaishi, C., & Nagasawa, N. (1972). *Functional Anatomy and Histology of the Lung*. University Park Press.
- Naghibalhossaini, F., Zamani, M., Mokarram, P., Khalili, I., Rasti, M., & Mostafavi-pour, Z. (2012). Epigenetic and genetic analysis of WNT signaling pathway in sporadic colorectal cancer patients from Iran. *Molecular Biology Reports*, 39(5), 6171-6178.
- Nakamoto, M., Matsuyama, A., Shiba, E., Shibuya, R., Kasai, T., Yamaguchi, K., & Hisaoka, M. (2014). Prognostic significance of WNT signaling in pancreatic ductal adenocarcinoma. *Virchows Archiv*, 465(4), 401-408.
- National Research Council. (1988). *Health Risks of Radon and Other Internally Deposited Alpha-Emitters: BEIR IV* (Vol. 4). National Academies Press.
- Neumann, A. S., Sturgis, E. M., & Wei, Q. (2005). Nucleotide excision repair as a marker for susceptibility to tobacco-related cancers: a review of molecular epidemiological studies. *Molecular Carcinogenesis*, 42(2), 65-92.
- Nguyen, D. X., Chiang, A. C., Zhang, X. H. F., Kim, J. Y., Kris, M. G., Ladanyi, M., ... & Massagué, J. (2009). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell*, 138(1), 51-62.
- Nicholson, S. A., Beasley, M. B., Brambilla, E., Hdtasleton, P. S., Colby, T. V., Sheppard, M. N., & Travis, W. D. (2002). Small cell lung carcinoma (SCLC): a clinicopathologic study of 100 cases with surgical specimens. *The American Journal of Surgical Pathology*, 26(9), 1184-1197.
- Niehrs, C. (2006). Function and biological roles of the Dickkopf family of Wnt modulators. *Oncogene*, 25(57), 7469-7481.
- Niehrs, C., & Acebron, S. P. (2012). Mitotic and mitogenic Wnt signalling. *The EMBO Journal*, 31(12), 2705-2713.

- Nojima, M., Suzuki, H., Toyota, M., Watanabe, Y., Maruyama, R., Sasaki, S., ...& Hirata, K. (2007). Frequent epigenetic inactivation of SFRP genes and constitutive activation of Wnt signaling in gastric cancer. *Oncogene*, 26(32), 4699-4713.
- Noordermeer, J., Klingensmith, J., & Perrimonl, N. (1994). dishevelled and armadillo act. *Nature*, 367, 80-82.
- Nozaki, I., Tsuji, T., Iijima, O., Ohmura, Y., Andou, A., Miyazaki, M., ...& Namba, M. (2001). Reduced expression of REIC/Dkk-3 gene in non-small cell lung cancer. *International Journal of Oncology*, 19(1), 117-121.
- Nuciforo, P., & Fraggetta, F. (2004). Cancer stem cell theory: pathologists' considerations and ruminations about wasting time and wrong evaluations. *Journal of Clinical Pathology*, 57(7), 782-782.
- Nusse, R., & Varmus, H. E. (1982). Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell*, 31(1), 99-109.
- Nusse, R., Brown, A., Papkoff, J., Scambler, P., Shackleford, G., McMahon, A., & Varmus, H. (1991). A new nomenclature for int-1 and related genes: the Wnt gene family. *Cell*, 64(2), 231.
- Okino, K., Nagai, H., Hatta, M., Nagahata, T., Yoneyama, K., Ohta, Y., ...& Emi, M. (2003). Up-regulation and overproduction of DVL-1, the human counterpart of the Drosophila dishevelled gene, in cervical squamous cell carcinoma. *Oncology Reports*, 10(5), 1219-1223.
- Oren, M. (2003). Decision making by p53: life, death and cancer. *Cell Death & Differentiation*, 10(4), 431-442.
- Osada, H., Tomida, S., Yatabe, Y., Tatematsu, Y., Takeuchi, T., Murakami, H., ...& Takahashi, T. (2008). Roles of achaete-scute homologue 1 in DKK1 and E-cadherin repression and neuroendocrine differentiation in lung cancer. *Cancer Research*, 68(6), 1647-1655.
- Paez, D., Gerger, A., Zhang, W., Yang, D., Labonte, M. J., Benhanim, L., ...& Wakatsuki, T. (2014). Association of common gene variants in the WNT/ β -catenin pathway with colon cancer recurrence. *The Pharmacogenomics Journal*, 14(2), 142-150.
- Panov, S. Z. (2005). Molecular biology of the lung cancer. *Radiology and Oncology*, 39(3).
- Pao, W., Miller, V. A., Politi, K. A., Riely, G. J., Somwar, R., Zakowski, M. F., ... & Varmus, H. (2005). Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain. *PLoS Medicine*, 2(3), e73.
- Parkin, D. M., Whelan, S. L., Ferlay, J., Teppo, L., & Thomas, D. B. (2002). Cancer incidence in five continents. Vol. VIII. *IARC Scientific Publication* No. 155. Lyon, France: International Agency for Research on Cancer.

- Peifer, M., Sweeton, D., Casey, M., & Wieschaus, E. (1994). Wingless signal and Zeste-white 3 kinase trigger opposing changes in the intracellular distribution of Armadillo. *Development*, 120(2), 369-380.
- Perrimon, N., Engstrom, L., & Mahowald, A. P. (1989). Zygotic lethals with specific maternal effect phenotypes in *Drosophila melanogaster*. I. Loci on the X chromosome. *Genetics*, 121(2), 333-352.
- Peters, J. M., McKay, R. M., McKay, J. P., & Graff, J. M. (1999). Casein kinase I transduces Wnt signals. *Nature*, 401(6751), 345-350.
- Pfeifer, G. P., Denissenko, M. F., Olivier, M., Tretyakova, N., Hecht, S. S., & Hainaut, P. (2002). Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking-associated cancers. *Oncogene*, 21(48), 7435.
- Phelan, C. M., Borg, Å., Cuny, M., Crichton, D. N., Baldersson, T., Andersen, T. I., ... & Kelsell, D. (1998). Consortium study on 1280 breast carcinomas: allelic loss on chromosome 17 targets subregions associated with family history and clinical parameters. *Cancer Research*, 58(5), 1004-1012.
- Pheesse, T. J., Parry, L., Reed, K. R., Ewan, K. B., Dale, T. C., Sansom, O. J., & Clarke, A. R. (2008). Deficiency of Mbd2 attenuates Wnt signaling. *Molecular and Cellular Biology*, 28(19), 6094-6103.
- Pinson, K. I., Brennan, J., Monkley, S., Avery, B. J., & Skarnes, W. C. (2000). An LDL-receptor-related protein mediates Wnt signalling in mice. *Nature*, 407(6803), 535-538.
- Pinto, D., Gregorieff, A., Begthel, H., & Clevers, H. (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & Development*, 17(14), 1709-1713.
- Pinto, D., Gregorieff, A., Begthel, H., & Clevers, H. (2003). Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes & Development*, 17(14), 1709-1713.
- Pohl, S., Scott, R., Arfuso, F., Perumal, V., & Dharmarajan, A. (2015). Secreted frizzled-related protein 4 and its implications in cancer and apoptosis. *Tumor Biology*, 36(1), 143-152.
- Polakis, P. (2000). Wnt signaling and cancer. *Genes & Development*, 14(15), 1837-185
- Polakis, P. (2007). The many ways of Wnt in cancer. *Current Opinion in Genetics & Development*, 17(1), 45-51.
- Polakis, P. (2012). Wnt signaling in cancer. *Cold Spring Harbor Perspectives in Biology*, 4(5), a008052.
- Pongracz, J. E., & Stockley, R. A. (2006). Wnt signalling in lung development and diseases. *Respiratory Research*, 7(1), 15.

- Preisinger, A. C., Hedge, P. H. I. L. I. P., McKechnie, D. O. U. G. L. A. S., Finniear, R., Markham, A., Groffen, J., ... & Miyoshi, Y. (1991). Identification of FAP locus genes from chromosome 5q21. *Science*, 253(5020), 661-665.
- Prelaj, A., Tay, R., Ferrara, R., Chaput, N., Besse, B., & Califano, R. (2019). Predictive biomarkers of response for immune checkpoint inhibitors in non-small-cell lung cancer. *European Journal of Cancer*, 106, 144-159.
- Previdi, S., Maroni, P., Matteucci, E., Broggin, M., Bendinelli, P., & Desiderio, M. A. (2010). Interaction between human-breast cancer metastasis and bone microenvironment through activated hepatocyte growth factor/Met and β -catenin/Wnt pathways. *European Journal of Cancer*, 46(9), 1679-1691.
- Pu, X., Ye, Y., Spitz, M. R., Wang, L., Gu, J., Lippman, S. M., ... & Yang, P. (2012). Predictors of Survival in Never-Smokers with Non-Small Cell Lung Cancer: A Large-Scale, Two-Phase Genetic Study. *Clinical Cancer Research*, 18(21), 5983-5991.
- Qian, D. C., Han, Y., Byun, J., Shin, H. R., Hung, R. J., McLaughlin, J. R., ... & Amos, C. I. (2016). A novel pathway-based approach improves lung cancer risk prediction using germline genetic variations. *Cancer Epidemiology and Prevention Biomarkers*.
- Quint, L. E., Tummala, S., Brisson, L. J., Francis, I. R., Krupnick, A. S., Kazerooni, E. A., & Orringer, M. B. (1996). Distribution of distant metastases from newly diagnosed non-small cell lung cancer. *The Annals of Thoracic Surgery*, 62(1), 246-250.
- Rami-Porta, R., Crowley, J. J., & Goldstraw, P. (2009). Review the revised TNM staging system for lung cancer. *Annals of Thoracic and Cardiovascular Surgery*, 15(1), 5.
- Rasmussen, H. B. (2012). Restriction fragment length polymorphism analysis of PCR-amplified fragments (PCR-RFLP) and gel electrophoresis-valuable tool for genotyping and genetic fingerprinting. In *Gel Electrophoresis-Principles and Basics*. InTech.
- Reed, J. C. (1997). Double identity for proteins of the Bcl-2 family. *Nature*, 387(6635), 773-776.
- Reya, T., Morrison, S. J., Clarke, M. F., & Weissman, I. L. (2001). Stem cells, cancer, and cancer stem cells. *Nature*, 414(6859), 105-111.
- Riese, J., Yu, X., Munnerlyn, A., Eresh, S., Hsu, S. C., Grosschedl, R., & Bienz, M. (1997). LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic. *Cell*, 88(6), 777-787.
- Riggleman, B., Schedl, P., & Wieschaus, E. (1990). Spatial expression of the Drosophila segment polarity gene armadillo is posttranscriptionally regulated by wingless. *Cell*, 63(3), 549-560.

- Rijsewijk, F., Schuermann, M., Wagenaar, E., Parren, P., Weigel, D., & Nusse, R. (1987). The *Drosophila* homology of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell*, 50(4), 649-657.
- Ritchie, M. D., & Motesinger, A. A. (2005). Multifactor dimensionality reduction for detecting gene–gene and gene–environment interactions in pharmacogenomics studies. *Pharmacogenomics*, 6(8), 823-834.
- Rodríguez-Paredes, M., & Esteller, M. (2011). Cancer epigenetics reaches mainstream oncology. *Nature Medicine*, 330-339.
- Rom, W. N., Hay, J. G., Lee, T. C., Jiang, Y., & Tchou-Wong, K. M. (2000). Molecular and genetic aspects of lung cancer. *American Journal of Respiratory and Critical Care Medicine*, 161(4), 1355-1367.
- Roman-Gomez, J., Jimenez-Velasco, A., Agirre, X., Castillejo, J. A., Navarro, G., Barrios, M., ...& Torres, A. (2004). Transcriptional silencing of the Dickkopf-3 (Dkk-3) gene by CpG hypermethylation in acute lymphoblastic leukaemia. *British Journal of Cancer*, 91(4), 707-713.
- Rossant, J., & Tam, P. P. (2004). Emerging asymmetry and embryonic patterning in early mouse development. *Developmental Cell*, 7(2), 155-164.
- Rubin, B. P., Skarin, A. T., Pisick, E., Rizk, M., & Salgia, R. (2001). Use of cytokeratins 7 and 20 in determining the origin of metastatic carcinoma of unknown primary, with special emphasis on lung cancer. *European Journal of Cancer Prevention*, 10(1), 77-82.
- Rubinfeld, B., Robbins, P., El-Gamil, M., Albert, I., Porfiri, E., & Polakis, P. (1997). Stabilization of β -catenin by genetic defects in melanoma cell lines. *Science*, 275(5307), 1790-1792.
- Rubinfeld, B., Souza, B., Albert, I., Muller, O., Chamberlain, S. H., Masiarz, F. R., & Polakis, P. (1993). Association of the APC gene product with beta-catenin. *Science*, 262(5140), 1731-1735.
- Rubinfeld, B., Tice, D. A., & Polakis, P. (2001). Axin-dependent phosphorylation of the adenomatous polyposis coli protein mediated by casein kinase 1 ϵ . *Journal of Biological Chemistry*, 276(42), 39037-39045.
- Ryu, J. S., Hong, Y. C., Han, H. S., Lee, J. E., Kim, S., Park, Y. M., ...& Hwang, T. S. (2004). Association between polymorphisms of ERCC1 and XPD and survival in non-small-cell lung cancer patients treated with cisplatin combination chemotherapy. *Lung Cancer*, 44(3), 311-316.
- Sakanaka, C., Weiss, J. B., & Williams, L. T. (1998). Bridging of β -catenin and glycogen synthase kinase-3 β by axin and inhibition of β -catenin-mediated transcription. *Proceedings of the National Academy of Sciences*, 95(6), 3020-3023.

- Salomon, D., Sacco, P. A., Roy, S. G., Simcha, I., Johnson, K. R., Wheelock, M. J., & Ben-Ze'ev, A. (1997). Regulation of β -catenin levels and localization by overexpression of plakoglobin and inhibition of the ubiquitin-proteasome system. *The Journal of Cell Biology*, 139(5), 1325-1335.
- Sambrook, J., & Russell, D. W. (2001). Molecular cloning: a laboratory manual. *Cold spring Harbor Laboratory Press*, New York.
- Samet, J. M. (1989). Radon and lung cancer. *JNCI:Journal of the National Cancer Institute*, 81(10), 745-758.
- Samet, J. M., Avila-Tang, E., Boffetta, P., Hannan, L. M., Olivo-Marston, S., Thun, M. J., & Rudin, C. M. (2009). Lung cancer in never smokers: clinical epidemiology and environmental risk factors. *Clinical Cancer Research*, 15(18), 5626-5645.
- Sato, H., Suzuki, H., Toyota, M., Nojima, M., Maruyama, R., Sasaki, S., ...& Sonoda, T. (2007). Frequent epigenetic inactivation of DICKKOPF family genes in human gastrointestinal tumors. *Carcinogenesis*, 28(12), 2459-2466.
- Sato, N., Fukushima, N., Maitra, A., Matsubayashi, H., Yeo, C. J., Cameron, J. L., ...& Goggins, M. (2003). Discovery of novel targets for aberrant methylation in pancreatic carcinoma using high-throughput microarrays. *Cancer Research*, 63(13), 3735-3742.
- Satoh, S., Daigo, Y., Furukawa, Y., Kato, T., Miwa, N., Nishiwaki, T., ...& Sasaki, Y. (2000). AXIN1 mutations in hepatocellular carcinomas, and growth suppression in cancer cells by virus-mediated transfer of AXIN1. *Nature Genetics*, 24(3), 245-250.
- Schneider, S., Steinbeisser, H., Warga, R. M., & Hausen, P. (1996). β -catenin translocation into nuclei demarcates the dorsalizing centers in frog and fish embryos. *Mechanisms of Development*, 57(2), 191-198.
- Schork, N. J., Fallin, D., & Lanchbury, J. S. (2000). Single nucleotide polymorphisms and the future of genetic epidemiology. *Clinical Genetics*, 58(4), 250-264.
- Schwarz-Romond, T., Fiedler, M., Shibata, N., Butler, P. J. G., Kikuchi, A., Higuchi, Y., & Bienz, M. (2007). The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization. *Nature Structural & Molecular Biology*, 14(6), 484-492.
- Seeling, J. M., Miller, J. R., Gil, R., Moon, R. T., White, R., & Virshup, D. M. (1999). Regulation of β -catenin signaling by the B56 subunit of protein phosphatase 2A. *Science*, 283(5410), 2089-2091.
- Serman, L., Martic, T. N., Serman, A., & Vranic, S. (2014). Epigenetic alterations of the Wnt signaling pathway in cancer: a mini review. *Bosnian Journal of Basic Medical Sciences*, 14(4), 191.

- Sethi, T., Langdon, S., Smyth, J., & Rozengurt, E. (1992). Growth of small cell lung cancer cells: stimulation by multiple neuropeptides and inhibition by broad spectrum antagonists in vitro and in vivo. *Cancer Research*, 52(9 Supplement), 2737s-2742s.
- Shanmugam, K. S., Brenner, H., Hoffmeister, M., Chang-Claude, J., & Burwinkel, B. (2007). The functional genetic variant Arg324Gly of frizzled-related protein is associated with colorectal cancer risk. *Carcinogenesis*, 28(9), 1914-1917.
- Shapiro, M., Akiri, G., Chin, C., Wisnivesky, J. P., Beasley, M. B., Weiser, T. S., ...& Aaronson, S. A. (2013). Wnt pathway activation predicts increased risk of tumor recurrence in patients with stage I non-small cell lung cancer. *Annals of Surgery*, 257(3), 548.
- Shastry, B. S. (2009). SNPs: impact on gene function and phenotype. *Single Nucleotide Polymorphisms: Methods and Protocols*, 3-22.
- Shen, L. X., Basilion, J. P., & Stanton, V. P. (1999). Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proceedings of the National Academy of Sciences*, 96(14), 7871-7876.
- Shi, Y., & He, L. (2009). Shesis, A powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *American Journal of Medical Genetics-A*, 138(1), 143.
- Shi, Y., He, B., You, L., & Jablons, D. M. (2007). Roles of secreted frizzled-related proteins in cancer. *Acta Pharmacologica Sinica*, 28(9), 1499-1504.
- Shih, Y. L., Shyu, R. Y., Hsieh, C. B., Lai, H. C., Liu, K. Y., Chu, T. Y., & Lin, Y. W. (2006). Promoter methylation of the secreted frizzled-related protein 1 gene SFRP1 is frequent in hepatocellular carcinoma. *Cancer*, 107(3), 579-590.
- Shine Raju, R. J., & Sehgal, S. (2018). Review of checkpoint immunotherapy for the management of non-small cell lung cancer. *ImmunoTargets and therapy*, 7, 63.
- Sholl, L. M. (2016). The Molecular Pathology of Lung Cancer. *Surgical Pathology Clinics*, 9 (3), 353 – 378.
- Shtutman, M., Zhurinsky, J., Simcha, I., Albanese, C., D'Amico, M., Pestell, R., & Ben-Ze'ev, A. (1999). The cyclin D1 gene is a target of the β -catenin/LEF-1 pathway. *Proceedings of the National Academy of Sciences*, 96(10), 5522-5527.
- Shu, J., Jelinek, J., Chang, H., Shen, L., Qin, T., Chung, W., ...& Issa, J. P. J. (2006). Silencing of bidirectional promoters by DNA methylation in tumorigenesis. *Cancer Research*, 66(10), 5077-5084.

- Sidhu, S. S., Nawroth, R., Retz, M., Lemjabbar-Alaoui, H., Dasari, V., & Basbaum, C. (2010). EMMPRIN regulates the canonical Wnt/ β -catenin signaling pathway, a potential role in accelerating lung tumorigenesis. *Oncogene*, 29(29), 4145-4156.
- Siegfried, E., Chou, T. B., & Perrimon, N. (1992). Wingless signaling acts through zeste-white 3, the Drosophila homolog of glycogen synthase kinase-3, to regulate engrailed and establish cell fate. *Cell*, 71(7), 1167-1179.
- Silverman, J. F., & Geisinger, K. R. (1996). *Fine Needle Aspiration Cytology of the Thorax And Abdomen*. WB Saunders Company.
- Simsek Kiper, P. O., Saito, H., Gori, F., Unger, S., Hesse, E., Yamana, K., ...& Boduroglu, K. (2016). Cortical-bone fragility-Insights from sFRP4 deficiency in Pyle's Disease. *New England Journal of Medicine*, 374(26), 2553-2562.
- Sjöblom, T., Jones, S., Wood, L. D., Parsons, D. W., Lin, J., Barber, T. D., ...& Szabo, S. (2006). The consensus coding sequences of human breast and colorectal cancers. *Science*, 314(5797), 268-274.
- Soda, M., Choi, Y. L., Enomoto, M., Takada, S., Yamashita, Y., Ishikawa, S., & Bando, M. (2007). Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. *Nature*, 448(7153), 561-566.
- Soria, J. C., Jang, S. J., Khuri, F. R., Hassan, K., Liu, D., Hong, W. K., & Mao, L. (2000). Overexpression of cyclin B1 in early-stage non-small cell lung cancer and its clinical implication. *Cancer Research*, 60(15), 4000-4004.
- Spiro, S. G., & Porter, J. C. (2002). Lung cancer—where are we today? Current advances in staging and nonsurgical treatment. *American Journal of Respiratory and Critical Care Medicine*, 166(9), 1166-1196.
- Srivastava, A., Sharma, K. L., Srivastava, N., Misra, S., & Mittal, B. (2012). Significant role of estrogen and progesterone receptor sequence variants in gallbladder cancer predisposition: a multi-analytical strategy. *PLoS One*, 7(7), e40162.
- Staal, F. J. (2016). Wnt signalling meets epigenetics. *Stem Cell Investigation*, 3.
- Stamos, J. L., & Weis, W. I. (2013). The β -catenin destruction complex. *Cold Spring Harbor Perspectives in Biology*, 5(1), a007898.
- Stewart, D. J. (2014). Wnt signaling pathway in non-small cell lung cancer. *Journal of the National Cancer Institute*, 106(1), djt356.

- Stewart, D. J., Chang, D. W., Ye, Y., Spitz, M., Lu, C., Shu, X., ...& Wu, X. (2014). Wnt signaling pathway pharmacogenetics in non-small cell lung cancer. *The Pharmacogenomics Journal*, 14(6), 509-522.
- Stirling, D. (2003). DNA extraction from plasma and serum. *PCR Protocols*, 63-63.
- Su, L.K., & Kinzler, K. W. (1992). Multiple intestinal neoplasia caused by a mutation in the murine homolog of the APC gene. *Science*, 256(5057), 668.
- Subramanian, J., & Govindan, R. (2007). Lung cancer in never smokers: a review. *Journal of Clinical Oncology*, 25(5), 561-570.
- Surana, R., Sikka, S., Cai, W., Shin, E. M., Warriar, S. R., Tan, H. J. G., ... & Kumar, A. P. (2014). Secreted frizzled related proteins: Implications in cancers. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1845(1), 53-65.3
- Suzuki, H., Gabrielson, E., Chen, W., Anbazhagan, R., van Engeland, M., Weijenberg, M. P., ...& Baylin, S. B. (2002). A genomic screen for genes upregulated by demethylation and histone deacetylase inhibition in human colorectal cancer. *Nature Genetics*, 31(2), 141-149.
- Suzuki, H., Toyota, M., Caraway, H., Gabrielson, E., Ohmura, T., Fujikane, T., ...& Mori, M. (2008). Frequent epigenetic inactivation of Wnt antagonist genes in breast cancer. *British Journal of Cancer*, 98(6), 1147-1156.
- Suzuki, H., Watkins, D. N., Jair, K. W., Schuebel, K. E., Markowitz, S. D., Chen, W. D., ... & Toyota, M. (2004). Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nature Genetics*, 36(4), 417-422.
- Suzuki, M., Shigematsu, H., Nakajima, T., Kubo, R., Motohashi, S., Sekine, Y., ...& Gazdar, A. F. (2007). Synchronous Alterations of Wnt and Epidermal Growth Factor Receptor Signaling Pathways through Aberrant Methylation and Mutation in Non-Small Cell Lung Cancer. *Clinical Cancer Research*, 13(20), 6087-6092.
- Szkandera, J., Pichler, M., Absenger, G., Stotz, M., Weissmueller, M., Samonigg, H., ...& Renner, W. (2014). A functional germline variant in GLI1 implicates hedgehog signaling in clinical outcome of stage II and III colon carcinoma patients. *Clinical Cancer Research*, 20(6), 1687-1697.
- Taguchi, Y., Iwadate, M., & Umeyama, H. (2016). SFRP1 is a possible candidate for epigenetic therapy in non-small cell lung cancer. *BMC Medical Genomics*, 9(1), 28.
- Takada, T., Yagi, Y., Maekita, T., Imura, M., Nakagawa, S., Tsao, S. W., ...& Ushijima, T. (2004). Methylation-associated silencing of the Wnt antagonist SFRP1 gene in human ovarian cancers. *Cancer Science*, 95(9), 741-744.

- Takeda, H., Lyle, S., Lazar, A. J., Zouboulis, C. C., Smyth, I., & Watt, F. M. (2006). Human sebaceous tumors harbor inactivating mutations in LEF1. *Nature Medicine*, 12(4), 395-397.
- Takei, H., Asamura, H., Maeshima, A., Suzuki, K., Kondo, H., Niki, T., & Matsuno, Y. (2002). Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases. *The Journal of Thoracic and Cardiovascular Surgery*, 124(2), 285-292.
- Tam, P. P., Kanai-Azuma, M., & Kanai, Y. (2003). Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. *Current Opinion in Genetics & Development*, 13(4), 393-400.
- Tamamura, Y., Otani, T., Kanatani, N., Koyama, E., Kitagaki, J., Komori, T., ...& Iwamoto, M. (2005). Developmental regulation of Wnt/ β -catenin signals is required for growth plate assembly, cartilage integrity, and endochondral ossification. *Journal of Biological Chemistry*, 280(19), 19185-19195.
- Tang, M., Torres-Lanzas, J., Lopez-Rios, F., Esteller, M., & Sanchez-Cespedes, M. (2006). Wnt signaling promoter hypermethylation distinguishes lung primary adenocarcinomas from colorectal metastasis to the lung. *International Journal of Cancer*, 119(11), 2603-2606.
- Tennis, M., Van Scoyk, M., & Winn, R. A. (2007). Role of the wnt signaling pathway and lung cancer. *Journal of Thoracic Oncology*, 2(10), 889-892.
- Teo, J. L., & Kahn, M. (2010). The Wnt signaling pathway in cellular proliferation and differentiation: a tale of two coactivators. *Advanced Drug Delivery Reviews*, 62(12), 1149-1155.
- Thakkinian, A., McElduff, P., D'Este, C., Duffy, D., & Attia, J. (2005). A method for meta-analysis of molecular association studies. *Statistics in Medicine*, 24(9), 1291-1306.
- Three-Year Report of Population Based Cancer Registries 2012-2014. (March, 2016). Retrieved October 31, 2017, from http://www.ncrpinidia.org/ALL_NCRP_REPORTS/PBCR_REPORT_2012_2014/index.htm
- Thun, M. J., Henley, S. J., Burns, D., Jemal, A., Shanks, T. G., & Calle, E. E. (2006). Lung cancer death rates in lifelong nonsmokers. *Journal of the National Cancer Institute*, 98(10), 691-699.
- Toyokawa, G., Cho, H. S., Masuda, K., Yamane, Y., Yoshimatsu, M., Hayami, S., ...& Tsunoda, T. (2011). Histone lysine methyltransferase Wolf-Hirschhorn syndrome candidate 1 is involved in human carcinogenesis through regulation of the Wnt pathway. *Neoplasia*, 13(10), 887IN1-898IN11.
- Travis, W. D., & Harris, C. C. (2004). *Pathology and Genetics of Tumours of the Lung, Pleura, Thymus and Heart*. Feance: IARC Press, 2004.

- Tseng, R. C., Lin, R. K., Wen, C. K., Tseng, C., Hsu, H. S., Hsu, W. H., & Wang, Y. C. (2008). Epigenetic silencing of AXIN2/betaTrCP and deregulation of p53-mediated control lead to wild-type β -catenin nuclear accumulation in lung tumorigenesis. *Oncogene*, 27(32), 4488-4496.
- Tsuji, T., Nozaki, I., Miyazaki, M., Sakaguchi, M., Pu, H., Hamazaki, Y., ...& Namba, M. (2001). Antiproliferative activity of REIC/Dkk-3 and its significant down-regulation in non-small-cell lung carcinomas. *Biochemical and Biophysical Research Communications*, 289(1), 257-263.
- Ueda, M., Gemmill, R. M., West, J., Winn, R., Sugita, M., Tanaka, N., ...& Drabkin, H. A. (2001). Mutations of the beta-and gamma-catenin genes are uncommon in human lung, breast, kidney, cervical and ovarian carcinomas. *British Journal of Cancer*, 85(1), 64-68.
- Urakami, S., Shiina, H., Enokida, H., Hirata, H., Kawamoto, K., Kawakami, T., ...& Igawa, M. (2006). Wnt antagonist family genes as biomarkers for diagnosis, staging, and prognosis of renal cell carcinoma using tumor and serum DNA. *Clinical Cancer Research*, 12(23), 6989-6997.
- van der Meide, W. F., Snellenberg, S., Meijer, C. J., Baalbergen, A., Helmerhorst, T. J., van der Sluis, W. B., ... & Steenbergen, R. D. (2011). Promoter methylation analysis of WNT/ β -catenin signaling pathway regulators to detect adenocarcinoma or its precursor lesion of the cervix. *Gynecologic oncology*, 123(1), 116-122.
- Vaughan, A. E., Halbert, C. L., Wootton, S. K., & Miller, D. (2011). Lung cancer in mice induced by the jaagsiekte sheep retrovirus envelope protein is not maintained by rare cancer stem cells, but tumorigenicity does correlate with Wnt pathway activation. *Molecular Cancer Research*, molcanres-0285.
- Veeck, J., & Dahl, E. (2012). Targeting the Wnt pathway in cancer: the emerging role of Dickkopf-3. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, 1825(1), 18-28.
- Veeck, J., Niederacher, D., An, H., Klopocki, E., Wiesmann, F., Betz, B., ...& Huszka, C. (2006). Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. *Oncogene*, 25(24), 3479-3488.
- Veeck, J., Noetzel, E., Bektas, N., Jost, E., Hartmann, A., Knüchel, R., & Dahl, E. (2008). Promoter hypermethylation of the SFRP2 gene is a high-frequent alteration and tumor-specific epigenetic marker in human breast cancer. *Molecular Cancer*, 7(1), 83.
- Veeck, J., Wild, P. J., Fuchs, T., Schüffler, P. J., Hartmann, A., Knüchel, R., & Dahl, E. (2009). Prognostic relevance of Wnt-inhibitory factor-1 (WIF1) and Dickkopf-3 (DKK3) promoter methylation in human breast cancer. *BMC Cancer*, 9(1), 217.
- Veeman, M. T., Axelrod, J. D., & Moon, R. T. (2003). A second canon: functions and mechanisms of β -catenin-independent Wnt signaling. *Developmental Cell*, 5(3), 367-377.

- Vermeulen, L., Felipe De Sousa, E. M., Van Der Heijden, M., Cameron, K., De Jong, J. H., Borovski, T., ...& Sprick, M. R. (2010). Wnt activity defines colon cancer stem cells and is regulated by the microenvironment. *Nature Cell Biology*, 12(5), 468-476.
- Vincent, A., Omura, N., Hong, S. M., Jaffe, A., Eshleman, J. R., & Goggins, M. (2011). Genome-wide analysis of promoter methylation associated with gene expression profiles of pancreatic adenocarcinomas. *Clinical Cancer Research*, clincanres-3431.
- Walter, J. B., & Pryce, D. M. (1955). The histology of lung cancer. *Thorax*, 10(2), 107.
- Wang, C., Yue, Y., Shao, B., Qiu, Z., Mu, J., Tang, J., ...& Ren, G. (2017). Dickkopf-Related Protein 2 is Epigenetically Inactivated and Suppresses Colorectal Cancer Growth and Tumor Metastasis by Antagonizing Wnt/ β -Catenin Signaling. *Cellular Physiology and Biochemistry*, 41(5), 1709-1724.
- Wang, H., Duan, X. L., Qi, X. L., Meng, L., Xu, Y. S., Wu, T., & Dai, P. G. (2017). Concurrent Hypermethylation of SFRP2 and DKK2 Activates the Wnt/ β -Catenin Pathway and Is Associated with Poor Prognosis in Patients with Gastric Cancer. *Molecules and Cells*, 40(1), 45.
- Wang, S., Krinks, M., Lin, K., Luyten, F. P., & Moos, M. (1997). Frzb, a secreted protein expressed in the Spemann organizer, binds and inhibits Wnt-8. *Cell*, 88(6), 757-766.
- Wang, Y., Macke, J. P., Abella, B. S., Andreasson, K., Worley, P., Gilbert, D. J., ...& Nathans, J. (1996). A large family of putative transmembrane receptors homologous to the product of the *Drosophila* tissue polarity gene *frizzled*. *Journal of Biological Chemistry*, 271(8), 4468-4476.
- Waterman, M. L., Fischer, W. H., & Jones, K. A. (1991). A thymus-specific member of the HMG protein family regulates the human T cell receptor C α enhancer. *Genes & Development*, 5(4), 656-669.
- Weiss. (2010, April 5). An Introduction to Lung Cancer. Retrieved October 31, 2017, from <http://cancergrace.org/lung/2010/04/05/an-introduction-to-lung-cancer/>
- Weissferdt, A., & Moran, C. A. (2012). *Diagnostic Pathology of Pleuropulmonary Neoplasia*. Springer Science & Business Media.
- Willert, K., & Jones, K. A. (2006). Wnt signaling: is the party in the nucleus? *Genes & Development*, 20(11), 1394-1404.
- Willert, K., Brown, J. D., Danenberg, E., Duncan, A. W., Weissman, I. L., Reya, T., ...& Nusse, R. (2003). Wnt proteins are lipid-modified and can act as stem cell growth factors. *Nature*, 423(6938), 448-452.
- Willert, K., Shibamoto, S., & Nusse, R. (1999). Wnt-induced dephosphorylation of axin releases β -catenin from the axin complex. *Genes & Development*, 13(14), 1768-1773.

- Wils, L. J., & Bijlsma, M. F. (2017). Epigenetic regulation of the Hedgehog and Wnt pathways in cancer. *Critical reviews in oncology/hematology*.
- Winn, R. A., Bremnes, R. M., Bemis, L., Franklin, W. A., Miller, Y. E., Cool, C., & Heasley, L. E. (2002). [gamma]-Catenin expression is reduced or absent in a subset of human lung cancers and re-expression inhibits transformed cell growth. *Oncogene*, 21(49), 7497.
- Wissmann, C., Wild, P. J., Kaiser, S., Roepcke, S., Stoehr, R., Woenckhaus, M., ...& Knuechel, R. (2003). WIF1, a component of the Wnt pathway, is down-regulated in prostate, breast, lung, and bladder cancer. *The Journal of Pathology*, 201(2), 204-212.
- Wistuba, I. I., Behrens, C., Virmani, A. K., Mele, G., Milchgrub, S., Girard, L., & Lerman, M. I. (2000). High resolution chromosome 3p allelotyping of human lung cancer and preneoplastic/preinvasive bronchial epithelium reveals multiple, discontinuous sites of 3p allele loss and three regions of frequent breakpoints. *Cancer Research*, 60(7), 1949-1960.
- Wodarz, A., & Nusse, R. (1998). Mechanisms of Wnt signaling in development. *Annual Review of Cell and Developmental Biology*, 14(1), 59-88.
- Wogan, G. N., Hecht, S. S., Felton, J. S., Conney, A. H., & Loeb, L. A. (2004, December). Environmental and chemical carcinogenesis. In *Seminars in Cancer Biology* (Vol. 14, No. 6, pp. 473-486). Academic Press.
- Wu, K., House, L., Liu, W., & Cho, W. (2012). Personalized targeted therapy for lung cancer. *International Journal of Molecular Sciences*, 13(9), 11471-11496.
- Wu, X., Gu, J., Grossman, H. B., Amos, C. I., Etzel, C., Huang, M., ...& Spitz, M. R. (2006). Bladder cancer predisposition: a multigenic approach to DNA-repair and cell-cycle-control genes. *The American Journal of Human Genetics*, 78(3), 464-479.
- Wu, X., Zhao, H., Suk, R., & Christiani, D. C. (2004). Genetic susceptibility to tobacco-related cancer. *Oncogene*, 23(38), 6500-6523.
- Wu, Z., Sun, Y., Tang, S., Liu, C., Zhu, S., Wei, L., & Xu, H. (2015). AXIN 2 rs2240308 polymorphism contributes to increased cancer risk: evidence based on a meta-analysis. *Cancer Cell International*, 15(1), 68.
- Xiang, T., Li, L., Yin, X., Zhong, L., Peng, W., Qiu, Z., ...& Tao, Q. (2013). Epigenetic silencing of the WNT antagonist Dickkopf 3 disrupts normal Wnt/ β -catenin signalling and apoptosis regulation in breast cancer cells. *Journal of Cellular and Molecular Medicine*, 17(10), 1236-1246.
- Xu, Q., Wang, Y., Dabdoub, A., Smallwood, P. M., Williams, J., Woods, C., ...& Nathans, J. (2004). Vascular development in the retina and inner ear: control by Norrin and Frizzled-4, a high-affinity ligand-receptor pair. *Cell*, 116(6), 883-895.

- Yadav, A., Gupta, A., Yadav, S., Rastogi, N., Agrawal, S., Kumar, A., ...& Mittal, B. (2016). Association of Wnt signaling pathway genetic variants in gallbladder cancer susceptibility and survival. *Tumor Biology*, 37(6), 8083-8095.
- Yamaguchi, T. P. (2001). Heads or tails: Wnts and anterior-posterior patterning. *Current Biology*, 11(17), R713-R724.
- Yamamoto, H., Hinoi, T., Michiue, T., Fukui, A., Usui, H., Janssens, V., ...& Kikuchi, A. (2001). Inhibition of the Wnt signaling pathway by the PR61 subunit of protein phosphatase 2A. *Journal of Biological Chemistry*, 276(29), 26875-26882.
- Yang, B., Du, Z., Gao, Y. T., Lou, C., Zhang, S. G., Bai, T., ...& Song, W. Q. (2010). Methylation of Dickkopf-3 as a prognostic factor in cirrhosis-related hepatocellular carcinoma. *World Journal of Gastroenterology: WJG*, 16(6), 755.
- Yang, K., Wang, X., Zhang, H., Wang, Z., Nan, G., Li, Y., ...& Bi, Y. (2016). The evolving roles of canonical WNT signaling in stem cells and tumorigenesis: implications in targeted cancer therapies. *Laboratory Investigation*, 96(2), 116-136.
- Yang, Y. (2012). Wnt signaling in development and disease. *Cell & Bioscience*, 2(1), 14.
- Yardy, G. W., & Brewster, S. F. (2005). Wnt signalling and prostate cancer. *Prostate Cancer and Prostatic Diseases*, 8(2), 119-126.
- Yilmaz, M., Donmez, G., Kacan, T., Sari, I., Babacan, N. A., Sari, M., & Kilickap, S. (2015). Significant association between polymorphisms of Wnt antagonist genes and lung cancer. *Journal of Investigative Medicine*, 63(8), 935-941.
- Yin, D. T., Wu, W., Li, M., Wang, Q. E., Li, H., Wang, Y., ...& Xing, M. (2013). DKK3 is a potential tumor suppressor gene in papillary thyroid carcinoma. *Endocrine-Related Cancer*, 20(4), 507-514.
- Ying, J., Li, H., Chen, Y. W., Srivastava, G., Gao, Z., & Tao, Q. (2007). WNT5A is epigenetically silenced in hematologic malignancies and inhibits leukemia cell growth as a tumor suppressor. *Blood*, 110 (12), 4130-4131.
- Ying, Y., & Tao, Q. (2009). Epigenetic disruption of the WNT/ β -catenin signaling pathway in human cancers. *Epigenetics*, 4(5), 307-312.
- Yo, S. S., Hong, M. J., Choi, J. E., Lee, J. H., Baek, S. A., Lee, W. K., ... & Kim, C. H. (2016). Genetic Variants in the Wnt Signaling Pathway Are Not Associated with Survival Outcome of Non-Small Cell Lung Cancer in a Korean Population. *Journal of Korean Medical Science*, 31(3), 463-466.

- Yoder, L. H. (2006). An overview of lung cancer symptoms, pathophysiology, and treatment. *Medsurg Nursing*, 15(4), 231.
- Yokota, J., Shiraishi, K., & Kohno, T. (2010). Genetic basis for susceptibility to lung cancer: Recent progress and future directions. *Advances in Cancer Research*, 109, 52.
- Yu, J., Tao, Q., Cheng, Y. Y., Lee, K. Y., Ng, S. S., Cheung, K. F., ...& Ebert, M. (2009). Promoter methylation of the Wnt/ β -catenin signaling antagonist Dkk-3 is associated with poor survival in gastric cancer. *Cancer*, 115(1), 49-60.
- Yue, W., Sun, Q., Dacic, S., Landreneau, R. J., Siegfried, J. M., Yu, J., & Zhang, L. (2008). Downregulation of Dkk3 activates β -catenin/TCF-4 signaling in lung cancer. *Carcinogenesis*, 29(1), 84-92.
- Zeng, L., Fagotto, F., Zhang, T., Hsu, W., Vasicek, T. J., Perry, W. L., ...& Costantini, F. (1997). The mouse Fusedlocus encodes Axin, an inhibitor of the Wnt signaling pathway that regulates embryonic axis formation. *Cell*, 90(1), 181-192.
- Zerlin, M., Julius, M. A., & Kitajewski, J. (2008). Wnt/Frizzled signaling in angiogenesis. *Angiogenesis*, 11(1), 63-69.
- Zhang, X., Gaspard, J. P., & Chung, D. C. (2001). Regulation of vascular endothelial growth factor by the Wnt and K-ras pathways in colonic neoplasia. *Cancer Research*, 61(16), 6050-6054.
- Zhang, Y. W., Miao, Y. F., Yi, J., Geng, J., Wang, R., & Chen, L. B. (2009). Transcriptional inactivation of secreted frizzled-related protein 1 by promoter hypermethylation as a potential biomarker for non-small cell lung cancer. *Neoplasia*, 57(3), 228.
- Zhang, Y., Wang, R., Song, H., Huang, G., Yi, J., Zheng, Y., ...& Chen, L. (2011). Methylation of multiple genes as a candidate biomarker in non-small cell lung cancer. *Cancer Letters*, 303(1), 21-28.
- Zhu, J., Zhang, S., Gu, L., & Di, W. (2012). Epigenetic silencing of DKK2 and Wnt signal pathway components in human ovarian carcinoma. *Carcinogenesis*, 33(12), 2334-2343.
- Zi, X., Guo, Y., Simoneau, A. R., Hope, C., Xie, J., Holcombe, R. F., & Hoang, B. H. (2005). Expression of Frzb/secreted Frizzled-related protein 3, a secreted Wnt antagonist, in human androgen-independent prostate cancer PC-3 cells suppresses tumor growth and cellular invasiveness. *Cancer Research*, 65(21), 9762-9770.
- Zöchbauer-Müller, S., Gazdar, A. F., & Minna, J. D. (2002). Molecular pathogenesis of lung cancer. *Annual Review of Physiology*, 64(1), 681-708.
- Zorn, A. M. (2001). Wnt signalling: antagonistic Dickkopfs. *Current Biology*, 11(15), R592-R595.

APPENDIX-I

PATIENT CONSENT FORM

Name of the participant: _____

Name of the Principal (Co-) Investigators: _____

Name of the Institution: _____

Name and address of the sponsoring (funding) agency(ies): _____

Documentation of the informed consent

I, ... have read the information in this form (or it has been read to me). I was free to ask any questions and they have been answered. I am over 18 years of age and, exercising my free power of choice, hereby give my consent to be included as a participant in the project “**Association of Single Nucleotide Polymorphisms and Methylation status in the Wnt and AhR signaling pathways with risk for occurrence of Lung Cancer in a North Indian Population**”

(1) I have read and understood consent form and the information provided to me.

(2) I have had the consent document explained to me.

(3) I have been explained about the nature of the study.

(4) My rights and responsibilities have been explained to me by the investigator.

(5) I have the right to not be a subject in the above study and am not willing to give sample

(6) I have informed the investigator of all the treatments I am taking or have taken in the past months including any *desi* (alternative) treatments.

(7) Patient data and confidentiality has been explained to me.

(8) I hereby give permission to the investigators to release the information obtained from me as result of participation in this study to the sponsors, regulatory authorities, Government agencies, and ethics committee. I understand that they may inspect my original records.

(9) My identity will be kept confidential if my data are publicly presented.

(10) I have had my questions answered to my satisfaction.

(11) I have decided to be in the research study.

I am aware, that if I have any questions during this study, I should contact the concerned authorities. By signing this consent from, I attest that the information given in this document. I will be given a copy of this consent document.

For adult participants

Name and signature / thumb impression of the participant (or legal representative if participant incompetent):

_____ (Name) _____ (Signature)

Date: _____ Time: _____

Name and signature of impartial witness (required for illiterate patients):

_____ (Name) _____ (Signature)

Date: _____ Time: _____

Address and contact number of the impartial witness: _____

Name and signature of the Investigator or his representative obtaining consent:

_____ (Name) _____ (Signature)

_____ (Date)

Investigator Certificate

I certify that all the elements including the nature of the research study and possible risks of the above study as described in this consent document have been fully explained to the subject. In my judgment, the participant possesses the legal capacity to give informed consent to participate in this research and is voluntarily and knowingly giving informed consent to participate.

Signature of the Investigator: _____ Dated: _____

Name of the Investigator: _____

Appendix –II
PATIENT INFORMATION SHEET

Department of Biotechnology, Thapar University
&
Dept. of Pulmonary Medicine, PGIMER, Chandigarh

Investigator: Siddharth Sharma, Ph.D

Co-Investigator: Dr. Aushtosh Aggarwal, M.D, D.M

Co-Investigator: Dr. Navneet Singh, M.D, D.M

Name of Patient _____

Address (Rural/Urban) _____

C.R Number _____

Date of Sampling _____

Type of Cancer _____

Stage of Cancer _____

Sex _____

Age _____

Education (Years) _____

Occupation _____

Marital Status _____

Smoking Status

1. Smoker or Non-Smoker _____

2. Type of Smoking

a) Cigarette b) Beedi c) Hookah d) Cigarette+Beedi e) Beedi+Hookah f) Any other form of smoke

3. No. of Cigarettes/ Beedis smoked

a) Per day _____

b) Per Week _____

c) Amount of tobacco added to hookah _____

4. a) Age at which started Smoking _____

b) How many years have you smoked _____

5. Time taken to finish cigarette (in min) _____

6. Brand of smoke preferred _____

7. Do you chew tobacco/betelnut

a) Yes b) No

If yes amount chewed _____

8. Any familial case of Smoking related cancer in your family: a) Yes b) No

9. Are you only one to smoke in your family

Alcohol Consumption

1. Do you consume alcohol

a) Yes b) No

2. Intake of Alcohol

a) Daily b) 3-5 times a week c) 1-2 times a week d) once a week e) once a month

3. How much alcohol do you consume in ml _____

4. What type of alcohol do you consume _____

INSTITUTE ETHICS COMMITTEE
Postgraduate Institute of Medical Education and Research, Chandigarh

Prof. O.N. Bhakoo
Chairman



Prof. Dheeraj Gupta
Convener

No: PGI/IEC/2011/ 557

Date: 25/11/2011

CHAIRMAN

Prof. O. N. Bhakoo

MEMBERS

Justice P. K. Palli /
Justice M. R. Agnihotri

Dr. B. N. Goswami

Dr. Girish Varshney

Dr. Promila Vasudeva

Prof. Savita Malhotra

Prof. R. Kochhar

Prof. Anil Bhansali

Prof. Vanita Jain

Prof. R. K. Dhiman

Prof. Ashim Das

Dr. Jayashree M.

Prof. Dheeraj Gupta
(Convener)

Dr. Navneet Singh,
Assistant Professor,
Dept. of Pulmonary Medicine,
PGIMER, Chandigarh.

Title: Association of Single Nucleotide Polymorphisms and Methylation status in the Wnt and AhR signaling pathways with risk for occurrence of Lung Cancer in a North Indian Population.

Project Number: P-106

Dear Dr. Singh,

We received five copies of your above mentioned project proposal, which has been assigned the project number **P-106**. The project was discussed in the meeting of the Institute Ethics committee held on 23.11.2011. After due consideration and discussion, the committee has sought following clarifications:

"Since the samples are to be processed and banked at TIET, Patiala, the host institute should give a clear undertaking that "stored DNA will be available to PGI investigators as and when they need it and no further investigations will be carried out by the PI without concurrence and consent from PGI team."

You are requested to submit your response to the above comment/s along with two copies of the revised document/s at the earliest for further necessary action.

All future correspondence with respect to this project must mention **the project number**.

Kind Regards

Yours sincerely,

(Dheeraj Gupta)

CONTACT

Phone:

+91 172 2756820

Email:

ecpgi@gmail.com

APPENDIX-III

1. 0.5M EDTA: Dissolved 9.306g of disodium salt of EDTA in 20ml of deionised water, and then adjusted the pH to 8.0 by 1 M sodium hydroxide. Sterilized the solution by autoclaving.
2. 10% SDS: Dissolved 1g of SDS in 10ml of deionised water.
3. 100mM Tris-Cl (pH 8.0): Dissolved 0.32g of Tris-Cl in 10 ml of deionised water, then adjusted the pH to 8.0 by 1M sodium hydroxide. Sterilized the solution by autoclaving.
4. 10mg/ml Proteinase K: Dissolved 10mg Proteinase K in 1ml of double distilled water. Sterilized the solution by autoclaving.
5. 1mg/ml BSA: Dissolved 100mg of BSA in 100ml of deionised sterile water and kept at 4 C overnight.
6. 5M Sodium chloride (NaCl): Dissolved 5.85g of sodium chloride in 20ml of deionised water. Sterilized the solution by autoclaving.
7. 5X TBE buffer: Dissolved 54g of Tris base and 27.5g of boric acid in 980ml of double distilled water and then added 20ml of 0.5 EDTA. Sterilized the solution by autoclaving.
8. Ethidium Bromide (10mg/ml): Dissolved 1g of ethidium bromide in 100ml of water. Mixed the solution properly.
9. Magnesium chloride (MgCl₂) (100mM): Dissolved 0.41gms of MgCl₂ in 20ml of deionised water and sterilized by autoclaving.
10. Sucrose (1M): Dissolved 3.41 g of sucrose in 10 ml of deionised water and sterilized by autoclaving.
11. TE buffer (pH 8.0): Added 1ml of 100mM Tris-Cl (pH 8.0) and 200 µl of 0.5M EDTA solution to 8.8 ml of deionised water. Sterilized the solution.
12. Triton X- 100 (10%): Took 100 µl of TritonX-100 and mixed with 900 µl of deionised water and mixed properly.



High-order gene interactions between the genetic polymorphisms in Wnt and AhR pathway in modulating lung cancer susceptibility

Charu Bahl¹, Navneet Singh², Digambar Behera² & Siddharth Sharma^{*,1}

¹Department of Biotechnology, Thapar University, Patiala, Punjab 147002, India

²Department of Pulmonary Medicine, Post-Graduate Institute of Education & Medical Research (PGIMER), Sector 14, Chandigarh, India

* Author for correspondence: Tel.: +91 9501688366; siddharthsharma.phd@thapar.edu

Aim: Genetic variations present within Wnt and AhR pathway might be related to the lung cancer susceptibility. **Methods:** A total of 555 subjects were genotyped using PCR-RFLP technique for polymorphic sites in *DKK4*, *DKK3*, *DKK2*, *sFRP3*, *sFRP4*, *Axin2* and *AhR*. Multifactor dimensionality reduction method and classification and regression tree analysis was used. **Results:** Overall *sFRP4*_{rs1802073} which has a cross validation consistency of 10/10, prediction error = 0.43 ($p > 0.0001$) is the best factor model. The second best model was *sFRP4*_{rs1802073} and *DKK2*_{rs419558} with cross validation consistency of 9/10 and prediction error = 0.40. In classification and regression tree analysis, *DKK2*_{rs419558} came out to be a significant factor; *DKK2*_{rs17037102} (M)/*DKK2*_{rs419558} (M) showed a tenfold risk of acquiring lung cancer, $p = 0.0001$. *DKK2*_{rs17037102} (M)/*AhR*_{rs2066853} (W)/*AhR*_{rs10250822} (M) showed an 11-fold risk of developing lung cancer, $p = 0.00001$. **Conclusion:** Both *DKK2* and *sFRP4* polymorphisms are found to play a crucial role; especially for smokers towards modulating risk for lung cancer. *AhR* variants are contributing maximally toward lung cancer risk.

First draft submitted: 23 February 2017; Accepted for publication: 25 July 2017; Published online: 28 November 2017

Keywords: aryl hydrocarbon receptor • interaction • lung cancer • polymorphisms • risk • Wnt

Interplay of numerous genetic variabilities within the signaling pathways at multiple loci and environmental factors synergistically modulate the propensity of occurrence of a disease. Genetic susceptibility toward a disease has been recently accounted as an important contributing factor in the process of carcinogenesis [1]. Predisposition toward a disease such as cancer cannot be dramatically effected by a single predictor that might not show promising results when analyzed alone [2]. Several individual genetic variants have given inconsistent outcomes in the various studies pertaining to the contradiction about their involvement in development of the disease. Various factors which bring these differences in results include smaller sample size, variation in statistical models and lack of evaluation of interaction among the genes involved in same physiological pathway [3]. Etiology of lung cancer (LC) has been elucidated by distinct and complex signaling pathways and their interaction with environmental factors. Multiple pathways and their interaction with other environmental parameters might unfold the cumulative effect of low penetrance alleles in modulating the susceptibility [4].

Canonical Wnt signaling cascade is a ligand-induced β -catenin-mediated diverse pathway which has been implicated in the various cell growth-regulating functions [5]. It forms basic architecture for development of embryos and maintenance of tissues in adults [6]. The downstream targets of this pathway include matrix metalloproteins, *cox-2*, cyclin D1, VEGF, *c-myc* among others. Regulation of this centrally important pathway is crucial and is checked at multiple cellular stages [7]. The tight regulatory machinery of the Wnt pathway has several major antagonists such as *sFRP*, *DKK* protein and *Axin2* protein [8]. *sFRP* proteins are extracellular protein which exhibit a structural homology with the frizzled-related receptors which is responsible for negatively regulating signal cascade by competing with the frizzled-related receptors for binding to the Wnt ligands which activate this pathway [9]. *DKK* proteins are secreted glycoproteins involved in the process of antagonizing of Wnt pathway at the next cellular

stage where it binds with the LRP co-receptor and due to this interaction there is disruption of Wnt ligand/frizzled receptor/LRP complex [10]. Axin2 being a scaffold protein forms the important part of the destruction complex along with dishevelled and adenomatous polyposis coli, which mediate the phosphorylation of β -catenin by GSK-3 β required for its ubiquitination so as to inhibit its translocation into nucleus to bind to transcription factors TCF/LEF regulating downstream target gene expression [11]. Germline variations in the genes of this pathway may lead to its constitutive expression. Chronic expression of this intermeshing transduction cascade has been reported in various human cancers such as colorectal, head and neck cancer, melanoma, leukemia and LC [12]. Evidently many studies have evaluated the role of genetic differences in affecting disease susceptibility for Wnt antagonists genes namely *sFRP3* [13–15], *sFRP4* [16–18], *DKK2* [14,17,18], *DKK3* [14,17–20], *DKK4* [17,18] and *Axin2* [21–27].

There have been very few reports that have attempted to intersect the Aryl hydrocarbon receptor (AhR) signaling pathway with Wnt cascade. AhR is a ligand-induced transcription factor, which is translocated into the nucleus leaving behind the chaperone proteins that bind to it in the cytoplasm. In the nucleus, it interacts with AhR nuclear translocator forming a heterodimer. This heterodimer has high affinity for dioxin response elements of the promoter region of target genes involved in phase I detoxification process such as *CYP1A1*, *CYP1A2* and *CYP1B1* among others. AhR cater to the wide variety of carcinogens as ligands namely poly aromatic hydrocarbons like benzo(a)pyrene and nitrosamines present in tobacco smoke [28–30]. It upregulates the *CYP1A1* activity when the cells are exposed to carcinogen metabolism. AhR is a major determinant in the process of smoke-driven lung carcinoma. Also, AhR has been found to be associated with other cellular processes such as cell proliferation, NF- κ B-induced inflammation and DNA adduct formation [31]. All these processes are playing a relevant part in the etiology of LC [32]. Few studies have focused on the sequence variations of *AhR* and their role in affecting cancer predisposition and prognosis [33].

A crosstalk between AhR and Wnt signaling is a rarely explored concept. The expression of canonical β -catenin pathway is majorly defined by the stability of β -catenin present in the nucleus, which is not governed by a single transduction network [34]. It is maintained by various upstream and downstream signals of proteins determining its fate in the cytoplasm. Also, another aspect is that altered expression of β -catenin cannot alone determine the status of downstream target genes, which have elucidated roles in oncogenesis [35,36]. A new dimension has been added to this intermeshing network, which is mediated by β -catenin, wherein AhR has been enumerated as one of the key protein which can alter this signal cascade and affect the β -catenin levels in the cytoplasm through other pathways [37]. β -catenin can augment AhR-driven expression by increasing the activity of AhR protein at hotspots in promoter sites called dioxin response elements [38,39].

In lieu of the above cited facts, the present study was designed to explore the high order gene–gene and gene–environment interactions that may cater to the difference in LC susceptibility in north Indians. Current study evaluates the role of several genetic variations with the Wnt antagonist genes and AhR pathway. The following SNPs are studied for assessing their association with LC risk, *sFRP3* (rs7775 and rs228236), *sFRP4* (rs1802073 and rs1802074), *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK2* (rs447372, rs419558 and rs17037102), *DKK4* (rs2073664), *Axin2* (148 C > T, 1365 G > A, 432 T > C, 956 + 16 A > G, 1386 C > T and 2062 C > T) and *AhR* (rs7811989, rs10250822, rs2282885 and rs2066853). Multifactor dimensionality reduction and classification and regression tree (CART) analysis were used as major tools to find out cumulative effects of all these polymorphic sites. The interlinking of the Wnt and AhR pathway is also studied by looking at their collaborative effect on LC risk.

Materials & methods

Study population

Current study recruited 292 cases and 263 controls from the Department of Pulmonary Medicine, Post Graduate Institute of Medical Education & Research, Chandigarh. A written informed consent was taken from each volunteer prior to blood collection. This study has been reviewed and approved by Ethics committee of Post Graduate Institute of Medical Education & Research. A questionnaire having all the details about various epidemiological factors was filled by a trained personnel during the recruitment process. The only exclusion criterion for LC patients was that they should not have any previous history of any carcinoma, otherwise there were no age, sex and tumor node metastasis (TNM) restrictions. The controls were healthy people who visited the hospital for health checkups. A major attempt was to avoid sampling bias, which may occur due to difference in age, sex and smoking status of the cases and controls. Smoking was quantified using pack years, which is calculated by this formula: (cigarettes

or beedis [type of Indian cigarette] per day/20) $\times$ number of years smoked. The other clinical details including histology, stage and TNM were obtained from the medical records of the patients in the hospital.

DNA isolation

Blood collected from each subject was 5 ml and it was used to isolate the genomic DNA using the protocol given by Sodhi *et al.* [40]. The isolated DNA was quantified using Nanodrop and stored at -4°C for further use.

Genotyping of *sFRP*, *DKK*, *Axin2* & *AhR* genetic variants

Genotyping of all the polymorphic sites under study was done using PCR–RFLP. The genotyping for two *sFRP3* gene variants was done as previously detailed by Shanmugam *et al.* [13]. For *sFRP4* variants (rs1802073 and rs1802074), *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK2* (rs447372, rs419558 and rs17037102) and *DKK4* (rs2073664) polymorphic sites, the genotyping was carried in a similar manner as reported by Hirata *et al.* [17]. In case of genetic variants of *Axin2* gene namely 148 C > T, 1365 G > A, 432 T > C, 956 + 16 A > G, 1386 C > T and 2062 C > T, the protocol described by Pinarbasi and colleagues [22] was applied to find out the genotype of the subjects. The genotyping of four AhR variants (rs7811989, rs10250822, rs2282885 and rs2066853) was done by PCR–RFLP using specific primer sequences and restriction enzymes as described previously by Bin *et al.* [41]. The PCR reaction (25 μl) used to amplify the desired fragment comprised of 1 $\times$ PCR buffer, 1.5 mM MgCl_2 , with 0.5 μM of both forward and reverse primer, 200 μM of each dNTP's, 100 $\mu\text{g}/\text{ml}$ bovine serum albumin and 1 U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific, MA, USA) and approximately 200 ng DNA. The amplified products were digested with their respective restriction enzymes as described above. The digested products were resolved on either agarose gel or polyacrylamide gel to find out the restriction patterns. Scoring of the patterns is done to find out the genotypic status of the sample. The genotyping of 15% samples was done twice in order to check the reproducibility of the results and it was 100%.

Statistical analysis

Demographic parameters included age, gender, smoking status and pack years. A *t*-test was used to compare the continuous variables such as age and pack years among cases and controls, on the other hand for categorical variables such as gender and smoking status χ^2 test was applied. Unconditional multivariate logistic regression was used to find adjusted odds ratio (OR) along with 95% CI associated with each genotype and various genotypic combinations. The OR was adjusted for age, gender and smoking status as these might act as confounding factors [40]. The major goal of the present study was to uncover the hidden interactions taking place in the presence of multiple factors, which cannot be determined by traditional statistical methods. Nonparametric approach was applied by using multifactor dimensionality reduction method in order to analyze various gene–gene (Wnt antagonists and AhR) and gene–environment interactions (Wnt antagonist, AhR with smoking as an environmental exposure) that contribute maximally to the scenario of LC predisposition. Multifactor dimensionality reduction (MDR) method involves the reduction of data to detect the multi-loci genotypic combinations that can help in predicting the risk for a complex disease such as LC. MDR reduces the multidimensional data into one dimension by pooling the genotypes into high and low-risk groups. The interaction models are then assessed on the basis of cross validation consistency (CVC), which is number of times a model is identified as the best model across the cross-validation sets. Higher CVC signifies higher support for the strength of the model. Average prediction error is also calculated (1-testing accuracy). Other important parameter is permutation testing (p-value), which determines the significance of the hypothesis generated [42]. To combat the effect of confounding factors in this interaction analysis, the author has employed the approach of doing stratified analysis on the basis of factors such as histology and smoking as this is the only way of overcoming this major drawback of MDR approach. The theoretical power of MDR in the stratified analysis when calculated for different histological subtypes came out to be nearly 80% and in case of smoker and nonsmoker category it came out to be 70%. Hence, the power of these statistical tests was ranging from 70 to 80%. The software used to apply MDR is version 0.5.1 of the open-source MDR software package that is available online [61]. The next important analysis which was carried out to find out high-order logistic regression complex interactions was CART using the CART software (6.0, Salford Systems, CA, USA). It is a recursive binary partitioning approach which divides the data subsequently on the basis of the risk associated with them and forms a decision tree depicting all the high- and low-risk subgroups. The most significant factor which contributes in disease susceptibility forms the first split in the tree and subsequent splits are made on the basis of the significance levels in order to control the tree growth. The tree is classified into nodes and terminal nodes. The splitting process

Table 1. Demographic characteristics among cases and controls.

Variable	Cases, n = 292 (%)	Controls, n = 263 (%)	p-value [†]
Age (years):			
– Mean $\pm$ SD	57.38 $\pm$ 10.74	53.23 $\pm$ 10.44	<0.0001
– Range	–	–	–
Gender:			
– Male	252 (86)	241 (92)	0.06
– Female	40 (14)	22 (8)	–
Smoking status:			
– Smokers	229 (78)	197 (74)	0.37
– Nonsmokers	63 (22)	66 (26)	–
– Pack years, mean $\pm$ SD	28.09 $\pm$ 34.09	19.4 $\pm$ 19.56	0.0003
Histological types:			
– SQCC	124 (42.4)	–	–
– ADCC	95 (32.5)	–	–
– SCLC	68 (23.3)	–	–
– Others	4 (1.4)	–	–
– Unknown	1 (0.3)	–	–

[†]p-values were derived from Pearson χ^2 test except age; student's t-test was used for age. All p-values are two sided. $p < 0.05$ was considered statistically significant.
n: Total number of case patients or control subjects.

ADCC: Adenocarcinoma; SCLC: Small-cell lung cancer; SD: Standard deviation; SQCC: Squamous cell carcinoma.

continues until the terminal nodes have no subsequent statistically significant values or very less number of subjects. This aids in the estimation of different genotypic combinations affecting LC susceptibility which are not obtained by traditional logistic regression. It takes into account a very large number of variables at a time, and find the high-risk subgroups. This data mining exercise yields a decision tree-like structure which depicts various factors and their interaction with each other along with the risk associated with these combinations. The nodes present at the initial splits in the tree formation are biologically meaningful in modulating the LC predisposition. Terminal node with a lowest case rate is used as a reference to estimate the OR and 95% CI for all genotypes depicted in other nodes. CART employs the use of easy to compute, Gini index. The impurity (or purity) measure used in building decision tree in CART is Gini Index. Gini impurity is a measure of how often a randomly chosen element from the set would be incorrectly labeled if it was randomly labeled according to the distribution of labels in the subset [43].

Results

Demographic characteristics

Study population consists of 292 cases with a mean age of 57.38 ± 10.74 years and 263 controls with a mean age of 53.23 ± 10.44 years. A maximum number of study participants were males (86.3% of cases and 91% of controls). An attempt was made to match cases and controls in terms of gender during sampling; however, due to difference in the sample size of cases and controls the males were over-represented in cases. The percentages of smokers were almost comparable (78.4 and 74.9%) among both the study groups, which indicate adequate matching between the cases and controls in terms of smoking habits. All other clinical details such as stage, TNM and histological classification have been summarized in Table 1.

Genotypic distribution & risk associated with single SNP

The results of χ^2 test done for the control group for each genetic variant in order to check the deviation from Hardy–Weinberg equilibrium have been enlisted in Table 2, which shows these polymorphic sites (rs419558, rs447372, rs17037102, rs3206824, *Axin2* (148, 1365, 432, 956 + 16 and 2062) rs7811989, rs10250822, rs2282885) follow Hardy–Weinberg equilibrium. Using, unconditional logistic regression the risk associated with each polymorphic site of *DKK*, *sFRP3*, *sFRP4*, *Axin2* and *AhrR* was calculated after adjusting for age, gender and smoking status. Minor allele frequencies and adjusted OR are given in Table 2. Out of all the polymorphic sites of *DKK* gene, *DKK2* rs419558 C/T and rs17037102 G/A both were found to confer an increased risk toward LC on the other hand, *DKK3* rs7396187 G/C showed a protective effect. Within the *sFRP4* variants, the variant (AA) genotype of *sFRP4* rs1802073 C/A indicated a threefold risk of developing LC when compared with CC genotype. Also, the

Table 2. Minor allele frequencies and risk associated with single locus sites among cases and controls.

Gene	Polymorphism	Minor allele	MAF _{controls}	MAF _{cases}	AOR _{het} [†]	P _{het} [†]	FDR adjusted P _{het} [†]	AOR _{homo} [†]	P _{homo} [†]	FDR adjusted P _{homo} [†]	AOR _{dom} [†]	P _{dom} [†]	FDR-adjusted P _{dom} [†]	HWE _{controls}
DKK2	rs419558 C/T	T	0.10	0.16	1.82 (1.21–2.68)	0.003	0.021	–	–	–	–	–	–	$\chi^2 = 3.73$, p = 0.05
DKK2	rs447372 A/G	G	0.39	0.29	0.88 (0.62–1.26)	0.5	0.65	1.58 (0.80–3.15)	0.18	0.40	0.97 (0.68–1.36)	0.86	0.9	$\chi^2 = 0.18$, p = 0.66
DKK2	rs17037102 G/A	A	0.08	0.14	1.73 (1.14–2.62)	0.009	0.04	–	–	–	–	–	–	$\chi^2 = 2.53$, p = 0.11
DKK3	rs2291599 C/T	T	0.19	0.21	0.94 (0.63–1.40)	0.79	0.87	1.35 (0.68–2.66)	0.38	0.43	1.02 (0.71–1.47)	0.90	0.9	$\chi^2 = 5.80$, p = 0.01
DKK3	rs3206824 G/A	A	0.14	0.17	1.17 (0.80–1.71)	0.41	0.58	1.61 (0.43–5.95)	0.47	0.50	1.20 (0.82–1.74)	0.33	0.49	$\chi^2 = 0.75$, p = 0.38
DKK3	rs7396187 G/C	C	0.17	0.13	0.60 (0.44–0.94)	0.02	0.07	–	–	–	–	–	–	$\chi^2 = 12.45$, p = 0.0004
DKK4	rs2073664 C/T	T	0.26	0.25	1 (0.70–1.42)	0.97	0.97	0.44 (0.16–1.17)	0.10	0.32	0.94 (0.67–1.34)	0.76	0.87	$\chi^2 = 4.70$, p = 0.03
sFRP3	rs7775 G/C	C	0.14	0.12	0.93 (0.58–1.48)	0.77	0.87	0.65 (0.29–1.41)	0.27	0.40	0.85 (0.56–1.29)	0.47	0.58	$\chi^2 = 27.48$, p = 0.000
sFRP3	rs288326 C/T	T	0.08	0.11	1.31 (0.71–2.43)	0.38	0.58	1.46 (0.67–3.21)	0.33	0.40	1.35 (0.82–2.22)	0.23	0.43	$\chi^2 = 54.29$, p = 0.000
sFRP4	rs1802073 C/A	A	0.46	0.57	1.41 (0.89–2.24)	0.14	0.32	3.19 (1.81–5.63)	0.0001	8e-04	1.76 (1.12–2.75)	0.01	0.03	$\chi^2 = 15.1$, p = 0.0001
sFRP4	rs1802074 G/A	A	0.45	0.51	1.74 (1.07–2.83)	0.02	0.07	1.66 (0.66–4.18)	0.27	0.40	1.69 (1.04–2.74)	0.03	0.07	$\chi^2 = 53.09$, p = 0.000
Axin2	148 C/T	T	0.48	0.42	0.99 (0.66–1.47)	0.96	0.97	0.30 (0.19–0.48)	<0.0001	8e-04	0.84 (0.58–1.23)	0.39	0.53	$\chi^2 = 3.17$, p = 0.07
Axin2	1365 G/A	A	0.10	0.06	0.45 (0.26–0.76)	0.003	0.021	0.96 (0.28–3.35)	0.96	0.40	0.50 (0.30–0.81)	0.005	0.03	$\chi^2 = 2.22$, p = 0.13
Axin2	1386 C/T	T	0.37	0.39	1.16 (0.79–1.70)	0.42	0.58	1.38 (0.76–2.52)	0.28	0.40	1.21 (0.83–1.74)	0.30	0.49	$\chi^2 = 6.77$, p = 0.009
Axin2	2062 C/T	T	0.07	0.07	0.91 (0.57–1.47)	0.72	0.87	–	–	–	–	–	–	$\chi^2 = 1.87$, p = 0.17
Axin2	432 T/C	C	0.02	0.01	0.57 (0.23–1.41)	0.22	0.46	–	–	–	–	–	–	$\chi^2 = 0.19$, p = 0.65
Axin2	956 + 16 A/G	G	0.09	0.08	0.79 (0.50–1.25)	0.32	0.56	0.79 (0.50–1.24)	0.31	0.40	–	–	–	$\chi^2 = 2.67$, p = 0.10
Ahr	rs10250822 T/C	C	0.03	0.03	1.73 (1.20–2.51)	0.003	0.021	1.59 (0.76–3.33)	0.21	0.40	1.71 (1.19–2.44)	0.003	0.03	$\chi^2 = 0.37$, p = 0.53
Ahr	rs2066853 G/A	A	0.19	0.13	0.69 (0.44–1.09)	0.11	0.28	0.40 (0.19–0.86)	0.01	0.04	0.61 (0.41–0.92)	0.01	0.03	$\chi^2 = 24.02$, p = 0.0001
Ahr	rs7811989 A/G	G	0.29	0.37	1.39 (0.95–2.01)	0.08	0.24	2.43 (1.22–4.81)	0.01	0.04	1.52 (1.06–2.17)	0.02	0.06	$\chi^2 = 1.47$, p = 0.15
Ahr	rs2282885 T/C	C	–	–	1.25 (0.83–1.86)	0.27	0.51	1.67 (0.70–3.97)	0.23	0.40	1.30 (0.89–1.91)	0.16	0.34	$\chi^2 = 4.15$, p = 0.04

† Adjusted odds ratios, 95% CIs and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking status.

† Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

All p-values are two sided, p < 0.05 was considered statistically significant.

‘–’ denotes the blanks where the number of subjects were nil.

AOR_{dom}: Adjusted odds ratio for combined variant; AOR_{het}: Adjusted odds ratio for heterozygous; AOR_{homo}: Adjusted odds ratio for homozygous mutant; FDR: False discovery rate; HWE: Hardy–Weinberg equilibrium; MAF: Minor allele frequency; P dom: p-values for combined variant; P het: p-values for heterozygous; P homo: p-values for homozygous mutant.

Table 3. MDR analysis showing interactions of Wnt antagonists (sFRP, DKK and Axin2) variants with lung cancer risk.

Number of risk factors	Best interaction model	Cross validation consistency	Prediction error	p-value
MDR analysis showing interactions of Wnt antagonists (sFRP, DKK and Axin2) variants with lung cancer risk				
1	sFRP4 _{rs1802073}	10/10	0.43	<0.001
2	sFRP4 _{rs1802073} , DKK2 _{rs419558}	9/10	0.40	<0.0001
3	sFRP4 _{rs1802073} , sFRP4 _{rs1802074} , DKK3 _{rs7396187}	5/10	0.45	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) variants with lung cancer risk in smokers				
2	sFRP4 _{rs1802073} , DKK2 _{rs419558}	8/10	0.43	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) variants with lung cancer risk in nonsmokers				
2	sFRP4 _{rs1802074} , sFRP4 _{rs1802073}	8/10	0.47	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) variants with lung cancer risk in ADCC patients				
2	sFRP4 _{rs1802073} , DKK2 _{rs419558}	8/10	0.45	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) variants with lung cancer risk in SQCC patients				
1	DKK2 _{rs17037102}	8/10	0.48	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) variants with lung cancer risk in SQCC patients				
1	sFRP4 _{rs1802073}	9/10	0.43	<0.0001
MDR analysis showing interactions of Wnt antagonists (sFRP, DKK and Axin2) and AhR variants with lung cancer risk				
1	sFRP4 _{rs1802073}	8/10	0.45	0.001
2	sFRP4 _{rs1802073} , DKK2 _{rs419558}	8/10	0.43	<0.0001
3	sFRP4 _{rs1802073} , DKK2 _{rs419558} , Axin2 _{rs9915936}	3/10	0.46	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) and AhR variants with lung cancer risk in smokers				
1	AhR _{rs10250822}	10/10	0.42	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) and AhR variants with lung cancer risk in nonsmokers				
3	AhR _{rs10250822} , sFRP3 _{rs288326} , DKK2 _{rs17037102}	7/10	0.40	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) and AhR variants with lung cancer risk in ADCC patients				
2	AhR _{rs10250822} , sFRP3 _{rs288326}	10/10	0.36	<0.0001
MDR analysis showing interactions of (sFRP, DKK and Axin2) and AhR variants with lung cancer risk in SQCC patients				
1	DKK2 _{rs447372}	6/10	0.52	0.002
MDR analysis showing interactions of (sFRP, DKK and Axin2) and AhR variants with lung cancer risk in SQCC patients				
1	sFRP4 _{rs1802073}	9/10	0.44	<0.0001

ADCC: Adenocarcinoma; SQCC: Squamous cell carcinoma.

carrier genotype of *sFRP4* rs1802074 G/A exhibited a slightly increased risk of acquiring LC. *Axin2* 148 C/T and 1365 G/A variant showed a strong protective effect toward LC. Out of the four *AhR* variants, rs2066853 showed a decreased risk whereas rs7811989 showed 2.4-fold risk of acquiring LC.

A strong linkage disequilibrium (LD) was observed in case of *Axin2* 432 and 1365 ($D' = 0.89$, $r^2 = 0.001$). Similarly the linkage disequilibrium was also found to be strong in case of *Axin2* 148 & 432 SNPs ($D' = 0.628$, $r^2 = 0.007$) and *Axin* 956 & 432 ($D' = 0.243$, $r^2 = 0.012$). A strong LD was observed in case of *DKK2* rs419558 and rs17037102 ($D' = 0.68$, $r^2 = 0.01$), similarly *DKK2* rs419558 and rs447372 also showed a considerable LD with a $D' = 0.53$ $r^2 = 0.01$. In case of *DKK3* variants, *DKK3* rs2295199 and rs3206824 showed a moderate LD ($D' = 0.49$, $r^2 = 0.16$). It was found that the variants of both *sFRP3* and *sFRP4* genes were not in LD. *AhR* rs7811989 and rs2282885 also illustrated a strong linkage disequilibrium ($D' = 0.108$, $r^2 = 0.05$).

Multifactor dimensionality reduction

Table 3 summarizes the multifactor dimensionality reduction results for the Wnt antagonists gene namely *sFRP*, *DKK* and *Axin2*. This table demonstrates various gene–gene interactions models and gene–environment interaction models which have been studied. All the possible interaction models have been listed in the table for patients. Also, the analysis was done for various subgroups stratified on the basis of smoking and histology. The author has not assessed association of each possible pair of SNP's with age, gender and smoking. However, the double combinations have been analyzed for their effect on the LC risk (data not shown) after adjusting for age, gender

and smoking. Due to excess of the data, these data were not reported in this manuscript. Also, an attempt was made to pool age, gender and smoking along with variants, it was observed none of these factors were found to be part of any interaction model. So, not much affect would be seen in the results of the final interaction detection strategy in the presence of confounders. Out of all the interactions models, the one having the maximum CVC, a minimum prediction error (1-testing accuracy) and a permutation low p-value was considered as the best model.

Overall, the best interaction model was one factor model sFRP4_{rs1802073} which has CVC of 10/10, prediction error = 0.43 and $p > 0.0001$. The second best model was the two factor model comprising sFRP4_{rs1802073} and DKK2_{rs419558} with CVC of 9/10 and prediction error = 0.40. This model shows the interactive effect of two genes which are keenly involved in the regulation of Wnt signaling and are found to contribute significantly in governing LC predisposition. An attempt was made to analyze the interactions between these variants of these genes in different categories so as to find out the significant contributing factors for these groups. In case of smokers, the best interaction model was sFRP4_{rs1802073} and DKK2_{rs419558} with CVC of 8/10, prediction error = 0.43 and $p < 0.0001$, this model comprised of the same factors as obtained in case of overall analysis. However, in nonsmoker category the best interaction model was a two factor model comprising sFRP4_{rs1802074}, sFRP4_{rs1802073} with a CVC of 8/10, prediction error = 0.47 and $p = 0.002$. This denotes that the *DKK* and *Axin2* gene are not interacting with the secreted frizzled-related receptor to modulate the LC susceptibility in nonsmokers. Further, stratifying the patients on the basis of histological subtype, it was observed that in case of adenocarcinoma (ADCC) patients, the best interaction model is sFRP4_{rs1802073} and DKK2_{rs419558} having CVC of 8/10, prediction error = 0.45 and $p < 0.0001$. However, in the squamous cell carcinoma (SQCC) subtype, the one factor model composed of DKK2_{rs17037102} was the best one with highest CVC of 8/10, prediction error = 0.48 and $p < 0.0001$. Moreover in the case of small-cell lung cancer (SCLC) patients, sFRP4_{rs1802073} was found to be best factor in modulating the LC susceptibility with a CVC of 9/10, prediction error = 0.43. SQCC and SCLC are more likely to be attributed to smoking as compared with ADCC patients.

Nonparametric interaction analysis of the AhR gene with the Wnt antagonists revealed the absence of any AhR variant in the possible interactions model obtained by MDR approach as shown in Table 3. The best interaction model included sFRP4_{rs1802073} and DKK2_{rs419558} with CVC of 8/10 and prediction error = 0.43. However, when the MDR approach was used in various subgroups such as smokers, AhR_{rs10250822} came out to be the best model with CVC of 10/10, prediction error of 0.42 and $p < 0.0001$. On the contrary, in nonsmokers the best interaction model was a three factor one, including AhR_{rs10250822}, sFRP3_{rs288326} and DKK2_{rs17037102} with a CVC of 7/10, prediction error = 0.40 and $p < 0.0001$. This model shows the cumulative effect of AhR and Wnt antagonist genes in modulating risk toward LC. Looking into the different categories made on the basis of histology, the ADCC patients showed the two factor model as the best one which included AhR_{rs10250822}, sFRP3_{rs288326} with a CVC of 10/10, and minimum prediction error of 0.36, $p < 0.0001$. This model is having the minimum prediction error and maximum CVC out of all the observed models in all categories. In case of SQCC patients, the single factor model (DKK2_{rs447372}) was showing the maximum contribution in affecting LC susceptibility. However, in case of SCLC patients the AhR variants did not portray any significant effect on the model governing the risk in these patients, the single factor model, sFRP4_{rs1802073} was found to be the best with CVC of 9/10, prediction error = 0.44.

The entropy dendrogram for Wnt antagonist genes have been illustrated in Figure 1A. Similarly the dendrograms depicting the interactions of AhR and Wnt antagonists have been shown in Figure 1B. The dendrogram depicts the extent of the interaction between the risk factors which form the part of the interaction model. The longer the length of the lines connecting two risk factors, the interaction would be weaker. Also, the colors of the lines depict the degree of interaction; red and orange indicate synergy between SNPs, and yellow denotes independence. Green and blue indicate redundancy or no interaction.

Classification & regression tree

To have an insight about the interaction of smoking as a parameter with the Wnt antagonist genetic variants, CART was also performed using smoking as a categorical variable and important factor for the lung carcinogenesis. In this decision tree, smoking formed the first split into smokers and nonsmokers. Henceforth, terminal nodes were separated for smokers and nonsmokers. The data are tabulated in Table 4. The tree having 28 terminal nodes has been illustrated in Figure 2. The lowest case rate (10) was obtained for terminal node 28 having the genotype (DKK3_{rs7396187} [M]/DKK3_{rs3206824} [W]/DKK2_{rs17037102} [W]), this is taken as the reference. In case of smokers, the most significant factor was DKK2_{rs419558}. Where as in case of nonsmokers it was the other variant of DKK2, DKK2_{rs17037102}. Several high-risk subgroups were reported in case of smokers. Out of which the highest risk was

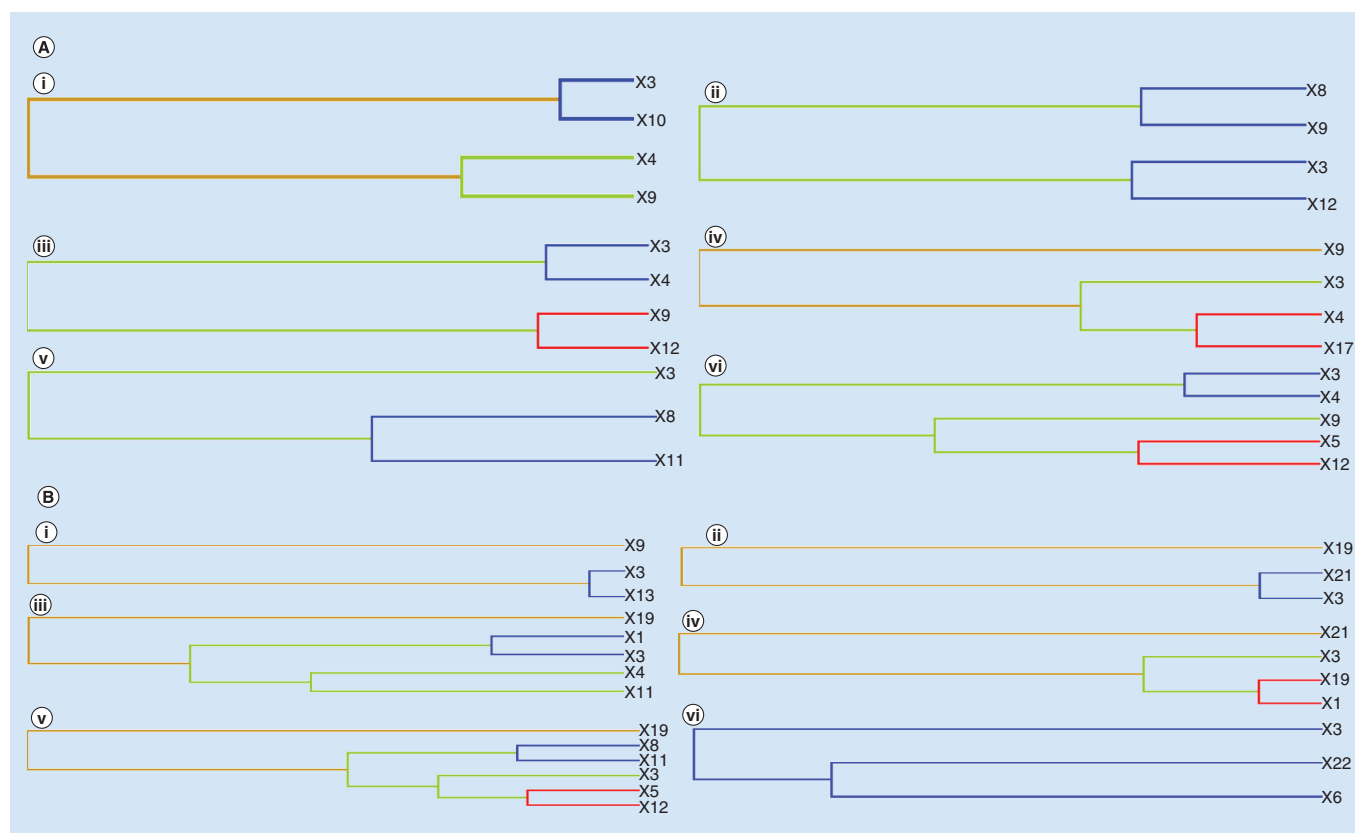


Figure 1. Multifactor dimensionality reduction (MDR) interaction dendrogram. (A) Shows MDR interactions between Wnt antagonist (sFRP, DKK, Axin2) and lung cancer risk: (i) in all the cases and controls; (ii) in smokers; (iii) in non-smokers; (iv) in adenocarcinoma (ADCC) patients; (v) in squamous cell carcinoma (SQCC) patients; and (vi) in small-cell lung cancer (SCLC) patients. **(B)** Shows MDR interactions between Wnt (sFRP, DKK, Axin2) and AhR with lung cancer risk: (i) in all the cases and controls; (ii) in smokers; (iii) in non-smokers; (iv) in ADCC patients; (v) in SQCC patients; and (vi) in SCLC patients. X1: sFRP3 rs288326; X2: sFRP3 rs7775; X3: sFRP4 rs1802073; X4: sFRP4 rs1802074; X5: DKK3 rs2291599; X6: DKK4 rs2073664; X7: DKK3 rs3206824; X8: DKK2 rs447372; X9: DKK2 rs419558; X10: DKK3 rs7391689; X11: DKK2 rs17037102; X12: Axin2 148; X13: Axin2 1365; X14: Axin2 1386; X15: Axin2 1712,12; X16: Axin2 432; X17: Axin2 2062; X18: Axin2 956+16; X19: AhRrs10250822; X20: AhR rs2282885; X21: AhRrs2066853; X22: AhR rs7811989.

associated with the genotype (DKK2_{rs17037102} [M]/Axin2_{rs9915936} [W]/DKK2_{rs419558} [M]) of terminal node1, which showed a significant 76-fold risk. Terminal node 7, 8, 10 and 13 showed 18-fold risk of developing LC. Several other terminal nodes such as 2, 3 and 4 which have DKK2_{rs419558} (M) as trailing genotype also showed a higher risk of developing LC. Within these genotypic combinations, the interaction of three Wnt antagonist gene is highly noticeable indicating their cumulative role in controlling the expression of Wnt cascade. On the other hand nonsmoker group showed a lesser number of terminal nodes, out of which the high-risk subgroups included the subjects with the genotype of terminal node 21, 22, 25 and 26. The highest risk was portrayed by subjects having sFRP4_{rs1802073} (M)/DKK3_{rs3206824} (M)/DKK2_{rs17037102} (W) genotype. The OR was 24 (4.22–136.22), $p = 0.003$.

CART analysis was done for Wnt antagonists (DKK, sFRP and Axin2) variants; the tree formed is illustrated in **Supplementary Figure 1**. The tree shows that initial split was formed by DKK2_{rs419558}, therefore this SNP becomes the most significant factor out of all the variants under study. The data showing all the terminal nodes have been listed in **Supplementary Table 1**. The lowest case rate of 16.6 was observed in the case of terminal node 6 having the genotype DKK3_{rs3206824} (W)/DKK3_{rs7391689} (M)/DKK2_{rs17037102} (W)/DKK2_{rs419558} (M), which is taken as the reference to calculate odds for other terminal nodes. The patients with the genotypic combination of terminal node 1 (DKK2_{rs17037102} [M]/DKK2_{rs419558} [M]) showed a tenfold risk of acquiring LC (OR = 10.35 [3.18–33.67], $p = 0.0001$). The other terminal nodes which formed the high-risk subgroups include the subjects with genotypic combination of terminal node 2, 3 and 8. Certain low-risk subgroups were also identified, they

Table 4. Risk estimates based on classification and regression tree analysis (sFRP, DKK and Axin2 variants).

Terminal node in smokers	Genotype for each terminal node	Control (n)	Case (n)	Case rate ¹	AOR (95% CI)	p-value ¹
Node 1	DKK2 _{rs17037102} (M)/Axin2 _{rs9915936} (W)/DKK2 _{rs419558} (M)	2	17	89.47	76.50 (9.66–605.65)	0.00004
Node 2	Axin2 _{rs2204038} (W)/DKK2 _{rs17037102} (W)/Axin2 _{rs9915936} (W)/DKK2 _{rs419558} (M)	7	23	76.66	29.57 (5.46–159.96)	0.00008
Node 3	DKK3 _{rs7396187} (W)/sFRP3 _{rs7775} (W)/Axin2 _{rs2204038} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs9915936} (W)/DKK2 _{rs419558} (M)	7	24	77.41	30.8 (5.71–166.58)	0.00007
Node 4	DKK3 _{rs7396187} (M)/sFRP3 _{rs7775} (W)/Axin2 _{rs2204038} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs9915936} (W)/DKK2 _{rs419558} (M)	7	4	36.36	5.14 (0.76–34.68)	0.09
Node 5	sFRP3 _{rs7775} (M)/Axin2 _{rs2204038} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs9915936} (W)/DKK2 _{rs419558} (M)	11	5	31.25	4.09 (0.67–24.82)	0.12
Node 6	Axin2 _{rs9915936} (M)/DKK2 _{rs419558} (M)	12	8	40	6 (1.08–33.27)	0.04
Node 7	DKK2 _{rs17037102} (M)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	12	25	67.56	18.75 (3.72–94.26)	0.0003
Node 8	DKK3 _{rs2291599} (W)/DKK3 _{rs7396187} (W)/DKK2 _{rs17037102} (W)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	18	36	66.66	18 (3.75–86.22)	0.0003
Node 9	DKK3 _{rs2291599} (M)/DKK3 _{rs7396187} (W)/DKK2 _{rs17037102} (W)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	7	6	46.15	7.71 (1.24–47.75)	0.028
Node 10	DKK3 _{rs2291599} (M)/Axin2 _{rs35285779} (W)/DKK3 _{rs7396187} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	4	8	66.66	18 (2.71–119.23)	0.002
Node 11	DKK3 _{rs2291599} (W)/Axin2 _{rs35285779} (W)/DKK3 _{rs7396187} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	5	1	16.66	1.80 (0.13–24.15)	0.65
Node 12	Axin2 _{rs35285779} (M)/DKK3 _{rs7396187} (M)/DKK2 _{rs17037102} (W)/Axin2 _{rs1133683} (M)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	5	0	0	0.00	0.98
Node 13	DKK3 _{rs3206824} (W)/Axin2 _{2062 C/T} (W)/DKK2 _{rs447372} (W)/Axin2 _{rs1133683} (W)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	9	18	66.66	18 (3.40–95.20)	0.0067
Node 14	DKK3 _{rs3206824} (M)/Axin2 _{2062 C/T} (W)/DKK2 _{rs447372} (W)/Axin2 _{rs1133683} (W)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	8	6	42.85	6.75 (1.11–41.00)	0.038
Node 15	Axin2 _{2062 C/T} (M)/DKK2 _{rs447372} (W)/Axin2 _{rs1133683} (W)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	6	3	33.33	4.50 (0.60–33.70)	0.14
Node 16	DKK2 _{rs447372} (M)/Axin2 _{rs1133683} (W)/sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	21	10	32.25	4.28 (0.82–22.17)	0.08
Node 17	sFRP3 _{rs288326} (M)/sFRP4 _{rs1802074} (W)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	1	3	75	27 (1.82–399.23)	0.01
Node 18	DKK2 _{rs17037102} (M)/sFRP3 _{rs288326} (W)/sFRP4 _{rs1802074} (W)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	3	5	62.50	15 (1.94–115.96)	0.009
Node 19	DKK2 _{rs17037102} (W)/sFRP3 _{rs288326} (W)/sFRP4 _{rs1802074} (W)/sFRP4 _{rs1802073} (M)/DKK2 _{rs419558} (W)	18	6	25	3 (0.53–16.89)	0.21
Node 20	sFRP4 _{rs1802074} (W)/sFRP4 _{rs1802073} (W)/DKK2 _{rs419558} (W)	1	4	80	36 (2.58–501.26)	0.007
Node 21	sFRP4 _{rs1802074} (M)/sFRP4 _{rs1802073} (W)/DKK2 _{rs419558} (W)	33	17	34	4.63 (0.96–22.36)	0.05
Terminal node in nonsmokers						
Node 22	DKK2 _{rs17037102} (M)	7	18	72	23.14 (4.21–126.91)	0.003
Node 23	sFRP4 _{rs1802073} (M)/DKK3 _{rs3206824} (M)/DKK2 _{rs17037102} (W)	6	16	72.72	24 (4.22–136.22)	0.003
Node 24	sFRP4 _{rs1802073} (W)/DKK3 _{rs3206824} (M)/DKK2 _{rs17037102} (W)	4	1	20	2.25 (0.16–31.32)	0.54

¹ Case rate is percentage lung cancer cases among all individuals in each terminal node (case/[case + control] × 100)
² p < 0.05 is considered statistically significant
³ Reference group
AOR: Adjusted odds ratio.

Table 4. Risk estimates based on classification and regression tree analysis (sFRP, DKK and Axin2 variants) (cont.).

Terminal node in smokers	Genotype for each terminal node	Control (n)	Case (n)	Case rate [†]	AOR (95% CI)	p-value [‡]
Node 25	Axin2 _{rs1133683} (W)/DKK3 _{rs7396187} (W)/DKK3 _{rs3206824} (W)/DKK2 _{rs17037102} (W)	7	13	65	16.71 (2.97–93.88)	0.001
Node 26	DKK3 _{rs2291599} (M)/Axin2 _{rs1133683} (M)/DKK3 _{rs7396187} (W)/DKK3 _{rs3206824} (W)/DKK2 _{rs17037102} (W)	5	8	61.50	14.4 (2.28–90.56)	0.004
Node 27	DKK3 _{rs2291599} (W)/Axin2 _{rs1133683} (M)/DKK3 _{rs7396187} (W)/DKK3 _{rs3206824} (W)/DKK2 _{rs17037102} (W)	19	5	20.83	2.36 (0.40–13.79)	0.33
Node 28	DKK3 _{rs7396187} (M)/DKK3 _{rs3206824} (W)/DKK2 _{rs17037102} (W)	18	2	10	1.0 [§]	

[†] Case rate is percentage lung cancer cases among all individuals in each terminal node (case/[case + control] × 100)

[‡] p < 0.05 is considered statistically significant

[§] Reference group

AOR: Adjusted odds ratio.

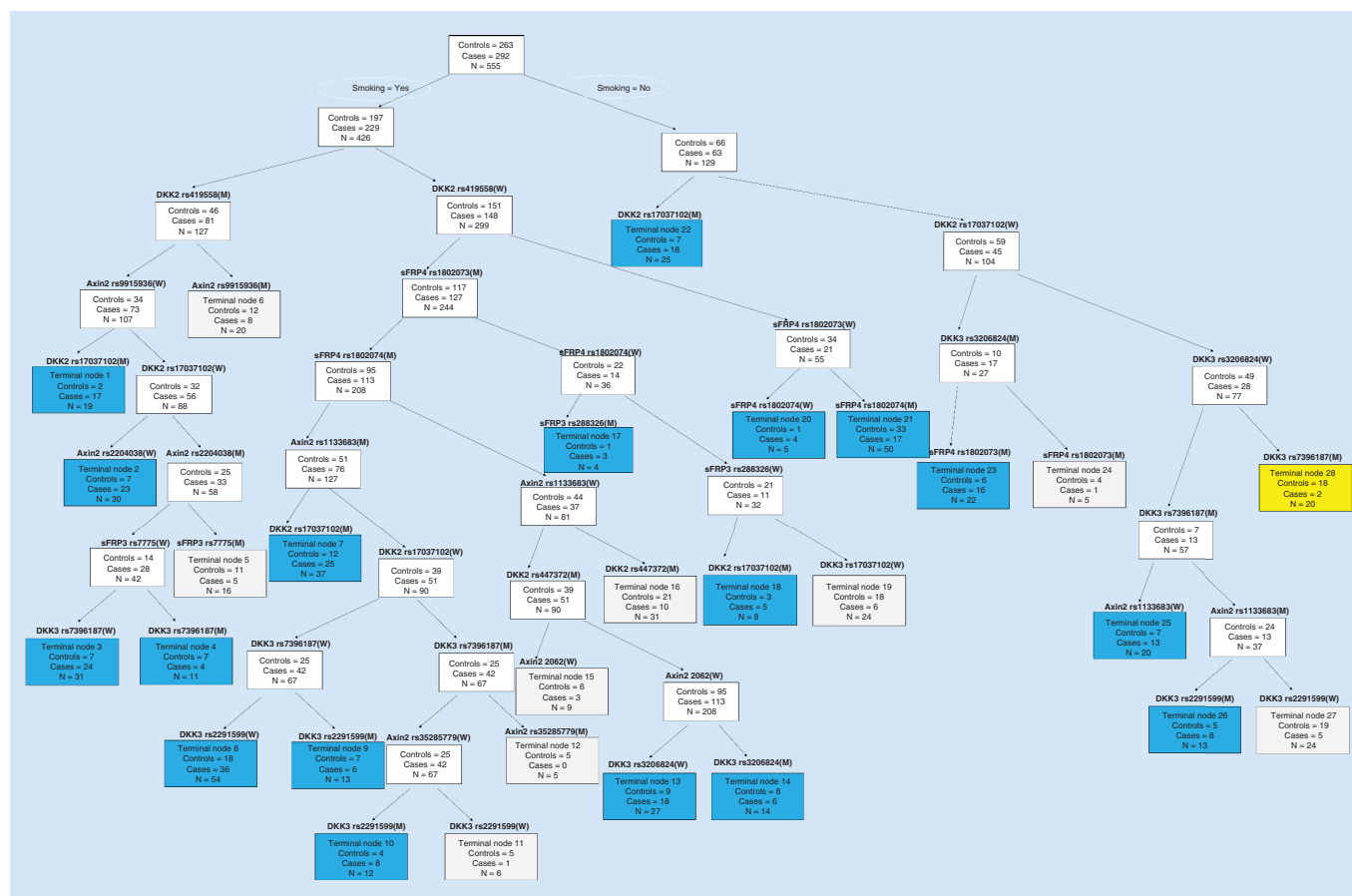


Figure 2. Classification and regression tree analysis of SFRP, DKK and DKK variants on the basis of smoking. Yellow box indicates reference terminal nodes and blue boxes indicate high-risk group. M: Variant type; W: Wild type.

harbored the genotype as shown in terminal node 5 (DKK3_{rs3206824} [M]/DKK3_{rs7396187} [M]/DKK2_{rs17037102} [W]/DKK2_{rs419558} [M]). The OR associated with terminal node 5 is 0.40 (0.10–1), p = 0.008.

Gene–gene interactions are analyzed using CART tool for finding out the possible cross relations between AhR and Wnt antagonist's genetic variants. This analysis revealed the genotypic combinations which were significantly contributing in affecting the predisposition toward LC. This multigene approach is scrutinizing two completely

different pathways, to infer their cumulative role in defining the propensity of these variants to modulate LC pathogenesis. When AhR variants along with the other polymorphic sites of Wnt antagonists were posed to CART analysis, the most important factor which formed the initial split was AhR_{rs10250822}. The tree illustrates the presence of 18 terminal nodes as in **Supplementary Figure 2**. OR associated with each terminal node has been tabulated along with the genotypic combinations in **Supplementary Table 2**. The lowest case rate was observed for terminal node 16 having the genotype DKK2_{rs17037102} (M)/sFRP4_{rs1802074} (W)/sFRP4_{rs1802073} (M)/DKK2_{rs419558} (W), which is used as a reference. Various high-risk subgroups were evident after this analysis, these majorly included the involvement of AhR_{rs10250822}, DKK2_{rs17037102} and DKK2_{rs419558}. The genotypic combinations of terminal node 1, 3, 10 and 15 conferred high risk in subjects with these genotypes. Terminal node 1 having the genotype DKK2_{rs17037102} (M)/AhR_{rs2066853} (W)/AhR_{rs10250822} (M) showed an 11-fold risk of developing LC (OR = 11.57[4.24–31.51], $p = 0.00001$).

AhR has been closely related with smoking as it is prime most receptor which encounters the carcinogens present in the smoke when inhaled. Consequently, it is important to find out the interactive association of AhR variants and Wnt antagonists along with smoking as the major factor. CART analysis was applied using smoking as the significant factor for occurrence of LC and the tree structured in this case has smoking as the first split into smokers and nonsmokers. As evident in **Supplementary Figure 3**, the initial split in case of smokers was formed by AhR_{rs10250822} whereas the first split in case of nonsmokers was formed by DKK2_{rs17037102}. Henceforth, it can be concluded that AhR variants are not significantly involved in modulating LC susceptibility in nonsmokers. In case of smokers, two terminal nodes having the presence of two AhR variants AhR_{rs2066853} and AhR_{rs10250822} in common were identified as high-risk subgroups. Subjects having the genotype AhR_{rs2066853} (W)/AhR_{rs10250822} (M) and Axin2_{rs2204038} (W)/AhR_{rs2066853} (M)/AhR_{rs10250822} (M) showed a high risk of developing LC with highly significant p -values as shown in **Supplementary Table 3**. Our data also suggests that, sub-group of subjects who were at high risk towards lung cancer risk were especially nonsmokers exhibiting the various polymorphic combinations of DKK2_{rs17037102} along with the other variants of DKK3 with an OR of 12.50 (2.25–69.18), $p = 0.003$.

Discussion

Increasingly growing knowledge about Wnt field is giving us new insights about its broad and effective spectrum in cancer therapeutics. The extend of involvement of Wnt pathway cannot be entangled to a single dimension as it has highly complex and intermeshing network which controls numerous cellular functions by impeding a central role in the process of carcinogenesis [44]. AhR signaling is imperative in toxicity-mediated response and caters to the various carcinogens so that they get metabolized and reach their desired fate. AhR pathway has various unraveled roles in mediating molecular toxicity and cellular homeostasis [45]. The interactive potential of these two pathways is intricately involved in transducing the transcriptional machinery of various target genes such as *matrix metalloproteins*, *cox-2*, *cyclin D1*, *VEGF*, *c-myc*, *Cyp1a1*, *Cyp1a2*, among others [37]. The present study has made an attempt to unfold the hidden interactive contributions conferred by the panel of genetic variants of these two pathways in LC predisposition. This multigene approach can help in overcoming the shortcomings of traditional single gene studies where the predictive power of the study is restricted. The pathway based study can help in providing a comprehensive conclusion about the associations of several polymorphic sites with cancer risk and prognosis.

In the current study, a systematic approach is used to interpret the combined and interactive ability of several SNPs within Wnt antagonist genes and AhR to predict the LC risk in north Indian population. This is probably the first attempt to evaluate the collaborative effect of the sequence variations among Wnt antagonists and AhR on LC risk, and the findings are new and have not been found in large-scale investigations. This is a highly unexplored arena which can add a significant dimension to the LC predisposition.

Several sequence variations of the Wnt antagonist genes have been evaluated for influencing the cancer risk in different ethnicities. Two widely studied polymorphic sites of sFRP3 gene namely, rs7775 and rs288326 have been studied in German colorectal cancer patients where only sFRP3 Arg³²⁴Gly was found to confer an increased risk [13]. Another report demonstrated a lack of association of these two variants with colorectal neoplasia [15]. Polymorphic variant of sFRP4 rs1802073 was found to be positively correlated with increased risk for rectal cancer and early-stage colorectal cancer in German population [16]. The findings of a study done in Turkish LC patients denied the functional significance of rs1802074 in influencing the susceptibility [20]. Recently, several different variants of the DKK4 (rs2073664), DKK3 (rs2291599, rs3206824 and rs7391689) and DKK2 (rs447372,

rs419558 and rs17037102) have been analyzed in north Indian LC patients. Out of them, *DKK3* rs7396187 showed a protective effect ($p = 0.01$). *DKK2* rs17037102 and rs419558 conferred an increased risk of developing LC. Genotypic combination, *DKK3* rs3206824 and *DKK2* rs419558 showed a twofold increased risk of developing LC ($p = 0.008$) [46]. Another study has investigated the role of *DKK2* (rs17037102), *DKK3* (rs3206824), *DKK3* intron4 G/C (rs7396187), *DKK4* (rs2073664) and *sFRP4* (rs1802074) variants in Turkish LC subjects and none of the sequence variation conferred a significant effect on LC risk. Only the gene combination *DKK3* rs3206824 and *sFRP4* rs1802074 conferred a significantly decreased risk of LC [20]. *Axin2* variants have been investigated in Turkish LC patients where the rs2204038 polymorphic site where in the TT genotype conferred a decreased risk of LC [22]. Similar findings have been observed in Chinese [25] and Japanese LC patients [23]. Reports are available in case of other type of cancers including breast [47], prostate [26], astrocytoma [27] and ovarian [24]. A study done in Chinese population explored a significant association of *Ahr* rs7811989 along with rs2158041 both residing in the intronic regions with higher risk of LC [48]. Prior studies conducted in Korean [49], Japanese [50], French [51,52] and Finnish [53,54] population showed a lack of association with regard to Arg⁵⁵⁴Lys polymorphic site of *Ahr*.

Previously some studies have taken into account a comprehensive panel of Wnt signaling pathway sequence variations and carried out the association analysis using traditional logistic regression method. A study done in renal cancer patients has analyzed the several variants of *DKK2* (rs447372, rs419558 and rs17037102), *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK4* (rs2073664), *sFRP4* (rs1802073 and rs1802074), *DAAM2* (rs6937133 and rs2504106) and *SMAD7* (rs12953717). Although *sFRP4* rs1802073 was not found to be correlated to renal cancer risk but being a nonsynonymous substitution (Pro³²⁰Thr), it has been characterized by Polyphen software as a 'probably damaging' change which can alter the functionalities of this extracellular glycoprotein [17]. In our study, *sFRP4* rs1802073 came out to be a significant contributor and was found to be a major component of the combinations and interaction models conferring to increasing the risk for LC. *DKK2* rs17037102 is found to be associated with survival of renal cancer patients [17]; on the contrary in our study it came out to be a significant predictor of LC susceptibility both in MDR and CART analysis. Another important study which evaluated polymorphic sites of Wnt cascade was done in non-small-cell lung cancer patients to find out the relapse-free survival after being treated for early-stage non-small-cell lung cancer [55]. A report on colon cancer investigated a range of SNPs of *DKK2*, *DKK3*, *sFRP4* and *Axin2* genes in predicting their clinical influence in tumor recurrence [14]. Recent findings from a breast cancer study done in Saudi Arabia also significantly established the role of germline variations of Wnt signaling pathway where 15 different SNPs with in *CTNNT1*, *DKK2*, *DKK3*, *sFRP3*, *sFRP4* and *Axin2* gene of Wnt cascade were examined to find their association with breast cancer risk [18].

The abovementioned studies have drawn the conclusions on the basis of individual gene association analysis, till there is no report demonstrating the collaborative effect of Wnt antagonists to modulate LC susceptibility. However, a few studies have looked into the interactive potential of several genes in various cancers such as bladder cancer [56,57], gallbladder cancer [43], gastric adenocarcinoma [58] and tobacco-related multiple primary neoplasms [59], among others. These studies have employed the high order statistics which has high power in forecasting the multiple-faceted role of these SNPs along with their interaction with the environmental parameters such as smoking. MDR and CART tools were used and some interesting conclusions about the contributions of all the germline variations under study became evident. These nonparametric methods have gained much more significance recently as they give us a clear picture about the various factors which might play an inevitable role in impeding cancer predisposition. Also, unfold the cumulative role of low penetrance alleles in essentially anticipating the relevance of their interactions in predicting the disease occurrence [59].

Interesting findings of the present study comprised the evident role of double loci genotypic combinations where *sFRP4* rs1802073 and *DKK2* variants (rs17037102 and rs419558) along with *sFRP4* rs1802074 and *DKK3* rs3206824 were found to contribute maximum in increasing the risk of LC. It is the first study which investigated the germline variations of two paramount signaling cascade and their accumulative effect on LC risk. Interaction models within Wnt pathway showed the evident partnership of *sFRP4* rs1802073 (nonsynonymous substitution) with *DKK2* rs419558 which is located in the 3'-UTR region. The synergistic effect of *Ahr* and Wnt antagonists was evidently explored through MDR analysis but none of the *Ahr* variant was interacting with Wnt antagonist but in case of smokers *Ahr* rs10250822 was alone determined as the best model for predicting the associated risk for LC. A study was conducted in the Chinese population where *Axin2* SNPs (rs2240308, rs22040307) and *MMP7* (rs4791169) as well as smoking were taken as parameters to find out the best interaction model that contributed towards LC risk [60]. Another study also used this exhaustive MDR method in ovarian cancer to find out the interactions between variants of *Axin2*, β -catenin and adenomatous polyposis coli [24].

Considering the CART output in case of Wnt antagonist, *DKK2* came out to be the most significant factor in determining the LC risk. *DKK2* is found to be interacting with other Wnt antagonists namely *sFRP3*, *Axin2* and *DKK3* as many of the genotypic combinations which fell in high-risk subgroups comprised the SNPs from these genes such as *DKK3*_{rs3206824} (M)/*DKK3*_{rs7396187} (M)/*DKK2*_{rs17037102} (W)/*DKK2*_{rs419558} (M) and *Axin2*_{rs2204038} (W)/*DKK3*_{rs7396187} (W)/*DKK2*_{rs17037102} (W)/*DKK2*_{rs419558} (M). Investigating smoking as a significant factor, it was evident that *sFRP4* and *DKK2* SNPs were prevalent predictors of LC risk and on the other hand in smokers the evident role of three stage Wnt regulatory machinery was illustrated by genotypic combinations such as *DKK3*_{rs2291599} (M)/*Axin2*_{rs35285779} (W)/*DKK3*_{rs7396187} (M)/*DKK2*_{rs17037102} (W)/*Axin2*_{rs1133683} (M)/*sFRP4*_{rs1802074} (M)/*sFRP4*_{rs1802073} (M)/*DKK2*_{rs419558} (W) where a high risk of LC was indicated. Synergistic role of *AhR* and Wnt pathway was examined using high order CART analysis, where two important SNPs of *AhR* gene (*rs10250822* and *rs2068553*) were cumulatively contributing in LC risk with *DKK2* rs17037102. High-risk subgroups among smokers were subjected to harbor a genotypic combination where *AhR* variants were found to be integrated with *Axin2* polymorphic site; however, in nonsmokers none of the *AhR* variant was found to determine the predisposition. This affirmed the interactive potential of *AhR* with Wnt antagonists as both have been reported to cater the stability of β -catenin, which is a primary protein which augments and transduces the expression of different oncogenic- and xenobiotic-metabolizing genes.

Conclusion

Our study is a primary effort to introduce the concept of polygenic approach to get an insight about the various polymorphic variants in determining cancer susceptibility. However, smaller sample size and lack of functional studies are a limitation to the study prospect. Also, lesser number of subjects were found in the high-risk subgroups but then future endeavors in this regard can give a new dimension about lung carcinogenesis. The therapeutic potential of these pathways can also be explored on the basis of the findings of the current study. Further studies with larger sample size and more accurate statistical methods are required to warranty the above findings.

Future perspective

The present study employed a multifactorial approach to investigate the significance of interactive effects of genes and environment on LC susceptibility. In the recent scenario, the pathway-based approach to decipher their roles in the process of carcinogenesis has gained importance. Identification of combinatorial markers has been the need of the hour. In order to improve the diagnostics and prognostics for LC, relying on the candidate gene approach is not going to aid in turning any milestones in the process of finding out new dimensions about LC. These kinds of studies will guide in process of finding the important pathways and their interactions which might play an important role in governing the heterogeneity of LC. The pathogenesis of LC and its different types still needs to be explored and hence, focusing on these multidimensional markers such as genetic variants within a single signaling cascade and other intermeshing pathway can help open new avenues for targeting their therapeutic and diagnostic potential.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/suppl/10.2217/pme-2017-0018

Financial & competing interests disclosure

This work was supported by grant from the Indian Council of Medical Research, New Delhi, India (Grant No. 5/13/126/2011/NCD-III). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Summary points

- The constitutive expression of Wnt and AhR pathways may be augmented to the genetic variations present within the genes regulating them which might be related to the lung cancer (LC) susceptibility.
- sFRP4_{rs1802073} which has cross validation consistency (CVC) of 10/10, prediction error = 0.43 ($p > 0.0001$) is the best factor determining the LC risk.
- The second best model was the two factor model comprising sFRP4_{rs1802073} and DKK2_{rs419558} with CVC of 9/10 and prediction error = 0.40.
- Adenocarcinoma patients showed the two factor model as the best one comprising AhR_{rs10250822}, sFRP3_{rs288326} with a CVC of 10/10, minimum prediction error of 0.36 ($p > 0.0001$).
- In classification and regression tree analysis, DKK2_{rs419558} came out to be a significant factor out of all the Wnt antagonists; DKK2_{rs17037102} (M)/DKK2_{rs419558} (M) showed a tenfold risk of acquiring LC (odds ratio = 10.35 [3.18–33.67], $p = 0.0001$). DKK2_{rs17037102} (M)/AhR_{rs2066853} (W)/AhR_{rs10250822} (M) showed an 11-fold risk of developing LC (odds ratio = 11.57 [4.24–31.51], $p = 0.00001$).
- DKK2 along with sFRP4 is certainly playing a crucial role in modulating susceptibility of LC patients, and particularly in case of smokers.
- AhR variants are contributing maximally toward LC risk. This study provides a new insight about the occurrence of LC in north Indians.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

- Mucci LA, Wedren S, Tamimi RM, Trichopoulos D, Adami HO. The role of gene–environment interaction in the aetiology of human cancer: examples from cancers of the large bowel, lung, and breast. *J. Intern. Med.* 249, 477–493 (2001).
- Neumann AS, Sturgis EM, Wei Q. Nucleotide excision repair as a marker for susceptibility to tobacco-related cancers: a review of molecular epidemiological studies. *Mol. Carcinog.* 42, 65–92 (2005).
- Horne BD, Anderson JL, Carlquist JF *et al.* Generating genetic risk scores from intermediate phenotypes for use in association studies of clinically significant endpoints. *Ann. Hum. Genet.* 69, 176–186 (2005).
- Qian DC, Han Y, Byun J *et al.* A novel pathway-based approach improves lung cancer risk prediction using germline genetic variations. *Cancer Epidemiol. Biomarkers Prev.* 25, 1208–1215 (2016).
- **Focused on the importance of pathway-based approach to predict lung cancer risk.**
- Mazieres J, He B, You L, Xu Z, Jablons DM. Wnt signaling in lung cancer. *Cancer Lett.* 222, 1–10 (2005).
- **Elucidates the role of Wnt signaling in the process of lung carcinogenesis.**
- Yang Y. Wnt signaling in development and disease. *Cell Biosci.* 2, 1 (2012).
- **Comprehends the therapeutic potential of Wnt pathway and its proteins in all the diseases including cancer.**
- Luo J, Chen J, Deng ZL *et al.* Wnt signaling and human diseases: what are the therapeutic implications? *Lab. Invest.* 87, 97–103 (2007).
- Akiyama T. Wnt/ β -catenin signaling. *Cytokine Growth Factor Rev.* 11, 273–282 (2000).
- Shi Y, He B, You L, Jablons DM. Roles of secreted frizzled-related proteins in cancer. *Acta Pharmacol. Sin.* 28, 1499–1504 (2007).
- Zorn AM. Wnt signalling: antagonistic Dickkopfs. *Curr. Biol.* 11, 592–595 (2001).
- Kikuchi A. Modulation of Wnt signalling by Axin and Axil. *Cytokine Growth Factor Rev.* 10, 255–265 (1999).
- Polakis P. Wnt signalling and cancer. *Genes Dev.* 14, 1837–1851 (2000).
- Shanmugam KS, Brenner H, Hoffmeister M, Chang-Claude J, Burwinkel B. The functional genetic variant Arg324Gly of frizzled-related protein is associated with colorectal cancer risk. *Carcinogenesis* 28, 1914–1917 (2007).
- **Provides the necessary information about the genetic variants within sFRP3 gene, the primers, RFLP pattern and the functional role of these SNPs.**
- Paez D, Gerger A, Zhang W *et al.* Association of common gene variants in the WNT/ β -catenin pathway with colon cancer recurrence. *Pharmacogenomics J.* 14, 142–150 (2014).
- Berndt SI, Huang WY, Yeager M, Weissfeld JL, Chanock SJ, Hayes RB. Genetic variants in frizzled-related protein (FRZB) and the risk of colorectal neoplasia. *Cancer Causes Control.* 20, 487–490 (2009).
- Frank B, Hoffmeister M, Klopp N, Illig T, Chang-Claude J, Brenner H. Single nucleotide polymorphisms in Wnt signaling and cell death pathway genes and susceptibility to colorectal cancer. *Carcinogenesis* 31, 1381–1386 (2010).
- Hirata H, Hinoda Y, Nakajima K *et al.* Wnt antagonist gene polymorphisms and renal cancer. *Cancer* 115, 4488–4503 (2009).
- **Evaluates all the genetic variants of DKK2, DKK3 and DKK4 genes along with sFRP4. It provided all the necessary information about the genotyping process and their significance in modulating the functionality of the protein.**

- 18 Alanazi MS, Parine NR, Shaik JP, Alabdulkarim HA, Ajaj SA, Khan Z. Association of single nucleotide polymorphisms in Wnt signaling pathway genes with breast cancer in Saudi patients. *PLoS ONE* 8, e59555 (2009).
- 19 Kim MS, Lee HN, Kim HJ, Myung SC. Single nucleotide polymorphisms in *DKK3* gene are associated with prostate cancer risk and progression. *Int. Braz. J. Urol.* 41, 869–897 (2015).
- 20 Yilmaz M, Donmez G, Kacan T *et al.* Significant association between polymorphisms of Wnt antagonist genes and lung cancer. *J. Investig. Med.* 63, 935–942 (2015).
- 21 Wu Z, Sun Y, Tang S *et al.* AXIN2 rs2240308 polymorphism contributes to increased cancer risk: evidence based on a meta-analysis. *Cancer Cell Int.* 15, 68 (2015).
- 22 Gunes EG, Pinarbasi E, Pinarbasi H, Silig Y. Strong association between lung cancer and the *Axin2* polymorphism. *Mol. Med. Rep.* 2, 1029–1035 (2009).
- **Provides the insight about the role of *Axin2* genetic variations in modulating lung cancer susceptibility.**
- 23 Kanzaki H, Ouchida M, Hanafusa H. Single nucleotide polymorphism of the *Axin2* gene is preferentially associated with human lung cancer risk in a Japanese population. *Int. J. Mol. Med.* 18, 279–284 (2006).
- 24 Mostowska A, Pawlik P, Sajdak S. An analysis of polymorphisms within the Wnt signaling pathway in relation to ovarian cancer risk in a Polish population. *Mol. Diagn. Ther.* 18, 85–91 (2014).
- 25 Liu D, Li L, Yang Y, Liu W, Wu J. The *Axin2* rs2240308 polymorphism and susceptibility to lung cancer in a Chinese population. *Tumor Biol.* 35, 10987–10991 (2014).
- 26 Ma C, Liu C, Huang P *et al.* Significant association between the *Axin2* rs2240308 single nucleotide polymorphism and the incidence of prostate cancer. *Oncol. Lett.* 8, 789–794 (2014).
- 27 Gunes EG, Pinarbasi E, Pinarbasi H. *Axin2* polymorphism and its association with astrocytoma in a Turkish population. *Mol. Med. Rep.* 3, 705–709 (2010).
- 28 Shimada T. Xenobiotic-metabolizing enzymes involved in activation and detoxification of carcinogenic polycyclic aromatic hydrocarbons. *Drug Metab. Pharmacokinet.* 21, 257–276 (2006).
- 29 Petrusis JR, Perdew GH. The role of chaperone proteins in the aryl hydrocarbon receptor core complex. *Chem. Biol. Interact.* 141, 25–40 (2002).
- 30 Morrow D, Qin C, Smith R, Safe S. Aryl hydrocarbon receptor-mediated inhibition of LNCaP prostate cancer cell growth and hormone-induced transactivation. *J. Steroid Biochem. Mol. Biol.* 88, 27–36 (2004).
- 31 Nakajima M, Yokoi T, Mizutani M, Kinoshita M, Funayama M, Kamataki T. Genetic polymorphism in the 5′-flanking region of human *CYP1A2* gene: effect on the *CYP1A2* inducibility in humans. *J. Biochem.* 125, 803–808 (1999).
- 32 Lin P, Chang H, Tsai WT *et al.* Overexpression of aryl hydrocarbon receptor in human lung carcinomas. Overexpression of aryl hydrocarbon receptor in human lung carcinomas. *Toxicol. Pathol.* 31, 22–30 (2003).
- 33 Tsay JJ, T'chou-Wong KM, Greenberg AK, Pass H, Rom WN. Aryl hydrocarbon receptor and lung cancer. *Anticancer Res.* 33, 1247–1256 (2013).
- 34 Valenta T, Hausmann G, Basler K. The many faces and functions of β -catenin. *EMBO J.* 31, 2714–2736 (2012).
- 35 Daugherty RL, Gottardi CJ. Phospho-regulation of β -catenin adhesion and signaling functions. *Physiology* 22, 303–309 (2007).
- 36 Gao C, Xiao G, Hu J. Regulation of Wnt/ β -catenin signaling by posttranslational modifications. *Cell Biosci.* 4, 1–20 (2014).
- 37 Schneider AJ, Branam AM, Peterson RE. Intersection of AHR and Wnt signaling in development, health, and disease. *Int. J. Mol. Sci.* 15, 17852–17885 (2014).
- **The primary report about the cross-linkage of two different pathways, namely, Wnt and AhR, on the basis of these facts there was a need to initiate such a study.**
- 38 Braeuning A, Kohle C, Buchmann A, Schwarz M. Coordinate regulation of cytochrome P450 1A1 expression in mouse liver by the aryl hydrocarbon receptor and the β -catenin pathway. *Toxicol. Sci.* 122, 16–25 (2011).
- 39 Procházková J, Kabátková M, Bryja V *et al.* The interplay of the aryl hydrocarbon receptor and β -catenin alters both AHR-dependent transcription and Wnt/ β -catenin signaling in liver progenitors. *Toxicol. Sci.* 122, 349–360 (2011).
- 40 Sodhi KK, Bahl C, Singh N, Sharma S, Behra D. Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in north Indians. *Future Oncol.* 11, 2159–2173 (2015).
- 41 Bin P, Leng S, Cheng J *et al.* Association of aryl hydrocarbon receptor gene polymorphisms and urinary 1-hydroxypyrene in polycyclic aromatic hydrocarbon-exposed workers. *Cancer Epidemiol. Biomarkers Prev.* 17, 1702–1708 (2008).
- 42 Ritchie MD, Moutsier AA. Multifactor dimensionality reduction for detecting gene–gene and gene–environment interactions in pharmacogenomics studies. *Future Med.* 6, 823–834 (2005).
- 43 Srivastava A, Sharma KL, Srivastava N, Misra S, Mittal B. Significant role of estrogen and progesterone receptor sequence variants in gallbladder cancer predisposition: a multi-analytical strategy. *PLoS ONE* 7, e40162 (2012).
- **Guides in finding the approach to analyze the high order interactions between Wnt and AhR pathway and their importance in predicting lung cancer predisposition.**

- 44 Luo J, Chen J, Deng ZL *et al.* Wnt signaling and human diseases: what are the therapeutic implications? *Lab. Invest.* 87, 97–103 (2007).
- 45 Marlowe JL, Fan Y, Chang X *et al.* The aryl hydrocarbon receptor binds to E2F1 and inhibits E2F1-induced apoptosis. *Mol. Biol. Cell.* 19, 3263–3271 (2008).
- 46 Bahl C, Singh N, Behera D, Sharma S. Association of polymorphisms in Dickkopf (DKK) gene towards modulating risk for lung cancer in north Indians. *Future Oncol.* 13, 213–232 (2016).
- 47 Aristizabal-pachon AF, Carvalho TI, Humberto H, Andrade J, Takahashi CS. *Axin2* polymorphisms, the β -catenin destruction complex expression profile and breast cancer susceptibility. *Asian Pac. J. Cancer Prev.* 16, 7277–7284 (2015).
- 48 Chen D, Tian T, Wang H *et al.* Association of human aryl hydrocarbon receptor gene polymorphisms with risk of lung cancer among cigarette smokers in a Chinese population. *Pharmacogenet. Genomics* 19, 25–34 (2009).
- 49 Kim JH, Kim H, Lee KY *et al.* Aryl hydrocarbon receptor gene polymorphisms affect lung cancer risk. *Lung Cancer* 56, 9–15 (2007).
- **Shows the functional significance of AhR variants on lung cancer risk and gives the primary information required for the same.**
- 50 Kawajiri K, Watanabe J, Eguchi H, Nakachi K, Kiyohara C, Hayashi SI. Polymorphisms of human Ahr receptor gene are not involved in lung cancer. *Pharmacogenetics* 5, 151–158 (1995).
- 51 Cauchi S, Stücker I, C  n  e S, Kremers P, Beaune P, Massaad-Massade L. Structure and polymorphisms of human aryl hydrocarbon receptor repressor (AhRR) gene in a French population: relationship with CYP1A1 inducibility and lung cancer. *Pharmacogenetics* 13, 339–347 (2003).
- 52 Cauchi S, Stücker I, Solas C *et al.* Polymorphisms of human aryl hydrocarbon receptor (*AhR*) gene in a French population: relationship with CYP1A1 inducibility and lung cancer. *Carcinogenesis* 22, 1819–1824 (2001).
- 53 Anttila S, Lei XD, Elovaara E *et al.* An uncommon phenotype of poor inducibility of CYP1A1 in human lung is not ascribable to polymorphisms in the *AHR*, *ARNT*, or *CYP1A1* genes. *Pharmacogenetics* 10, 741–775 (2000).
- 54 Anttila S, Tuominen P, Hirvonen A *et al.* CYP1A1 levels in lung tissue of tobacco smokers and polymorphisms of CYP1A1 and aromatic hydrocarbon receptor. *Pharmacogenetics* 11, 501–509 (2001).
- 55 Coscio A, Chang DW, Roth JA *et al.* Genetic variants of the Wnt signaling pathway as predictors of recurrence and survival in early-stage non-small cell lung cancer patients. *Carcinogenesis* 35, 1284–1291 (2014).
- 56 Chen M, Kamat AM, Huang M *et al.* High-order interactions among genetic polymorphisms in nucleotide excision repair pathway genes and smoking in modulating bladder cancer risk. *Carcinogenesis* 28, 2160–2165 (2007).
- 57 Wu X, Gu J, Grossman HB *et al.* Bladder cancer predisposition: a multigenic approach to DNA-repair and cell-cycle-control genes. *Am. J. Hum. Genet.* 78, 464–479 (2006).
- 58 Liu L, Wu C, Wang Y *et al.* Association of candidate genetic variations with gastric cardia adenocarcinoma in Chinese population: a multiple interaction analysis. *Carcinogenesis* 32, 336–342 (2011).
- 59 Kotnis A, Namkung J, Kannan S *et al.* Multiple pathway-based genetic variations associated with tobacco related multiple primary neoplasms. *PLoS ONE* 7, e30013 (2012).
- 60 Han S, Lv L, Wang X, Wang X, Zhao H. Association of *Axin2* and *MMP7* polymorphisms with non-small cell lung cancer in Chinese Han population. *Int. J. Clin. Exp. Pathol.* 9, 2253–2258 (2016).
61. Computational Genetics Laboratory. Epistasis and Human Health. www.epistasis.org

Genetic Variants in the Wingless Antagonist Genes (*sFRP*, *DKK*, and *Axin2*) Predict the Overall Survival and Prognosis of North Indian Lung Cancer Patients Treated with Platinum-Based Doublet Chemotherapy

Charu Bahl,¹ Navneet Singh,² Digambar Behera,² and Siddharth Sharma¹

Abstract

Background: To investigate the prognostic implication of genetic variants within the wingless (Wnt) antagonist genes (*DKK*, *sFRP*, and *Axin2*) in North Indian lung cancer patients.

Materials and Methods: A total of 212 subjects were genotyped using polymerase chain reaction–restriction fragment length polymorphism technique for 18 polymorphic sites in *DKK4*, *DKK3*, *DKK2*, *sFRP3*, *sFRP4*, and *Axin2*. Overall survival (OS) was estimated using the Kaplan–Meier survival analysis, and adjusted hazard ratio (HR) was obtained using the Cox regression method.

Results: It was observed that the unfavorable genotypes of the three *DKK2* variants collectively (rs447372, rs419558, and rs17037102) exhibited a highly decreased rate of death (adjusted HR=0.37, $p=0.03$). Adenocarcinoma (ADCC) patients carrying the heterozygous (CT) genotype for *DKK4* rs2073664 showed a better OS compared with wild genotype (log rank $p=0.01$). The two exonic variants (148 and 1386) of *Axin2* gene showed contrasting results, where the ADCC subjects having TT genotype for *Axin2* 148 showed a better prognosis (adjusted HR=0.48, $p=0.003$) and those with TT genotype for *Axin2* 1386 showed a poor prognosis in small-cell lung carcinoma patients (adjusted HR=2.33, $p=0.02$). The intronic *Axin2* 1712+19 variant on the other hand indicated a highly increased death risk in ADCC patients with GG genotype. Survival tree analysis depicted *DKK4* rs2073664 as the major contributor in predicting the survival of the lung cancer patients. Node 3 exhibited the lowest death rate (HR=0.04, $p=0.008$) and better median survival time (9 months vs. 3 months) when compared with reference node.

Conclusions: A cumulative effect of three variants of *DKK2* gene along with *DKK4* rs2073664 can jointly predict the survival as shown by tree analysis.

Keywords: interaction, lung cancer, polymorphisms, prognosis, survival, Wnt antagonists

Introduction

Lung cancer is characterized as multifactorial and complex carcinoma having diverse heterogeneity, which indicates the pathological understanding of this disease. Several genetic and molecular perturbations consequently lead to the formation of primary tumor of the lung and then its metastasis. Mechanisms underlying the aggressive and robust capacity of lung tumors to develop and spread are largely unknown.¹ Uncovering the genes and

pathways involved in this abrupt initiation and rapid spread of this disease can help in introducing the early diagnostic methods, treatment modulations, and also targeting new molecules that have hidden therapeutic potential. Despite numerous efforts, the average 5-year survival of lung cancer patients has not improved from 15%.² The clinical outcomes of lung cancer have been very poor compared with the other cancers, such as breast, prostate, and colorectal; this difference is attributed to variations in the biology of lung tumors.³

¹Department of Biotechnology, Thapar University, Patiala, India.

²Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER), Chandigarh, India.

Address correspondence to: Siddharth Sharma; Department of Biotechnology, Thapar University; Patiala, Punjab 147002, India
E-mail: siddharthsharma.phd@thapar.edu

The ontogeny of lung cancer cells has been recently related to the expression of those genes, which recapitulate the events taking place during embryonic development of lung tissue.⁴ Considering the role of the developmental pathways in lung carcinogenesis and progression has indicated the need of exploring the signal transduction cascades such as canonical wingless (Wnt) pathway, which has elucidated roles in maintenance of embryo, adult cells, and also other variety of processes such as cell polarity, differentiation, apoptosis, and survival.⁵ Another important aspect of the β -catenin-mediated Wnt signaling is that it is found to nurture and dictate the differentiation of cancer stem cells that have great tumorigenic capacity. Hyperactivity of β -catenin-mediated pathway when controlled by the optimum activity of the Wnt antagonists, such as secreted frizzled-related protein (sFRP), Dickkopf (DKK) protein, and Axin2 protein, has been found responsible for decreasing the dividing capacity of this stem cell population.^{6,7}

The localization of β -catenin from cytoplasm to nucleus in the presence of the Wnt ligand binding to the frizzled-related receptor (Fzd) and LDL-receptor-related protein (LRP) coreceptor leads to the transcription of downstream target genes, such as *matrix metalloproteins*, *cox-2*, *cyclin D1*, *VEGF*, and *c-myc*, through the transcription-regulating proteins T-cell factor (TCF)/lymphoid-enhancing factor (LEF).⁸ The level of β -catenin is strictly regulated in the cytoplasm by the degradation complex, which comprises of Dishevelled (Dsh), adenomatous polyposis coli, and Axin2 proteins. The phosphorylation of β -catenin is mediated by glycogen synthase kinase-3 β (GSK-3 β) present in the complex.⁹ Wnt antagonists control the transduction of this cascade at various subcellular levels, for example, sFRPs inhibit the Wnt signaling extracellularly by competing with the Fzd to interact with Wnt ligand as they both share a homology in structure.¹⁰ Similarly, DKK proteins regulate the transfer of signal by binding with the LRP coreceptor, and due this interaction, there is disruption of Wnt ligand/frizzled receptor/LRP complex.¹¹ Axin2 is the scaffold protein of the degradation complex required for the ubiquitination of the β -catenin. Chronic expression of Wnt cascade may lead to molecular alterations governing the process of lung pathogenesis along with other environmental factors.¹² The epigenetic silencing of the genes in this pathway has been observed to alter the level of expression of target genes in various tumors.¹³ Methylation of Wnt antagonists, such as sFRPs, has been elucidated in breast,¹⁴ hepatic,¹⁵ lung,¹⁶ and ovarian cancers.¹⁷ Similarly, the hypermethylation of the *DKK3* promoter has been observed in high percentage of lung tumors.¹⁸ The expression of *DKK2* gene is reported to decrease in melanoma and gastric cancer.¹⁹ There are evidences supporting the oncogenic role of *DKK3* gene in process of lung cancer progression, which is frequently inactivated in lung cancer.²⁰

Exploration of Wnt signaling pathway as a prognostic and predictive marker has been carried out in various malignancies, such as hematological malignancies,²¹ breast cancer,²² ovarian cancer,²³ pancreatic ductal adenocarcinoma (ADCC),²⁴ and esophageal squamous cell carcinoma (SQCC).²⁵ The involvement of the genetic variations within this developmental pathway has been evaluated in non-small-cell lung carcinoma (NSCLC) in Caucasians²⁶ and Korean ethnicities.²⁷ Another study has been performed in renal cancer patients where the polymorphic sites of negative regulators of Wnt pathway were

examined.²⁸ Keeping the above-cited facts in mind, the single-nucleotide polymorphisms (SNPs) within the various genes of the Wnt antagonist family, namely *sFRP3* (rs7775 and rs228236), *sFRP4* (rs1802073 and rs1802074), *DKK3* (rs2291599, rs3206824, and rs7391689), *DKK2* (rs447372, rs419558, and rs17037102), *DKK4* (rs2073664), and *Axin2* (148 C/T, 1365 G/A, 432 T/C, 956+16 A/G, 1386 C/T, 1712+19 G/A, and 2062 C/T), have been examined for their association with overall survival (OS) of North Indian lung cancer patients. Also, these SNPs were screened to find their ability to act as predictive and prognostic markers.

Methodology

Study population

In the present study, a total of 212 lung cancer subjects were recruited from the Department of Pulmonary Medicine, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh. Informed consent was obtained from each patient, and a questionnaire was filled by every patient with the help of a trained personnel. The questionnaire provided all the details including the epidemiological information of the volunteers. There was no age, sex, cancer stage, and histology restrictions in selecting the subjects who were enrolled for the study. The only exclusion criterion was the presence of history of any kind of tumor. Five microliters of blood was collected from each volunteer in Ethylene diamine tetra acetic acid (EDTA)-coated Vacutainer. All the subjects were histopathologically classified as SQCC, ADCC, and small-cell lung carcinoma (SCLC). This study has been reviewed and approved by the Institutional Ethics Committee of PGIMER. The clinical details of the patients such as tumor, node, and metastases (TNM) staging, smoking habits, Karnofsky performance status (KPS), and Eastern Cooperative Oncology Group performance status (ECOG PS) were collected prospectively from the medical records of the hospital. Telephonic follow-up of the patients was carried out after six sessions of chemotherapy to find out the OS of the patients. The patients were followed for 3 years from the date of their diagnosis.

Chemotherapeutic regimen

Doublet platinum-based chemotherapeutic drug combination was administered to all the lung cancer patients enrolled in the present study. This combination composed of different drugs such as docetaxel, irinotecan, or pemetrexed along with a platinum-based drug, such as cisplatin or carboplatin. These drugs were administered intravenously at a fixed dosage according to the surface area of the body. The dose of drug that was administered to the patients was as follows: 75 mg/m² docetaxel or 500 mg/m² pemetrexed or 75 mg/m² irinotecan was given as 1 h infusion, which was then followed by cisplatin (70 mg/m²) infused for over 3 h. Standard folate and vitamin B12 formulations were given to each patient as supplements. Certain drugs such as dexamethasone, granisetron, and ranitidine were also administered as a part of premedication routine. Tumor response assessment was carried out after four cycles of chemotherapy. In certain cases, due to high toxicity and massive disease progression reported by radiological assessment, chemotherapy procedure was stopped even before the completion of four cycles. The patients showing a good objective

response were administered with additional two cycles of chemotherapy.

Follow-up and response assessment

Study subjects were followed up for survival through telephonic conversations. The subjects were called upon telephonically every 2 months until the death of the patient or till the end of the study period was over. Primarily, the OS was assessed by considering the survival time from the date of diagnosis till the date of last follow-up or death due to other reasons. Tumor response was evaluated after every two cycles of chemotherapy according to the Response Evaluation Criteria in Solid Tumor Group (RECIST).

DNA extraction and genotyping

Genomic DNA was extracted from each sample using the detailed procedure as described by Sodhi et al.²⁹ Genotyping for each SNP was carried out using polymerase chain reaction (PCR)–restriction fragment length polymorphism technique. The genotyping for two *sFRP3* gene variants was performed as previously detailed by Shanmugam et al.³⁰ In *sFRP4* variants (rs1802073 and rs1802074), *DKK3* (rs2291599, rs3206824, and rs7391689), *DKK2* (rs447372, rs419558, and rs17037102), and *DKK4* (rs2073664) polymorphic sites, the genotyping was carried in a similar manner as reported by Hirata et al.²⁸ In genetic variants of *Axin2* gene, namely 148 C>T, 1365 G>A, 432 T>C, 1712+19 G>A, 956+16 A>G, 1386 C>T, and 2062 C>T, the protocol described by Gunes et al.³¹ was followed to find out the genotype of the subjects.

The PCR (25 μ L) was used to amplify the desired fragment, and it composed of 1 $\times$ PCR buffer, 1.5 mM MgCl₂, 0.5 μ M of both forward and reverse primers, 200 μ M of each deoxynucleotide triphosphates, 100 μ g/mL bovine serum albumin, 1 U Taq polymerase (DyNAzyme II DNA Polymerase; Thermo Scientific), and $\sim$ 200 ng DNA. The amplified product was digested with their respective restriction enzymes as described above. The digested products were resolved on either agarose gel or polyacrylamide gel to find out the restriction patterns. Scoring of the patterns was performed to determine the genotypic status of the sample. The genotyping of 15% samples was performed twice to check the reproducibility of the results, and it was found to be 100%.

Statistical analyses

Genotypic and allelic frequencies for each polymorphic site under study were analyzed to check whether they have followed the Hardy–Weinberg equilibrium ($p^2 + 2pq + q^2 = 1$) by using the chi-square test. For finding out the role of these polymorphisms in predicting the survival of the lung cancer patients, the survival time (in months) of each patient was calculated. Survival time can also be censored when there is a follow-up time, but the event which is death of the patient in this case has not yet occurred or is not known to have occurred. The concept of Kaplan–Meier was used to find out the median survival time (MST) and survival curve for each category in all the polymorphic sites. Univariate analysis by log-rank test was used as a statistical method to compare the MST of patients based on the genotype of that particular variant. To find out whether these polymorphic variants can

act as independent predictor variable, multivariate Cox proportional hazard was used, where hazard ratio (HR) along with 95% confidence interval (CI) adjusted for various covariates such as age, sex, smoking, cancer stage, KPS, ECOG PS, and histology were obtained. Cox proportional hazards model is analogous to a multiple regression model and enables the difference between survival times of particular groups of patients to be tested while allowing for other factors. In this model, the response (dependent) variable is the “hazard.” The HR is the probability of dying that patients have survived up to a given point in time or the risk for death at that moment. The Kaplan–Meier and Cox regression methods were also used in various subgroups stratified on the basis of sex, smoking, histology, and performance status.³² All *p*-values were two-sided, and a *p*-value of <0.05 was considered statistically significant.

Multiple comparison corrections were also carried out by false discovery rate (FDR) method using online software based on Benjamini and Hochberg approach (<http://sdmproject.com/utilities/?show=FDR>). The adjusted *p*-values are mentioned, and the FDR-adjusted *p*-value of <0.1 was considered as significant. All the statistical analyses were performed with MedCalc version 16.8.1 (MedCalc Software, Ostend, Belgium) and SPSS version 20.0. Furthermore, high-order gene interactions were taken into account by using the recursive partitioning approach. This approach produces the initial split taking into consideration the primary factor affecting the OS of patients. The software used for making the decision tree is STREE program (<http://c2s2.yale.edu/software/stree>). Log-rank method was employed as the splitting criteria. The recursive method keeps on producing the offspring nodes till statistically insignificant split occurs. The resulting tree was binary, and each terminal node represented a subgroup of patients carrying a particular genotypic combination and hence shows a varied survival time and prognosis. HRs and their 95% CIs for each terminal node were calculated using the multivariate Cox proportional hazard model adjusted for sex, age, histology, KPS, ECOG PS, smoking, and stage of tumor.

Results

Demographic and clinical parameters

Both demographic and clinical parameters are summarized in Table 1. The study population was characterized for various epidemiological variables, such as age, gender, and smoking habits. The mean age of the study subjects was 57.87 ± 10.81 years. Of 212 subjects, 86.8% were males and the remaining 19.4% were females. As smoking is an established cause for the occurrence of lung cancer, so the majority of sample population composed of smokers (80.6%), whereas 19.4% were characterized as nonsmokers. Different clinical parameters such as histology, TNM staging, performance status, and chemotherapeutic regimen are also listed in Table 1. The patients are classified on the basis of their tumor histology into ADCC, SQCC, and SCLC. Considering the TNM staging data, maximum proportion of lung cancer patients were diagnosed with stage III of lung cancer. To evaluate the clinical performance status of the patients, two criteria namely ECOG PS and KPS were used. One hundred thirty-three (62.7%) patients had KPS of 60–80 and 74 (34.9%) patients harbored an ECOG PS score between 1 and 2. Furthermore, the sample population was

TABLE 1. DISTRIBUTION OF DEMOGRAPHIC AND CLINICAL PARAMETERS

Variable	Total (N)	Cases, n (%)	HR (95% CI)	p
Mean $\pm$ SD	212	57.87 $\pm$ 10.81		
Range		28–86		
Male	212	184 (86.8)	0.93 (0.61–1.40)	0.74
Female		28 (13.2)	Reference	
Smokers	212	171 (80.6)	1.04 (0.72–1.49)	
Nonsmokers		41 (19.4)	Reference	0.83
Histology				
ADCC	212	72 (34)		
SCLC		51 (24)		
SQCC		86 (40)		
Others		3 (2)		
TNM staging				
I	212	2 (0.9)		
II		6 (2.8)		
III		105 (49.5)	Reference	
IV		89 (41.9)	1.23 (0.92–1.65)	0.16
Others		10 (4.7)		
Overall survival				
Dead	212	181 (85.3)		
Alive		31 (14.6)		
Performance status				
KPS (30–50)	212	7 (3.3)	Reference	
KPS (60–80)		133 (62.7)	0.66 (0.29–1.49)	0.32
KPS (90–100)		72 (33.9)	0.80 (0.52–1.21)	0.30
ECOG PS (0)	212	31 (14.6)	Reference	
ECOG PS (1)		74 (34.9)	0.66 (0.42–1.03)	0.07
ECOG PS (2)		74 (34.9)	0.84 (0.68–1.05)	0.14
ECOG PS (3 and 4)		33 (15.5)	0.94 (0.79–1.12)	0.53
Chemotherapy regimen				
Docetaxel/cisplatin	156	55 (35.2)		
Irinotecan/cisplatin		41 (26.2)		
Paclitaxel/cisplatin		12 (7.6)		
Pemetrexed/cisplatin		40 (25.6)		
Gemcitabine/cisplatin		1 (0.6)		
Others		7		

ADCC, adenocarcinoma; CI, confidence interval; ECOG PS, Eastern Cooperative Oncology Group performance status; HR, hazard ratio; KPS, Karnofsky performance status; SCLC, small-cell lung carcinoma; SQCC, squamous cell carcinoma; TNM, tumor, node, and metastasis.

stratified on the basis of their chemotherapeutic regimen; highest percentage of patients were treated with docetaxel and cisplatin doublet combination. SCLC patients were given irinotecan and cisplatin doublet combination. The HR was also calculated for various parameters, such as gender, smoking, TNM staging, and performance status. It was evident that none of these factors could act as an independent predictor of survival for lung cancer patients. Lung cancer patients having an ECOG PS score of 1 were found to have a better survival (HR = 0.66) compared with those with ECOG PS score of 0, but the *p*-value of 0.07 was not found to be statistically significant.

Association of Wnt antagonist polymorphism and OS of lung cancer patients

Table 2 summarizes the univariate and multivariate analyses for the various polymorphic sites of *DKK*, *sFRP3*, *sFRP4*, and *Axin2* genes in relation to the OS of lung cancer patients. The MST was calculated using the Kaplan–Meier statistics, and survival curves were obtained for each ge-

notype. Genotypic influence on the survival of the patients was examined by using log-rank test, which compares the MST of different groups considering the wild genotype as reference group. Univariate analysis gave the probability of death (HR), whereas multivariate Cox regression analysis provided the adjusted HR. The HR was adjusted using the Cox regression model for various covariates such as age, sex, smoking, histology, stage, KPS, and ECOG PS to find whether a particular SNP can act as an independent predictor and prognostic factor. The data showed that the MST of lung cancer patients having heterozygous (*CT*) genotype for *DKK4* rs2073664 was higher compared with wild (*CC*) and mutant (*TT*) genotypes. Better OS was observed in heterozygous patients both in univariate (HR = 0.74) and multivariate (adjusted HR = 0.75) analyses.

None of the three variants (rs2295199, rs3206824, and rs7396187) of *DKK3* gene showed any significant role in predicting the survival of lung cancer patients. However, in *DKK3* rs3206824 polymorphism, lung cancer subjects harboring the heterozygous (*CT*) genotype did show a higher MST (8.8 months, log-rank *p*-value = 0.07) with HR of 0.74;

TABLE 2. UNIVARIATE AND MULTIVARIATE ANALYSES FOR WINGLESS ANTAGONISTS

	Cases, n (N=212)	Death, n (N=181)	Alive, n (N=31)	Univariate analysis			Multivariate analysis		
				MST (months)	Log- rank p	Unadjusted HR ^a (95% CI)	Adjusted HR ^b (95% CI)	p	FDR- adjusted p
<i>DKK4</i> rs2073664 C/T									
CC	112	101	11	6.2		Reference	Reference		
CT	96	77	19	7.5	0.06	0.75 (0.55–1.01)	0.74 (0.55–1.01)	0.06	0.66
TT	4	3	1	6.5	0.74	0.97 (0.51–1.60)	1.01 (0.56–1.85)	0.95	0.95
CT + TT	100	80	20	7.5	0.07	0.76 (0.57–1.02)	0.79 (0.58–1.06)	0.12	0.76
<i>DKK3</i> rs2291599 C/T									
CC	133	114	19	6.4		Reference	Reference		
CT	63	52	11	7.5	0.52	0.89 (0.64–1.24)	0.92 (0.65–1.29)	0.64	0.76
TT	16	15	1	6.7	0.98	1 (0.76–1.31)	1.07 (0.81–1.43)	0.60	0.9
CT + TT	79	67	12	7.3	0.55	0.91 (0.67–1.23)	0.96 (0.70–1.30)	0.96	0.96
<i>DKK3</i> rs3206824 G/A									
GG	140	121	19	5.8		Reference	Reference		
GA	66	55	11	8.8	0.07	0.74 (0.54–1.02)	0.76 (0.54–1.06)	0.10	0.66
AA	6	5	1	4.9	0.91	1.02 (0.65–1.60)	1.07 (0.67–1.72)	0.76	0.91
GA + AA	72	60	12	8.3	0.11	0.78 (0.58–1.05)	0.81 (0.58–1.12)	0.20	0.76
<i>DKK3</i> rs7396187 G/C									
GG	149	128	21	6.3	0.74	Reference	Reference		
GC	62	52	10	7.8		0.89 (0.65–1.23)	0.89 (0.64–1.23)	0.49	0.76
CC	1	1	0	16.7		—	—		
GC + CC	63	53	10	8	0.47	0.89 (0.65–1.23)	0.87 (0.63–1.20)	0.87	0.95
<i>DKK2</i> rs447372 A/G									
AA	105	92	13	5.9		Reference	Reference		
AG	87	74	13	8.2	0.22	0.82 (0.61–1.12)	0.77 (0.56–1.05)	0.11	0.66
GG	20	15	5	10.3	0.30	0.86 (0.65–1.13)	0.81 (0.61–1.08)	0.15	0.32
AG + GG	107	89	18	8.3	0.12	0.79 (0.59–1.06)	0.73 (0.54–0.99)	0.04	0.60
<i>DKK2</i> rs17037102 G/A									
GG	152	127	25	6.7	0.78	Reference	Reference		
GA	59	53	6	6.4		1.11 (0.81–1.54)	1.09 (0.78–1.53)	0.59	0.76
AA	1	1	0	10.6		—	—		
GA + AA	60	54	6	6.8	0.31	1.17 (0.84–1.63)	1.13 (0.81–1.58)	0.45	0.76
<i>DKK2</i> rs419558 C/T									
CC	143	125	18	6.3		Reference	Reference		
CT	69	56	13	7.3	0.38	0.87 (0.64–1.18)	0.82 (0.60–1.13)	0.24	0.76
<i>sFRP3</i> rs7775 C/G									
CC	167	147	20	6.4		Reference	Reference		
CG	33	26	7	7.3	0.41	0.83 (0.55–1.27)	0.84 (0.55–1.28)	0.44	0.76
GG	12	8	4	9.4	0.13	0.75 (0.53–1.08)	0.74 (0.51–1.06)	0.10	0.32
CG + GG	45	34	11	8.3	0.12	0.73 (0.53–1.05)	0.73 (0.50–1.07)	0.11	0.76
<i>sFRP3</i> rs288326 C/T									
CC	179	153	26	6.7		Reference	Reference		
CT	20	19	1	6.8	0.48	1.18 (0.71–1.97)	1.30 (0.80–2.12)	0.28	0.76
TT	13	9	4	5.4	0.68	0.93 (0.66–1.30)	0.98 (0.69–1.39)	0.94	0.9
CT + TT	33	28	5	6.1	0.74	1.06 (0.70–1.61)	1.21 (0.79–1.83)	0.37	0.76
<i>sFRP4</i> rs1802073 C/A									
CC	34	28	6	6.1		Reference	Reference		
CA	118	99	19	7.5	0.54	0.87 (0.57–1.33)	0.89 (0.57–1.38)	0.22	0.76
AA	60	54	6	5.8	0.58	1.06 (0.84–1.33)	1.20 (0.92–1.57)	0.16	0.32
CA + AA	178	153	25	6.7	0.73	0.93 (0.61–1.41)	1.04 (0.68–1.59)	0.83	0.95
<i>sFRP4</i> rs1802074 G/A									
GG	28	24	4	6.4		Reference	Reference		
GA	174	149	25	6.8	0.72	0.92 (0.60–1.41)	0.93 (0.58–1.48)	0.77	0.81
AA	10	8	2	5.7	0.35	0.82 (0.55–1.23)	0.74 (0.45–1.21)	0.74	0.9
GA + AA	184	157	27	6.8	0.57	0.88 (0.56–1.38)	0.85 (0.53–1.36)	0.51	0.76
<i>Axin2</i> 148 C/T									
CC	57	47	10	5.5		Reference	Reference		
CT	109	92	17	6.7	0.57	0.90 (0.571.42)	0.86 (0.52–1.44)	0.38	0.76
TT	46	42	8	9.4	0.22	0.84 (0.64–1.18)	0.76 (0.55–1.05)	0.09	0.32
CT + TT	155	134	21	7.3	0.81	0.96 (0.68–1.34)	0.91 (0.64–1.29)	0.60	0.81

(continued)

TABLE 2. (CONTINUED)

	Cases, n (N=212)	Death, n (N=181)	Alive, n (N=31)	MST (months)	Univariate analysis		Multivariate analysis		
					Log-rank p	Unadjusted HR ^a (95% CI)	Adjusted HR ^b (95% CI)	p	FDR-adjusted p
<i>Axin2</i> 1365 G/A									
GG	188	160	28	7.3		Reference	Reference		
GA	20	17	3	6.7	0.82	1.08 (0.54–2.13)	1.27 (0.61–2.63)	0.52	0.76
AA	4	4	0	6.1	0.41	1.23 (0.75–2.02)	1.50 (0.89–2.50)	0.12	0.32
GA + AA	24	21	3	6.4	0.18	1.35 (0.81–2.26)	1.33 (0.82–2.15)	0.23	0.76
<i>Axin2</i> 956+16 A/G									
AA	179	153	26	6.4	0.60	Reference	Reference		
AG	31	26	5	7.3		0.89 (0.59–1.35)	0.90 (0.59–1.37)	0.62	0.76
GG	2	2	0	10		—	—		
AG + GG	33	28	5	8.1	0.49	0.87 (0.59–1.27)	0.86 (0.57–1.29)	0.47	0.76
<i>Axin2</i> 432 T/C									
TT	206	176	30	6.4		Reference	Reference		
TC	6	5	1	16.4	0.26	0.60 (0.24–1.46)	0.64 (0.26–1.58)	0.34	0.76
<i>Axin2</i> 1712+19 G/T									
TT	60	48	12	6.7		Reference	Reference		
GT	144	125	19	7.2	0.55	1.10 (0.79–1.54)	1.13 (0.80–1.58)	0.47	0.76
GG	8	8	0	4.3	0.29	1.22 (0.84–1.77)	1.47 (0.98–2.20)	0.06	0.32
GT + TT	152	133	19	7.1	0.45	1.13 (0.82–1.56)	1.17 (0.84–1.64)	0.34	0.76
<i>Axin2</i> 2062 C/T									
CC	178	151	27	7.2		Reference	Reference		
CT	34	30	4	6.2	0.69	1.08 (0.72–1.61)	1.07 (0.72–1.60)	0.71	0.79
<i>Axin2</i> 1386 C/T									
CC	71	61	10	7.3		Reference	Reference		
CT	113	97	16	6.7	0.88	1.02 (0.74–1.41)	1.03 (0.74–1.44)	0.81	0.81
TT	28	23	5	6.1	0.81	0.97 (0.76–1.23)	1.10 (0.85–1.44)	0.43	0.73
CT + TT	141	120	21	6.4	0.98	0.99 (0.73–1.35)	0.97 (0.71–1.33)	0.89	0.95

^aUnadjusted HR for the Kaplan–Meier analysis.

^bHR adjusted for age, sex, smoking, histology, cancer stage, KPS, and ECOG PS; FDR-adjusted *p*-value. FDR, false discovery rate; MST, median survival time.

although on adjusting with various covariates, no significant change in death probability was observed (adjusted HR = 0.75, *p* = 0.10). The three polymorphic sites of *DKK2* gene, that is, rs447372, rs17037102, and rs419558, also exhibited a lack of association with the prognosis of North Indian lung cancer patients. On the contrary, the polymorphic site of *DKK2*, that is, rs447372, showed a good prognostic effect in patients carrying AG + GG genotype; however, the *p*-value was on the borderline of significance (HR = 0.73, *p* = 0.04), which after adjusting for multiple correction increased to *p* = 0.60.

Furthermore, in *sFRP3* gene of the two SNPs studied, none of them showed significant contribution in predicting the survival of the patients. For *sFRP3* rs7775, the subjects having the mutant (GG) genotype had a better survival time (9.4 months) compared with those carrying the wild-type (6.4 months) genotype. But, both univariate and multivariate analyses revealed insignificant HR. Similarly, both polymorphic variations of *sFRP4* gene (rs1802073 and rs1802074) reported a lack of association with the OS of lung cancer patients. Although in rs1802073 the mutant (AA) genotype did reveal a lower MST (5.8 months), the difference was not found to be significant.

Finally, the seven variants of *Axin2* gene under study also negated the role of Wnt antagonists sequence variations in predicting the prognosis of the patients. None of the SNPs

showed any contribution in modulating the MST of the particular genotypic group. In *Axin2* 148, the mutant (TT) genotype had a better survival time (9.4 months vs. 5.5 months, log-rank *p* = 0.39) compared with wild (CC) genotype, which was however not significant. A similar trend was observed in *Axin2* 956+16 polymorphic site, where the mutant (GG) genotype had a better MST (10 months) compared with wild (AA) genotype. For *Axin2* 1712+19, a decreased MST (4.3 months) was observed in subjects carrying GG genotype compared with those having TT genotype. The HR (1.22, *p* = 0.29) in these subjects increased to 1.47 (*p* = 0.06) after adjusting for covariates using the Cox regression analysis. Hence, a slight increase in death probability was observed.

Association of Wnt antagonist polymorphism and OS according to the histological subtype of lung cancer patients

To evaluate the relevance of these polymorphic sites in predicting the survival of patients in different subgroups stratified on the basis of histology, the association of Wnt antagonist variants with survival of patients was analyzed and data are presented in Supplementary Tables S1–S3 (Supplementary Data are available online at www.liebertpub.com/cbr).

TABLE 3. ASSOCIATION OF COMBINATION OF UNFAVORABLE GENOTYPES WITH OVERALL SURVIVAL

No. of unfavorable genotypes	Cases (N)	MST (months)	Unadjusted HR ^a (95% CI)	Log-rank p	Adjusted HR ^b (95% CI)	p	FDR-adjusted p
DKK2							
0	51	5.5	Reference		Reference		
2 (rs419558 and rs17037102)	15	6.8	1.03 (0.55–1.92)	0.91	0.64 (0.31–1.33)	0.23	0.23
2 (rs447372 and rs17037102)	22	9	0.80 (0.48–1.35)	0.43	0.58 (0.31–1.08)	0.09	0.12
2 (rs419558 and rs447372)	47	9.6	0.76 (0.49–1.18)	0.22	0.53 (0.33–0.85)	0.01	0.04
3	9	9	0.72 (0.35–1.47)	0.41	0.37 (0.14–0.95)	0.03	0.06
DKK3							
0	99	5.5	Reference		Reference		
2 (rs3206824 and rs7396187)	38	9.2	0.77 (0.52–1.14)	0.22	0.81 (0.53–1.24)	0.34	0.45
2 (rs3206824 and rs2291599)	48	9.2	0.73 (0.50–1.05)	0.11	0.74 (0.49–1.11)	0.15	0.45
2 (rs7396187 and rs2291599)	48	7.3	0.89 (0.62–1.30)	0.57	0.90 (0.61–1.33)	0.62	0.62
3	33	10.1	0.75 (0.50–1.14)	0.21	0.79 (0.50–1.24)	0.31	0.45
sFRP3							
0	137	6.4	Reference		Reference		
2	4	6.08	0.89 (0.29–2.41)	0.76	0.76 (0.23–2.47)	0.65	0.65
sFRP4							
0	5	5.3	Reference		Reference		
2	155	6.7	0.75 (0.24–2.35)	0.58	0.95 (0.33–2.73)	0.93	0.93

The bold values represent the significant p -values, $p < 0.05$.

^aUnadjusted HRs for the Kaplan–Meier analysis.

^bHR adjusted for age, sex, histology, smoking, cancer stage, KPS, ECOG PS. 0: wild genotype for all SNPs of a gene; 2: variant (hetero + mutant) genotype of any two variants; 3: variant (hetero + mutant) genotype of all the three variants; FDR-adjusted p -value.

sFRP, secreted frizzled-related protein; SNP, single-nucleotide polymorphism.

In this study, as shown in Supplementary Table S1, ADCC patients carrying the heterozygous (CT) genotype for *DKK4* rs2073664 showed a better OS compared with those with wild genotype (8.8 months vs. 6.3 months; Figure 1A). The HR obtained from both univariate and multivariate analyses was 0.53 with significant p -values of 0.01 and 0.03 (adjusted $p = 0.1$ and 0.2). Of the three *DKK3* polymorphic sites, the two SNPs, namely rs2291599 and rs7396187, were found to play no role in predicting the prognosis of patients in any of the histological subtypes. Only SCLC patients who were harboring the mutant (AA) genotype for *DKK3* rs3206824 had a lower MST compared with the lung cancer patients with wild-type genotype (1.8 months vs. 8.0 months) and exhibited a three-fold increased death ratio (adjusted HR = 3.45, $p = 0.03$; adjusted $p = 0.12$) when compared with the SCLC subjects carrying the wild (GG) genotype.

In the three SNPs of *DKK2* gene, it was observed that only ADCC patients having heterozygous (CT) genotype for *DKK2* rs419558 showed a better MST compared with ADCC patients carrying wild-type (CC) genotype (8.3 months vs. 6.2 months, log-rank $p = 0.11$), as illustrated in Figure 1B. However, on adjusting with various covariates, the HR was found to be 0.50 ($p = 0.03$, adjusted $p = 0.10$). Hence, the heterozygous ADCC patients showed a good prognosis compared with ADCC patients with wild-type (CC) genotype. The other two variants of *DKK2*, that is, rs447372 and rs170371, did not contribute toward predicting the survival of lung cancer patients.

Of the two variants of *sFRP3* gene, rs7775 exhibited a penchant in predicting the survival of SCLC patients only. SCLC patients having heterozygous (CG) genotype showed a poor survival time when compared with wild-type (CC) genotype (1.5 months vs. 7.2 months, log-rank $p = 0.01$), as shown in Figure 1D. Death rate was found to be two times higher in these subjects (adjusted HR = 2.03, $p = 0.05$; ad-

justed $p = 0.2$). However, *sFRP4* genetic variants were found to play a negligible role in prognosis of patients having any of the histological subcategory, as shown in Supplementary Table S2.

Seven *Axin2* variants were analyzed for finding their relevance in predicting the prognosis of lung cancer patients; the data are summarized in Supplementary Table S3. For *Axin2* 148 C/T, ADCC patients having TT genotype showed a better survival compared with ADCC patients having wild-type (CC) genotype (9.4 months vs. 5.7 months), as shown in Figure 1E. The HR obtained from univariate analysis was not statistically significant (HR = 0.63, log-rank $p = 0.19$), but after applying the Cox regression, the death ratio (0.48) decreased and was highly significant ($p = 0.003$; adjusted $p = 0.006$). Furthermore, for *Axin2* 1712 + 19, it was found that ADCC patients possessing mutant GG genotype had a very poor MST compared with subjects with TT genotype (0.27 months vs. 8.3 months, $p < 0.0001$), as illustrated in Figure 1C. The adjusted HR depicted four-fold higher death probability (adjusted HR = 4.14, $p = 0.0003$; adjusted $p = 0.001$) in these patients when compared with those with TT genotype. SCLC patients having heterozygous (CT) genotype for *Axin2* 1386 showed a twofold higher death risk (adjusted HR = 2.33, $p = 0.02$; adjusted $p = 0.08$) with reference to subjects with wild-type SCLC genotype.

Association of different unfavorable genotypic combinations of *DKK2*, *DKK3*, *sFRP3*, and *sFRP4* gene with OS of lung cancer patients

To find out the cumulative effect of the polymorphic variants of each gene, different genotypic combination of unfavorable genotypes of the Wnt antagonist polymorphic variants were analyzed for predicting the OS of lung cancer patients, as demonstrated in Table 3. In *DKK2* gene, the

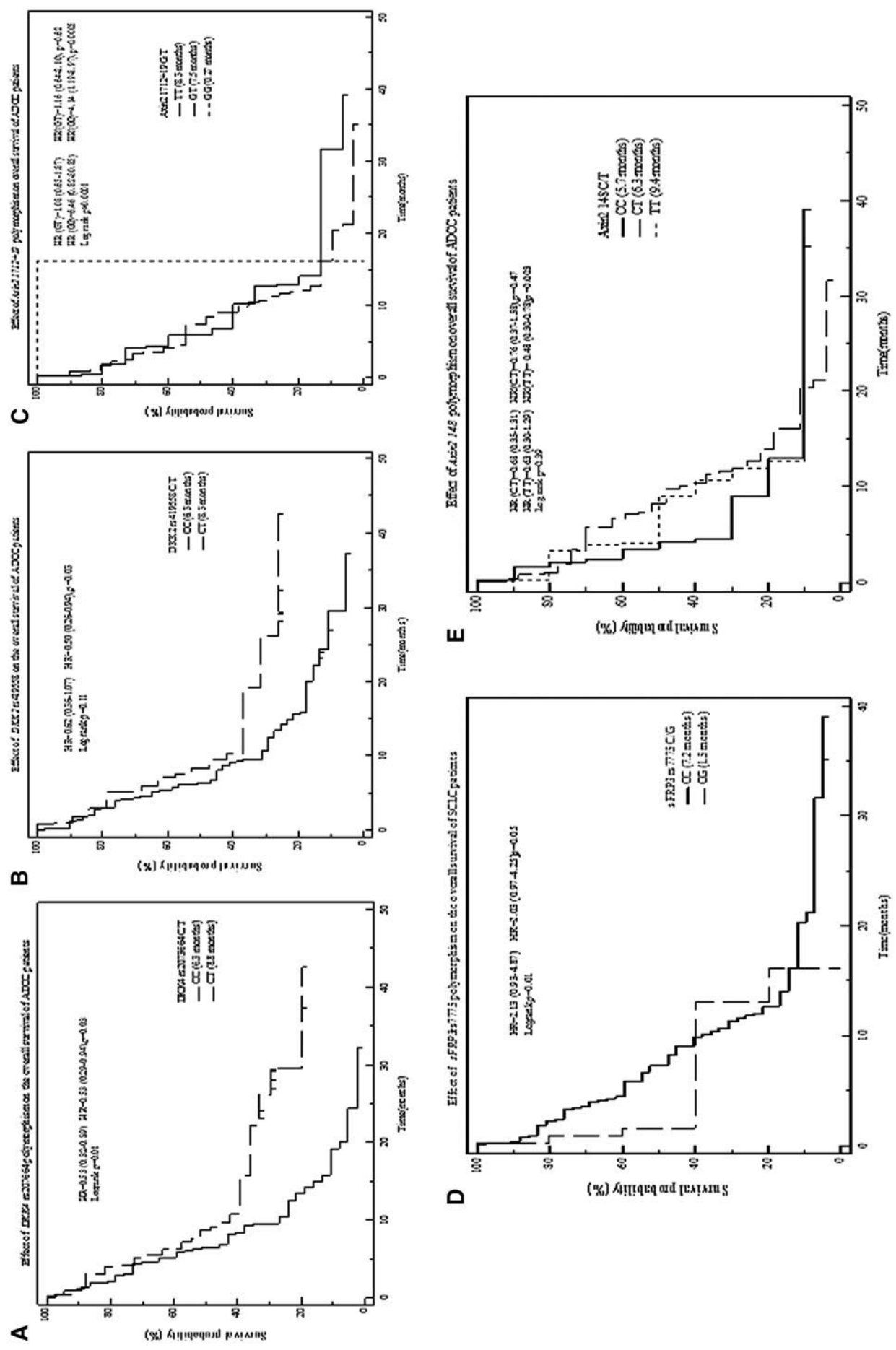


FIG. 1. Kaplan-Meier survival curves. (A) Effect of *DKK4* rs2073664 on the overall survival of ADCC patients. (B) Effect of *DKK2* rs419558 on the overall survival of ADCC patients. (C) Effect of *Axin2* 1712+12 on the overall survival of ADCC patients. (D) Effect of *sFRP3* rs7715 on the overall survival of ADCC patients. (E) Effect of *Axin2* 148 on the overall survival of ADCC patients. ADCC, adenocarcinoma; SCLC, small-cell lung carcinoma; sFRP, secreted frizzled-related protein.

genotypic combination comprising the unfavorable genotypes for both rs419558 and rs447372 SNPs showed a better survival time (9.6 months vs. 5.5 months) compared with the patients lacking any unfavorable genotypes for the three variants. The HR obtained from univariate analysis was not statistically significant (log-rank $p=0.22$), but when it was adjusted for several covariates, the death rate was found to decrease significantly (adjusted HR=0.53, $p=0.01$; adjusted $p=0.04$). Patients harboring all the three unfavorable genotypes for *DKK2* variants showed a significantly better prognosis (adjusted HR=0.37, $p=0.03$; adjusted $p=0.06$) during multivariate analysis. In *DKK3* gene, none of the combination of unfavorable genotype showed any cumulative effect on the survival of lung cancer patients. A similar trend was also observed in *sFRP3* and *sFRP4* genetic variants where none of the combination of unfavorable genotypes was found to act as an independent predictor of survival. Several different combinations of unfavorable genotypes were also examined for SNPs of *Axin2* gene, but none of them showed any significant association with prognosis of the patients (data not shown).

Association of Wnt antagonist polymorphism with treatment response for different chemotherapeutic regimens and also their correlation with objective response

In the present study, the effect of genotypic differences was evaluated on the treatment response in patients treated with different double platinum-based chemotherapeutic regimen, such as docetaxel/cisplatin, irinotecan/cisplatin, and pemetrexed/cisplatin. However, none of the polymorphic SNPs of the Wnt pathway exhibited any significant association with the prognosis of the patients administered with any of the above-mentioned chemotherapeutic regimens (data not shown). Furthermore, the correlation of these genetic variants with objective response of the patients was investigated after six cycles of chemotherapy. Patients were divided into four categories on the basis of the objective response, that is, complete response, partial response, stable disease, and progressive disease. However, none of the variants showed any significant association with the any of the response category (data not shown).

High-order gene–gene interactions for prognosis of lung cancer patients using survival tree analysis

Multiple interaction studies including different genes involved in the Wnt signaling pathway are a new dimension to find out the high-order parametric interactions taking place within them and effect of these interactions on OS of the patients. In the present study, the recursive partitioning approach was used to create a survival tree, which depicts the involvement of different SNPs in prognosis. The tree obtained by STREE software is illustrated in Figure 2A. Of all the algorithms available, log-rank method was used to create the tree. A total of nine terminal nodes were observed. The most significant factor contributing to the prognosis of patients was *DKK4* rs2073664, as it is root node. The terminal node 7 with shortest MST (3 months) was taken as a reference. MST corresponding to each terminal node is shown in Figure 2A. Also, the HR for each node is mentioned. Figure 2B shows the comparison of survival curves for different nodes. It depicts a significant difference between the MST of different nodes

when compared with reference (log-rank $p<0.0001$). Supplementary Table S4 summarizes the results of the Cox regression analysis for each terminal node. For the Cox regression analysis, the HR was adjusted for various covariates, such as age, sex, smoking, histology, ECOG PS, KPS, and stage. All the nodes showed a significantly good prognosis for lung cancer patients except terminal node 4. Node 3 exhibited the lowest death rate (HR=0.04, $p=0.008$) and better MST (9 months vs. 3 months) when compared with reference node. A similar trend was observed in all other nodes showing good prognostic effect.

Discussion

Wnt cascade has been increasingly gaining importance due to its influence on a myriad of cellular and biological processes. This evolutionary conserved pathway has an established role in embryonic development and is also critically involved in maintaining both normal and cancer stem cells that have an indispensable ability of self-renewability and differentiation.³³ Constitutive activation of this ligand-induced pathway is found to be associated with different human cancers, such as breast, prostate, colorectal, hepatocellular, and melanoma.³⁴ The aberrant expression of Wnt cascade has been implicated in lung tumors due to dysregulation of various Wnt inhibitors. *DKK*, an antagonist of Wnt signaling, is expressed in NSCLC cell lines,²⁰ whereas *sFRP* was found to be downregulated.¹⁶ Evidences supporting the clinical significance of Wnt pathway in predicting the prognosis of NSCLC patients are very few. These reports used immunostaining techniques for analyzing the level of β -catenin in NSCLC tissues, which depicts the activation of Wnt pathway. However, the inferences about the prognostic implications of this signaling pathway have been conflicting.³⁵ In the present study, genetic differences within the Wnt antagonist genes such as *DKK*, *sFRP*, and *Axin2* have been examined to find their role in predicting the survival and prognosis of North Indian lung cancer patients. This is probably the first attempt to explore the relevance of these polymorphic sites to act as an independent predictor of survival.

Findings from the current study suggested a lack of any significant association of the variants of *DKK4*, *DKK3*, *DKK2*, *sFRP3*, *sFRP4*, and *Axin2* genes with OS of lung cancer patients. However, certain trends were observed in *DKK4* rs2073664 where in the heterozygous *CT* genotype showed a better survival compared with wild *CC* type. In *Axin2* 1712 + 19, the subjects having *GG* genotype showed a trend of poor prognosis compared with those with *TT* genotype.

Lung cancer is a complex and intricate solid tumor, so different histological subtypes possess diverse pathological pathways for initiation and progression. An attempt was made to find the influence of these variants on the prognosis of patients having a particular histological subtype. When patients were stratified on the basis of their histological subtypes, several interesting clinical outcomes were obtained. *DKK4* rs2073664 emerged as an independent predictor of survival in ADCC patients. This variant has prominently emerged as a good prognostic tool for heterozygous (*CT*) subjects, and on the contrary, the death rate was threefold higher for patients with variant genotype (*TT*). *DKK4* is an established inhibitor of Wnt signaling, which is ligand induced; therefore, it can be considered for having antitumor potential. Antitumor properties of

A

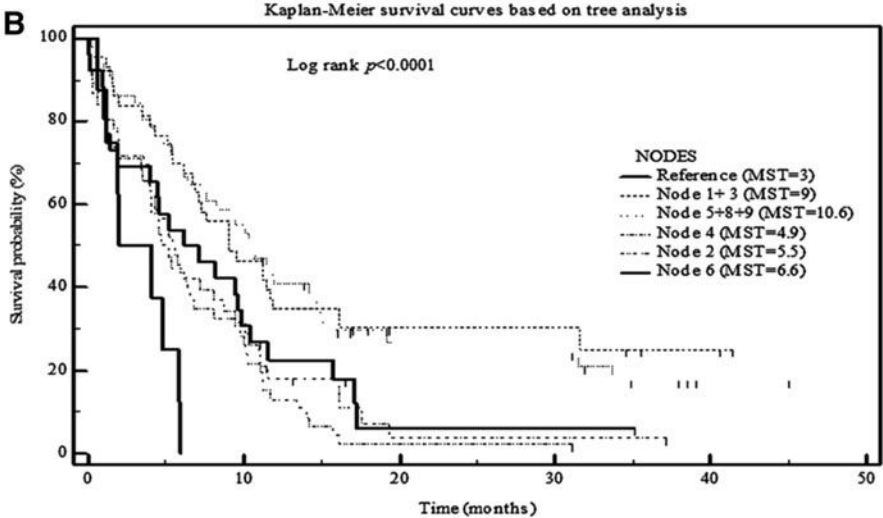
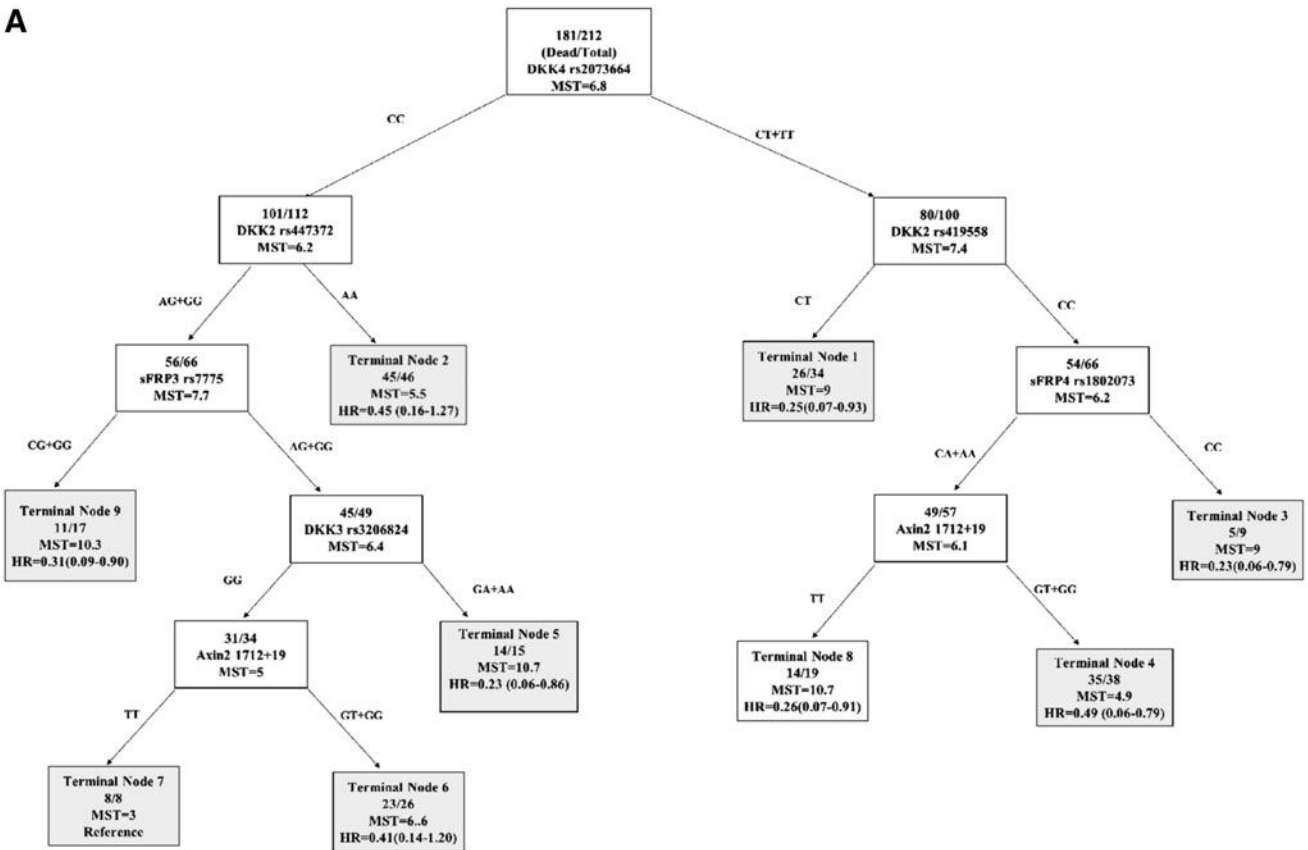


FIG. 2. (A) Survival tree analysis of sFRP, DKK, and Axin2 variants. MST in months; HR with 95% confidence interval (univariate analysis). (B) Overall survival according to different nodes identified by survival tree analysis. HR, hazard ratio; MST, median survival time.

DKK4 pertain to its ability of activating GSK-3 β and the subsequent degradation of cytoplasmic β -catenin.³⁶ The polymorphic site rs2073664 lies in exon4 of *DKK4* gene and results in synonymous valine substitution at 169th position. However, the functional relevance of this sequence variation is yet unclear.³⁷ Very few reports are available in regard of this SNP where its effect on prognosis of cancer patients was investigated. A study performed in renal cancer patients negated

the role of this polymorphic site in predicting their survival after radical nephrectomy.²⁸

Among the three genetic variations explored within the *DKK3* gene, only rs3206824 showed an increased death risk in SCLC patients after adjusting for various covariates. This outcome was not in line with the study performed in renal cell carcinoma patients undergoing radical nephrectomy, where all the three SNPs did not contribute in predicting the prognosis of

the patients.²⁸ DKK3 protein has been elucidated as key player in tumorigenesis process and is indicated to possess tumor-suppressive abilities. This property of inhibiting the cancer cell growth has been confirmed in various cancer cell line models, such as lung, gastric, glioma, colon, and hepatic. Lung cancer was the first primary tumor, where the dysregulated expression pattern of DKK3 was confirmed.³⁸

In *DKK2* gene, which has a strong affinity for Wnt receptor complex, only rs419558 polymorphic site was found to have a good prognostic implication in ADCC patients. This site lies in the 3'-UTR region, which is responsible for stability of mRNA.³⁴ In contrast to this finding, Hirata et al. suggested that other variant of *DKK2* gene, rs17037102, can act as a predictor of survival of renal cancer patients, whereas the patients harboring the combined variant (GA + AA) had a decreased risk of death compared with those having GG genotype.²⁸ One of the interesting finding of the present study obtained after investigating the cumulative role of the unfavorable genotypes of the three *DKK2* variants is a decreased rate of death (HR = 0.37, $p = 0.03$) in North Indian lung cancer patients. No reports are available till date about the cumulative effect of these SNPs on the survival and prognosis of cancer patients. Hence, it is plausible that these single point mutations if investigated in other ethnicities can act as biomarker for lung cancer patients.

Another outcome, which was evident from the present study, was the association of *sFRP3* rs7775 variant with the survival time of SCLC patients. Subjects having CG genotype had a significantly higher risk of death (adjusted HR = 2.03, $p = 0.05$) compared with those having CC genotype. The polymorphic site rs7775 at codon 324 is located at the C-terminal end of *sFRP3* gene, which is known extracellular glycoprotein and binds to Wnt ligands to negatively regulate the Wnt cascade. sFRPs interact with Wnt ligands through N-terminal frizzled cysteine-rich domain.^{39,40} However, a study has shown that the direct binding between sFRP1 and its ligand Wnt was strongly influenced by heparin, and it should be noted that the C-terminus shows the heparin-binding activity. Certain studies have suggested the fact that to antagonize the Wnt signal, sFRP requires both N- and C-terminal, although it is able to bind the Wnt ligand when truncated.³⁹⁻⁴¹

Axin2 is the major scaffold protein of the degradation complex, which facilitates the GSK-3 β -mediated phosphorylation of β -catenin to avoid its internalization into the nucleus and consequently inhibiting the activation of transcription factors TCF/LEF responsible for regulating the expression of downstream target genes of Wnt cascade involved in the process of oncogenesis.^{42,43} Of all the seven variants under study, only three variants, namely 148 C/T, 1712 + 19 G/T, and 1386 C/T, were impeding to have a certain implication as a prognostic marker. The two exonic variants (148 and 1386) showed contrasting results, where the ADCC subjects having TT genotype for 148 showed a better prognosis and those with TT genotype for *Axin2* 1386 showed a poor prognosis in SCLC subtype. The intronic *Axin2* 1712 + 19 variant on the other hand indicated a highly increased death risk in ADCC patients with GG genotype.

Prognostic role of Wnt pathway has been investigated in various cancers, including breast cancer,²² pancreatic ADCC,²⁴ ovarian cancer,²³ NSCLC,³⁵ and hematological malignancies.²¹ However, the genetic variations of the Wnt pathway have been evaluated in NSCLC. One such study performed in

Korean population examined the different sequence variations of Wnt 5A, β -catenin, Wnt16, and Fzd4 and reported lack of any significant association with survival of lung cancer patients.²⁷ Another study performed in early NSCLC patients has explored around 40 SNPs of 20 genes of Wnt pathway, where some of the genetic variations in Wnt16 and β -catenin showed a significant association with survival of these patients before and after surgical treatment.²⁶ An important highlight of the present study is the use of high-order interaction tools to find out the subgroups showing the best prognosis and higher OS due to the presence of the different genotypic combinations, which cannot be inferred by traditional statistical tools. This was probably the first attempt to analyze the interaction of Wnt antagonists in predicting the survival of lung cancer patients depicted by a survival tree, where DKK4 rs2073664 came out as most significant factor in modulating the survival of lung cancer patients.

In conclusion, this study revealed the association of the sequence differences within the Wnt antagonist genes with the survival of North Indian lung cancer patients. The present findings illuminate the scope of *DKK2* variants to act as prognostic marker in lung cancer. The role of *DKK4* variant rs2073664 is plausible in ADCC progression and initiation as it came across as an independent predictor of survival in this case. However, a major limitation of this study was a limited sample size in stratified categories, but these findings guide a new track for uncovering the host differences, which play a role in progression of lung cancer that has immense metastatic potential. Furthermore, studies are needed to validate these facts and pave a way for exploring the Wnt antagonists, which have a potential to act as a prognostic biomarker.

Acknowledgments

The authors would like to express their gratitude to all the subjects who participated in this current study. This work was supported by the grant from the Indian Council of Medical Research, New Delhi, India (Grant No. 5/13/126/2011/NCD-III).

Ethical Approval

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

Compliance with Ethical Standards

The current study is in compliance with ethical standards.

Informed Consent

Informed consent was obtained from all individual participants included in the study.

Disclosure Statement

No competing financial interests exist.

References

1. Nguyen DX, Chiang AC, Zhang XHF, et al. WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 2009;138:51.

2. Chen Z, Fillmore CM, Hammerman PS, et al. Non-small-cell lung cancers: A heterogeneous set of diseases. *Nat Rev Cancer* 2014;14:535.
3. Borczuk AC, Gorenstein L, Walter KL, et al. Non-small-cell lung cancer molecular signatures recapitulate lung developmental pathways. *Am J Pathol* 2003;163:1949.
4. Reya T, Morrison SJ, Clarke MF, et al. Stem cells, cancer, and cancer stem cells. *Nature* 2001;414:105.
5. Pongracz JE, Stockley RA. Wnt signalling in lung development and diseases. *Respir Res* 2006;7:1.
6. Pinto D, Gregorieff A, Begthel H, et al. Canonical Wnt signals are essential for homeostasis of the intestinal epithelium. *Genes Dev* 2003;17:1709.
7. Korinek V, Barker N, Moerer P, et al. Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet* 1998;19:379.
8. Akiyama T. Wnt/ β -catenin signaling. *Cytokine Growth Factor Rev* 2000;11:273.
9. Giles RH, van Es JH, Clevers H. Caught up in a Wnt storm: Wnt signaling in cancer. *Biochim Biophys Acta* 2003;1653:1.
10. Shi Y, He B, You L, et al. Roles of secreted frizzled-related proteins in cancer. *Acta Pharmacol Sin* 2007;28:1499.
11. Zorn AM. Wntsignalling: Antagonistic Dickkopfs. *Curr Biol* 2001;11:592.
12. Tetsu O, McCormick F. β -catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* 1999;398:422.
13. Tennis M, Van Scoyk M, Winn RA. Role of the wnt signaling pathway and lung cancer. *J Thorac Oncol* 2007;2:889.
14. Veeck J, Niederacher D, An H, et al. Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. *Oncogene* 2006;25:3479.
15. Shih YL, Shyu RY, Hsieh CB, et al. Promoter methylation of the secreted frizzled-related protein 1 gene SFRP1 is frequent in hepatocellular carcinoma. *Cancer* 2006;107:579.
16. Tang M, Torres-Lanzas J, Lopez-Rios F, et al. Wnt signaling promoter hypermethylation distinguishes lung primary adenocarcinomas from colorectal metastasis to the lung. *Int J Cancer* 2006;119:2603.
17. Takada T, Yagi Y, Maekita T, et al. Methylation-associated silencing of the Wnt antagonist SFRP1 gene in human ovarian cancers. *Cancer Sci* 2004;95:741.
18. Yue W, Sun Q, Dacic S, et al. Downregulation of Dkk3 activates β -catenin/TCF-4 signaling in lung cancer. *Carcinogenesis* 2008;29:84.
19. Sato H, Suzuki H, Toyota M, et al. Frequent epigenetic inactivation of DICKKOPF family genes in human gastrointestinal tumors. *Carcinogenesis* 2007;28:2459.
20. Nozaki I, Tsuji T, Iijima O, et al. Reduced expression of REIC/Dkk-3 gene in non-small cell lung cancer. *Int J Oncol* 2001;19:117.
21. Ashihara E, Takada T, Maekawa T. Targeting the canonical Wnt/ β -catenin pathway in hematological malignancies. *Cancer Sci* 2015;106:665.
22. Kazi MM, Trivedi TI, Kobawala TP, et al. The potential of Wnt signaling pathway in cancer: A focus on breast cancer. *Cancer Transl Med* 2016;2:55.
23. Bodnar L, Stanczak A, Cierniak S, et al. Wnt/ β -catenin pathway as a potential prognostic and predictive marker in patients with advanced ovarian cancer. *J Ovarian Res* 2014;7:1.
24. Nakamoto M, Matsuyama A, Shiba E, et al. Prognostic significance of WNT signaling in pancreatic ductal adenocarcinoma. *Virchows Archiv* 2014;465:401.
25. Deng F, Zhou K, Cui W, et al. Clinicopathological significance of wnt/ β -catenin signaling pathway in esophageal squamous cell carcinoma. *Int J Clin Exp Pathol* 2015;8:3045.
26. Coscio A, Chang DW, Roth JA, et al. Genetic variants of the Wnt signaling pathway as predictors of recurrence and survival in early-stage non-small cell lung cancer patients. *Carcinogenesis* 2014;35:1284.
27. Yoo SS, Hong MJ, Choi JE, et al. Genetic variants in the Wnt signaling pathway are not associated with survival outcome of non-small cell lung cancer in a Korean population. *J Korean Med Sci* 2016;31:463.
28. Hirata H, Hinoda Y, Nakajima K, et al. Wnt antagonist gene polymorphisms and renal cancer. *Cancer* 2009;115:4488.
29. Sodhi KK, Bahl C, Singh N, et al. Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians. *Future Oncol* 2015;11:2159.
30. Shanmugam KS, Brenner H, Hoffmeister M, et al. The functional genetic variant Arg324Gly of frizzled-related protein is associated with colorectal cancer risk. *Carcinogenesis* 2007;28:1914.
31. Gunes EG, Pinarbasi E, Pinarbasi H, et al. Strong association between lung cancer and the AXIN2 polymorphism. *Mol Med Rep* 2009;2:1029.
32. Kumari A, Bahl C, Singh N, et al. Association of p53 codon 72 polymorphism and survival of North Indian lung cancer patients treated with platinum-based chemotherapy. *Mol Bio Rep* 2016;43:1383.
33. Willert K, Jones KA. Wnt signaling: Is the party in the nucleus? *Gene Dev* 2006;20:1394.
34. Alanazi MS, Parine NR, Shaik JP, et al. Association of single nucleotide polymorphisms in Wnt signaling pathway genes with breast cancer in Saudi patients. *PLoS One* 2013;8:e59555.
35. Shapiro M, Akiri G, Chin C, et al. Wnt pathway activation predicts increased risk of tumor recurrence in patients with stage I non-small cell lung cancer. *Ann Surg* 2013;257:548.
36. Niehrs C. Function and biological roles of the Dickkopf family of Wnt modulators. *Oncogene* 2006;25:7469.
37. Yilmaz M, Donmez G, Kacan T, et al. Significant association between polymorphisms of Wnt antagonist genes and lung cancer. *J Investig Med* 2015;63:935.
38. Veeck J, Dahl E. Targeting the Wnt pathway in cancer: The emerging role of Dickkopf-3. *Biochim Biophys Acta Rev* 2012;1825:18.
39. Frank B, Shanmugam KS, Beckmann L, et al. Death receptor 4 variants and colorectal cancer risk. *Cancer Epidemiol Biomarkers Prev* 2006;15:2002.
40. Lin K, Wang S, Julius MA, et al. The cysteine-rich frizzled domain of Frzb-1 is required and sufficient for modulation of Wnt signaling. *Proc Natl Acad Sci USA* 1997;94:11196.
41. Chong JM, Uren A, Rubin JS, et al. Disulfide bond assignments of secreted frizzled related protein-1 provide insights about frizzled homology and netrin modules. *J BiolChem* 2002;277:5134.
42. Dong X, Seelan RS, Qian C, et al. Genomic structure, chromosome mapping and expression analysis of the human AXIN2 gene. *Cytogenet Cell Genet* 2001;28:26.
43. Phelan CM, Borg A, Cuny M, et al. Consortium study on 1280 breast carcinoma; allelic loss on chromosome 17 targets subregions associated with family history and clinical parameters. *Cancer Res* 2000;58:1004.

Association study between genetic variations in *Axin2* gene and lung cancer risk in North Indian population: A multiple interaction analysis

Charu Bahl¹, Siddharth Sharma¹, Navneet Singh²
and Digamber Behera²

Tumor Biology
April 2017: 1–18
© The Author(s) 2017
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1010428317695533
journals.sagepub.com/home/tub


Abstract

Wnt pathway has been implicated in the process of human carcinogenesis. Axis inhibition protein2 (*Axin2*), a major scaffold protein is an antagonist of Wnt pathway and is potent to act as a tumor suppressor gene in various human cancers. Therefore, the seven polymorphic sites of *Axin2* gene were analyzed, in relation to lung cancer susceptibility in North Indians. A total of 608 subjects were genotyped using PCR-RFLP technique for each polymorphic site including 303 cases and 305 controls. Further association analysis was carried out using logistic regression approach to obtain adjusted odds ratio and statistical significance. MDR and CART analysis were applied to evaluate high order interactions between the SNP's. Three out of seven studied polymorphic sites showed a strong protective effect in subjects having mutant genotype for *Axin2* 148 C > T and heterozygous genotype for 1365 G > A and 1712 + 19 G > T towards lung cancer risk. The other important finding was the significant association of *Axin2* 148 C > T in SQCC patients having variant (TT) genotype. *Axin2* 1712 + 19 G > T showed a decreased risk for all the histological subtypes in patients with heterozygous (GT) genotype. MDR analysis predicted a best interaction model (*Axin2* 148, *Axin2* 2062 and *Axin2* 1712 + 19) with maximum CVC (10/10) and minimum prediction error (0.38) along with significant permutation *p*-value. CART analysis gave a wide spectrum of interactive combinations which exhibited a major contribution in modulating lung cancer susceptibility. *Axin2* 148 and *Axin2* 1712 + 19 were found to play a major role in modulating lung cancer risk.

Keywords

Lung cancer, *Axin2*, Wnt pathway, single-nucleotide polymorphism

Date received: 11 August 2016; accepted: 23 December 2016

Introduction

Wnt signaling is one of the primary pathways which form the basic framework of cellular development and differentiation in organisms. Cellular polarity in embryos and tissue maintenance in adults are some of the major dimensions of this intermeshing pathway.¹ This cascade directs the expression of all the primary genes involved in the processes of cell polarity, cellular adhesion, apoptosis, and tissue differentiation.² Canonical Wnt pathway is a β -catenin-mediated signaling network which is switched on upon the binding of the Wnt ligands to the cell

surface-frizzled receptors and low-density lipoprotein receptor-related protein 5/6 (LRP5/6) complex.³ This is the primary interaction which disrupts the destruction

¹Department of Biotechnology, Thapar University, Patiala, India

²Department of Pulmonary Medicine, Post Graduate Institute of Education and Medical Research (PGIMER), Chandigarh, India

Corresponding author:

Siddharth Sharma, Department of Biotechnology, Thapar University, Patiala, Punjab 147002, India.

Email: siddharthsharma.phd@thapar.edu



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons

Attribution-NonCommercial 4.0 License (<http://www.creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

complex comprising disheveled (Dsh), adenomatous polyposis coli (APC), and Axin2 proteins which is formed for the ubiquitination of β -catenin.^{4,5} But, due to Wnt ligand–surface receptor binding, the phosphorylation of β -catenin by glycogen synthase kinase- β (GSK-3 β) in this destruction complex is inhibited, and then, consequently, β -catenin is translocated to the nucleus where it binds to the several transcription factors such as T-cell factor/lymphoid enhancer factor (TCF/LEF) which regulate the expression of target genes including *matrix metalloproteinases*, *cox-2*, *cyclin D1*, vascular endothelial growth factor (*VEGF*), and *c-myc*.³

Wnt pathway can be regulated by various proteins including Wnt antagonists such as sFRP and DKK and other proteins, for example, *Axin2*, which forms the part of the destruction complex that inhibits the translocation of β -catenin into the nucleus.³ *Axin2* is an architectural protein which forms the scaffold of this destruction complex. In a way, *Axin2* negatively controls the Wnt signaling pathway, and it has been established as one of the potent tumor suppressor gene. *Axin2* is located on chromosome 17q24 in humans and codes for 843 amino acid–long protein. *Axin2* gene comprises 10 exons.⁶ Dysregulation of Wnt pathway may lead to its constitutive activation which enhances the unregulated division of cells,^{7,8} and henceforth, this augments the process of carcinogenesis in various tissues such as skin, breast, liver, ovary, and lung.⁹ Overexpression of target genes is evident in various tumor tissues.¹⁰ Lung cancer is a multifactor-driven cellular abnormality which arises because of certain heterogeneous and unknown genetic and molecular anomalies,¹¹ and recently, some reports have established the role of this developmental pathway in the oncogenesis of lung tissue.¹² Several types of cancers have evidently reported the modifications in *Axin2* protein.¹³ The variations in the *Axin2* gene may cause certain modulations in the protein structure, and as it forms the scaffold of the destruction complex, its binding with APC may be affected and the process of addition of phosphate moiety to β -catenin may be altered, which is the primary step for its ubiquitination.¹⁴ Various findings have illustrated the loss of heterozygosity and also certain mutations in various cancers.¹⁵ *Axin2* gene has been found to be dysregulated in various tumors such as breast, colorectal, ovarian, and hepatocellular.¹⁶ Elevated levels of β -catenin were certainly attributed to the decreased expression of *Axin2* in lung cancer tissue, and in case of non-small-cell lung cancer (NSCLC), accumulation of β -catenin in nucleus is attributed to the epigenetic silencing of *Axin2* gene.¹⁷ There are no findings about the specific mutations in *Axin2* gene in relation to lung cancer.

Polymorphic sites in the gene sequence do greatly influence the orchestration of the protein structure and function. Several studies have been carried out exploring

the effect of certain polymorphic sequence differences on the disease susceptibility.¹⁸ Several genetic factors intricately influence the individual's predisposition to acquire lung cancer. There have been a few reports where the sequence variation within the *Axin2* gene was analyzed to find out its effect on lung cancer risk.¹⁹ One of the commonly studied single-nucleotide polymorphism (SNP) includes rs22040308, which has been assessed in different populations such as Turkish,²⁰ Japanese,²¹ Polish,²² and Chinese.²³ These studies have been carried out in different carcinomas including ovarian,²² lung,²⁰ prostate,²⁴ and astrocytoma.²⁵ A few studies have also evaluated several other intronic and exonic polymorphic sites including *Axin2* 148 C/T, *Axin2* 1365 G/A, *Axin2* 432 T/C, *Axin2* 956+16 A/G, *Axin2* 1712+19 G/T, *Axin2* 1386 C/T, and *Axin2* 2062 C/T in Turkish lung cancer patients²⁰ and astrocytoma patients.²⁵ However, the sample size was very low and further studies are required to establish the role of these genetic variants within *Axin2* gene in modulating toward lung cancer susceptibility.

Taking into account the above-mentioned facts, this study was designed to find the association of seven *Axin2* SNPs, including 148 C>T, 1365 G>A, 432 T>C, 956+16 A>G, 1712+19 G>T, 1386 C>T, and 2062 C>T, with lung cancer susceptibility in North Indian population. An attempt was made to identify the high-risk subgroups by stratifying the subjects on the basis of histology and smoking parameters. The interaction of these SNPs was analyzed using novel computational algorithms, so as to accomplish a comprehensive appraisal of lung cancer risk. Multifactor dimensionality reduction (MDR) approach and Classification and Regression Tree (CART) analysis yielded some interesting patterns and combinations of these polymorphic sites that contribute evidently in modulating lung cancer risk.

Material and methods

Study population

A total of 608 subjects have been enrolled in this study, comprising 303 cases and 305 controls. The patients have been recruited from Lung Cancer Clinic of Department of Pulmonary Medicine, Postgraduate Institute of Medical Education and Research (PGIMER), Chandigarh. The study was approved by the Institutional Ethics Committee of PGIMER. Subjects were explained about this study and an informed written consent was also obtained from each participant. A volume of 5 mL of peripheral blood was taken from each volunteer, and a detailed questionnaire was filled by the volunteers with the help of a trained personnel in order to obtain all the epidemiological details such as age, sex, smoking history, and region they belong from. The clinical details of the patients were obtained

from the medical records which included histology of the tumor, stage of the disease, and other specific details. The selection criteria for cases were only that the patient should not have past history of any other carcinoma; otherwise, there were no age, gender, smoking, histology, TNM, and familial history restrictions. The controls were the subjects who visited the hospital for normal health checkups. Controls were selected on the basis of the fact that they did not have any record of previous tumors; however, their familial history did not pose any restriction. The sampling was carried in such a manner so that there is age (± 10 years), gender, and smoking matching among cases and controls. This attempt was done in order to avoid sampling bias. In order to obtain a clear idea about the total smoking exposure, pack-years were obtained using the following equation: ((cigarettes or beedis per day/20) $\times$ number of years smoked).

DNA extraction

The isolation of genomic DNA was done using the phenol chloroform extraction method with certain modifications as detailed by Sobti et al.²⁶

Genotyping of *Axin2* variants using polymerase chain reaction–restriction fragment length polymorphism

The genotyping of the seven polymorphic variants of *Axin2* gene was carried out by polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP). The primers and restriction enzymes used in this protocol have been previously detailed by Gunes et al.²⁰ The PCR protocol used for the amplification of these genotypic variants is as follows. The 25 μ L reaction mixture comprised 1 $\times$ PCR buffer, 1.5 mM $MgCl_2$ with 0.5 μ M of both forward and reverse primers, 200 μ M of each deoxynucleotide triphosphates (dNTPs), 100 μ g/mL bovine serum albumin (BSA), 1 U Taq polymerase (DNAzyme II DNA Polymerase; Thermo Scientific, Waltham, MA, USA), and approximately 200 ng DNA. The annealing temperature was optimized for each variant and their respective restriction enzymes have been mentioned in Table 1. The amplified product was digested with the respective restriction enzymes (5U) and was incubated at desired temperature overnight. RFLP products were resolved on the agarose gel electrophoresis/polyacrylamide gel electrophoresis depending on the difference in the size of fragments to be separated. The gels were visualized using either ethidium bromide or silver staining. Ethidium bromide can be replaced by some safe staining agent as it is toxic, and if it is used, proper precautions have to be taken. The restriction patterns were studied and scoring was done accordingly as mentioned in Table 1. The genotyping of 15% samples was repeated twice in order to check the reproducibility of the

Table 1. List of primers and restriction enzymes used.

Polymorphism	Marker	Location	Primer sequence: forward and reverse	Product (bp)	Restriction enzyme	Allele (bp)
148 C > T	rs2240308	Exon1	F: CCA CGC CGA TTG CTG AGA GG R: TTC CGC CTG GTG TTG GAA GAG ACA T	242	MphI / O3I C	C (218 + 24) T (242)
432 T > C	rs2240308	Exon1	F: CCT GGA GAG GGA GAA ATG C R: CAT CAC CGA CTG GAT CTC G	253	Esp3I T	T (253) C (145 + 108)
956 + 16 A > G	rs35285779	Intron2	F: GGC TTT CTG GTG AGG GTC AGT R: CAA GCA AGC CCA CGG AAG G	228	MisI	A (205 + 23) G (228)
1365 G > A	rs9915936	Exon5	F: ACG TCT TCC CTT TCA GGA TG R: GGT ACT GCG AAT GGT GGT G	248	MspI	G (100 + 80 + 36 + 32) A (116 + 100 + 32)
1386 C > T	rs1133683	Exon5	F: TGC GTA GGG AGC CGA ATG TTG R: GTG GTC CGG GGA GCG GGA TC	294	TaqI	C (294) T (274 + 20)
1712 + 19 G > T	–	Intron5	F: CAC CAC CAC TAC ATC CAC CAC R: GCT CCC ACC TCA CAC CTG	249	CpoI	G (184 + 65) T (249)
2062 C > T	–	Exon7	F: TCC TTC TGT TTT CTC TCT GCT CAT TCC R: AGC GTG TTG GGT GGG GTC G	209	TaqI	C (209) T (191 + 18)

genotyping process, and the results were found to be 100% concordant.

Statistical analysis

Difference in the demographic variables such as age, gender, and smoking status was evaluated using the test of significance according to the type of variable. For the categorical variables such as gender and smoking status, chi-square statistics was used. However, the continuous variables such as age and pack-years were compared using *t*-test among cases and controls. The genotypic frequencies of the control group were evaluated using chi-square goodness of fit in order to find whether they were in accordance with Hardy–Weinberg equilibrium (HWE; $p^2 + 2pq + q^2 = 1$). The difference in the genotypic frequencies among cases and controls was evaluated using Pearson's χ^2 test. Unconditional multivariate logistic regression approach was used to find the association between the polymorphic variants and lung cancer predisposition. This analytical approach gives an adjusted odds ratio (AOR) and 95% confidence interval (CI) which is adjusted for various confounding factors such as age, gender, and smoking status. The same approach was used for evaluating the risk associated with these polymorphic sites in the various subgroups formed on the basis of histology and smoking. All *p* values were two-sided, and a *p* value of <0.05 was considered statistically significant.²⁷ All the statistical analyses were performed with MedCalc version 16.8.1 (MedCalc Software, Ostend, Belgium) and SPSS Version 20.0. Haplotype and linkage disequilibrium (LD) analysis was carried out using the SHEsis software²⁸ ("<http://analysis.bio-x.cn/myAnalysis.php>"). For analyzing the inter SNP interaction, MDR approach was used to overcome the limitations of the traditional logistic regression approach. It is a non-parametric method where no hypothesis is made, and hence, no genetic inheritance model is applied. It basically analyzes all the data by reducing the data in a multidimensional manner and pooling the genotypes into high-risk and low-risk groups. It can easily access multiple loci at the same time. It finds the best interaction model on the basis of cross-validation consistency (CVC) which is number of times a model is identified as the best model and testing accuracy. The permutation testing is the method which is used by this approach to find the significance of the hypothesis generated. MDR analysis was performed using the version 0.5.1 of the open-source MDR software package that is available online (<http://www.epistasis.org>).²⁹ Another high-order interaction between all the polymorphic variants was analyzed using CART. CART was applied using the CART software (6.0; Salford Systems, San Diego, CA, USA). The CART software uses the data mining approach to create a decision tree which depicts the interaction of all the variables under study through a binary recursive partitioning method to identify subgroups at higher risk, which are not easily revealed by the traditional logistic regression models. The

analysis is conducted in such a manner where the most significant predictor is used to split the sample into subgroups and continues until the differences are not significant. The tree obtained has various nodes and sub nodes depicting their role in modulating the disease susceptibility. The node with a lower case rate is used as a reference to find out the odds ratio (OR) and 95% CI for all other genotypes depicted in other nodes.³⁰ Multiple testing corrections were not used for any of the statistical analyses carried out in this study.

Results

Demographic characteristics

Table 2 summarizes the various demographic details of the study subjects who were recruited for this study. The study comprised 303 cases and 305 controls. The mean age of both the groups has been illustrated in Table 2 which came out to be 57.38 ± 10.74 years for cases and 53.23 ± 10.44 years for controls. The numbers of males and females were also compared between the two groups, and it was observed that there was no significant difference in both cohorts ($p=0.60$). The percentage of males in cases and controls was 86% and 88%, respectively. The next important factor under consideration was smoking; an adequate attempt was made to match the number of smokers and non-smokers in both groups. It was seen that number of smokers was slightly higher in the case group (78.8% vs 73.1%). However, the difference was not found to be significant ($p=0.11$). The mean pack-years were also calculated for both the study groups. The mean pack-years were higher in the cases as compared to controls (28.09 ± 34.09 years vs 19.4 ± 19.56 years). A significant difference ($p=0.0003$) was found between the pack-years of cases and controls pertaining to the fact that smoking came out to be a significant risk factor for lung cancer which is an already established fact. Furthermore, considering the clinical parameters of the patients, it was found that the major percentage of patients were diagnosed with squamous cell carcinoma (SQCC; 43.23%) followed by adenocarcinoma (ADCC; 32%), and the least percentage of patients were found to suffer from small-cell lung cancer (SCLC; 23.76%). The other clinical factors under study include the TNM staging where the primary tumor extension is classified from T_0 to T_4 , the presence and extent of lymph node involvement are classified from N_0 to N_3 , and distant metastasis is classified into M_0 and $M_{1a/1b}$.

Allelic and genotypic frequencies of the *Axin2* variants and their association with overall lung cancer risk on the basis of histological subtypes

The genotypic and allelic frequencies of the seven different variants of *Axin2* gene have been tabulated in Table 3. The seven polymorphic sites which have been evaluated in this study include *Axin2* 148 C>T, 1365 G>A, 432 T>C,

Table 2. Demographic characteristics among cases and controls.

Variable	Cases (n (%), N=303)	Controls (n (%) N=305)	p value ^a
Age (years)			
Mean $\pm$ SD	57.38 $\pm$ 10.74	53.23 $\pm$ 10.44	$p < 0.0001$
Range			
Gender			
Male	262 (86.46)	269 (88.19)	0.60
Female	41 (13.53)	36 (11.8)	
Smoking status			
Smokers	239 (78.87)	223 (73.11)	0.11
Non-smokers	64 (21.12)	82 (26.88)	
Pack-years			
Mean $\pm$ SD	28.09 $\pm$ 34.09	19.4 $\pm$ 19.56	0.0003
Histological types			
SQCC	131 (43.23)		
ADCC	97 (32.01)		
SCLC	72 (23.76)		
Others	2 (0.66)		
Unknown	1 (0.33)		
TNM staging			
I	3 (1)		
II	12 (4)		
III	140 (46.7)		
IV	123 (41)		
Unclassified	25 (7.3)		
Tumor size			
Tx	10 (3.3)		
T1	14 (4.7)		
T2	39 (13)		
T3	79 (26.3)		
T4	128 (42.7)		
Unknown	33 (10)		
Lymph node			
N0	48 (16)		
N1	36 (12)		
N2	112 (37.3)		
N3	72 (24)		
Unknown	33 (10)		
Metastasis			
M0	154 (51.3)		
M1	116 (38.7)		
Unknown	33 (10)		

SD: standard deviation; ADCC: adenocarcinoma; SQCC: squamous cell carcinoma; SCLC: small-cell lung carcinoma; n: total number of case patients or control subjects.

^ap values were derived from Pearson chi-square test except age; Student's t-test was used for age. All p values are two-sided. $p < 0.05$ was considered statistically significant.

956 + 16 $A > G$, 1712 + 19 $G > T$, 1386 $C > T$, and 2062 $C > T$. The genotypic and allelic frequencies have been illustrated for each variant along with the minor allele frequencies (MAFs) in both the study groups. Table 3 also shows the HWE values for control group of each polymorphic site. It can be observed that the following variants were in accordance to HWE, namely, *Axin2* 148 ($\chi^2 = 0.94$, $df = 1$, and $p = 0.33$), 1365 ($\chi^2 = 1.53$, $df = 1$, and $p = 0.21$),

956 + 16 ($\chi^2 = 0.31$, $df = 1$, and $p = 0.57$), *Axin2* 432 ($\chi^2 = 0.22$, $df = 1$, and $p = 0.63$), and 2062 ($\chi^2 = 1.84$, $df = 1$, and $p = 0.17$). Only two variants, namely, *Axin2* 1386 ($\chi^2 = 8.79$, $df = 1$, and $p = 0.003$) and 1712 + 19 ($\chi^2 = 175$, $df = 1$, and $p = 0.000$), showed a certain deviation from HWE.

In case of *Axin2* 148 $C > T$, the frequency of the heterozygous (CT) genotype was consistent among the cases

Table 3. Genotypic distribution of the Axin2 genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histology.

Overall lung cancer									
SQCC				ADCC			SCLC		
Controls (305), N (%)	Cases (303), N (%)	AOR (95% CI) ^a	p ^b	Cases (131), N (%)	AOR (95% CI) ^a	p ^b	Cases (97), N (%)	AOR (95% CI) ^a	p ^b
Axin2 148 C/T									
CC	81 (26.5)	99 (32.6)	1.00 (reference)	47 (35.9)	1.00 (reference)		30 (30.9)	1.00 (reference)	
CT	144 (47.3)	150 (49.6)	0.87 (0.59–1.28)	61 (46.7)	0.79 (0.48–1.29)	0.35	49 (50.5)	0.89 (0.51–1.55)	0.68
TT	80 (26.2)	54 (17.8)	0.55 (0.34–0.87)	23 (17.5)	0.51 (0.28–0.94)	0.03	18 (18.5)	0.62 (0.31–1.22)	0.17
CT + TT	224 (73.5)	204 (67.4)	0.75 (0.52–1.08)	84 (64.2)	0.69 (0.43–1.06)	0.11	67 (69)	0.79 (0.47–1.34)	0.39
C (allele)	153	174							
T (allele)	152	129							
MAF	0.49	0.42							
HWE	$\chi^2=0.94$, df= 1, and p=0.33								
Axin2 1365 G/A									
GG	249 (81.6)	268 (88.4)	1.00 (reference)	116 (88.5)	1.00 (reference)		81 (83.5)	1.00 (reference)	
GA	51 (16.7)	29 (9.6)	0.45 (0.27–0.75)	13 (9.9)	0.53 (0.27–1.04)	0.06	12 (12.6)	0.57 (0.27–1.19)	0.13
AA	5 (1.6)	6 (2)	1.08 (0.37–3.73)	2 (1.5)	0.77 (0.14–4.28)	0.77	4 (4.1)	2.74 (0.66–11.57)	0.16
GA + AA	56 (18.3)	35 (11.6)	0.50 (0.31–0.81)	15 (11.4)	0.55 (0.29–1.04)	0.06	16 (16.5)	0.73 (0.38–1.40)	0.34
G (Allele)	274.5	282.5							
A (Allele)	30.5	20.5							
MAF	0.1	0.06							
HWE	$\chi^2=1.53$, df= 1, and p=0.21								
Axin2 956+16 A/G									
AA	248 (81.3)	255 (84.2)	1.00 (reference)	109 (83.2)	1.00 (reference)		84 (86.6)	1.00 (reference)	
AG	55 (18.0)	46 (15.2)	0.83 (0.53–1.29)	21 (16)	0.92 (0.52–1.64)	0.79	13 (13.4)	0.69 (0.35–1.36)	0.29
GG	2 (0.6)	2 (0.6)	–	1 (0.8)	–	–	–	–	–
AG + GG	57 (18.6)	48 (15.8)	0.83 (0.53–1.28)	22 (16.8)	0.92 (0.52–1.61)	0.78			
A (allele)	275.5	278							
G (allele)	29.5	25							
MAF	0.09	0.08							
HWE	$\chi^2=0.31$, df= 1, and p=0.57								
Axin2 432 T/C									
TT	289 (94.8)	295 (97.3)	1.00 (reference)	127 (96.9)	1.00 (reference)		94 (96.9)	1.00 (reference)	
TC	16 (5.2)	8 (2.7)	0.56 (0.23–1.35)	4 (3.1)	0.69 (0.22–2.19)	0.53	3 (3.1)	0.62 (0.17–2.27)	0.47
T (allele)	297	299							
C (allele)	8	4							
MAF	0.02	0.01							
HWE	$\chi^2=0.22$, df= 1, and p=0.63								

Table 3. (Continued)

Overall lung cancer																
SQCC				ADCC				SCLC								
Controls (305), N (%)		Cases (303), N (%)	AOR (95% CI) ^a	p ^b	Cases (131), N (%)		AOR (95% CI) ^a	p ^b	Cases (97), N (%)		AOR (95% CI) ^a	p ^b	Cases (72), N (%)		AOR (95% CI) ^a	p ^b
Axin2 1712+19 G/T																
TT	42 (13.8)	87 (28.7)	1.00 (reference)		34 (25.9)	1.00 (reference)			29 (29.9)	1.00 (reference)			22 (30.5)	1.00 (reference)		
GT	263 (86.2)	205 (67.7)	0.34 (0.22–0.52)	<0.0001	95 (72.5)	0.29 (0.16–0.52)	<0.0001	0.004	62 (63.9)	0.41 (0.23–0.75)	0.004	47 (65.3)	0.16 (0.08–0.35)	<0.0001		
GG	–	11 (3.6)	–		2 (1.6)	–		–	6 (6.2)	–		3 (4.2)				
GT+TT	263 (86.2)	216 (71.3)	0.36 (0.23–0.55)	<0.0001	97 (74.1)	0.30 (0.17–0.53)	<0.0001	0.009	68 (70.1)	0.45 (0.25–0.83)	0.009	50 (69.5)	0.18 (0.08–0.37)	<0.0001		
T (allele)	173.3	189.5														
G (Allele)	131.5	113.5														
MAF	0.43	0.37														
HWE	$\chi^2 = 175$, df = 1, and p = 0.000															
Axin2 2062 C/T																
CC	261 (85.6)	257 (84.8)	1.00 (Reference)		106 (81)	1.00 (Reference)			85 (87.6)	1.00 (Reference)		63 (87.5)	1.00 (Reference)			
CT	44 (14.4)	46 (15.2)	0.96 (0.60–1.52)	0.86	25 (19)	1.16 (0.66–2.03)	0.59	0.66	12 (12.4)	0.85 (0.41–1.74)	0.66	9 (12.5)	0.79 (0.36–1.77)	0.58		
C (allele)	283	280														
T (allele)	22	23														
MAF	0.07	0.07														
HWE	$\chi^2 = 1.84$, df = 1, p = 0.17															
Axin2 1386 C/T																
CC	103 (33.8)	94 (31)	1.00 (reference)		41 (31.3)	1.00 (reference)			34 (35.1)	1.00 (reference)		17 (23.6)	1.00 (reference)			
CT	169 (55.4)	172 (56.8)	1.12 (0.77–1.61)	0.53	76 (58)	1.20 (0.75–1.94)	0.43	0.84	53 (54.6)	0.94 (0.56–1.60)	0.84	42 (58.3)	1.77 (0.92–3.40)	0.08		
TT	33 (10.8)	37 (12.2)	1.24 (0.71–2.18)	0.43	14 (10.7)	1.11 (0.52–2.38)	0.78	0.95	10 (10.3)	0.97 (0.42–2.26)	0.95	13 (18.1)	2.82 (1.18–6.73)	0.01		
CT+TT	202 (66.2)	209 (69)	1.14 (0.80–1.62)	0.45	90 (68.7)	1.19 (0.75–1.88)	0.44	0.82	63 (64.9)	0.95 (0.56–1.57)	0.82	55 (76.4)	1.94 (1.03–3.65)	0.03		
C (allele)	187.5	180														
T (allele)	117.5	123														
MAF	0.38	0.40														
HWE	$\chi^2 = 8.79$, df = 1, and p = 0.003															

ADCC: adenocarcinoma; SQCC: squamous cell carcinoma; SCLC: small-cell lung carcinoma; AOR: adjusted odds ratio; CI: confidence interval; MAF: minor allele frequency; HWE: Hardy–Weinberg equilibrium.

^aAOR, 95% CI, and their corresponding p values were calculated by unconditional logistic analysis after adjusting for age, gender, and smoking status.

^bTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.
 p values in boldface depict statistically significant values.

(49.5%) and controls (47.2%). However, there lies a significant difference in between the genotypic distribution of the cases and controls ($\chi^2=6.96$, $df=2$, and $p=0.03$). MAF was found to be 0.42 in cases and 0.49 in controls. Using *CC* genotype as reference, the multivariate logistic regression showed that subjects carrying the variant (*TT*) genotype were exhibiting a protective effect toward lung cancer risk (OR=0.55, 95% CI: 0.34–0.87, and $p=0.01$). Out of the three histological subgroups, a decreased frequency of patients carrying variant (*TT*) genotype was observed only in SQCC subgroup (OR=0.51, 95% CI: 0.28–0.94, and $p=0.03$). No such association was observed in ADCC and SCLC.

In case of *Axin2* 1365 *G>A* polymorphic site, when a comparison between the genotypic frequencies of cases and controls was done, a significant difference was observed ($\chi^2=6.83$, $df=2$, and $p=0.03$). The MAF was found to be 0.1 and 0.06 in controls and cases, respectively. Taking *GG* as a reference genotype, logistic regression analysis revealed a protective effect in patients with heterozygous (*GA*) genotype (OR=0.45, 95% CI: 0.27–0.75, and $p=0.002$). However, patients with variant (*AA*) genotype did not show any such association with risk of lung cancer. Furthermore, stratifying the subjects on the basis of their histological subtypes, it was evident that lung cancer patients carrying the heterozygous and mutant genotypes did not exhibit any significant association in any of the histological subgroups; however, significant protective effect for heterozygous genotype was found in SCLC patients (OR=0.24, 95% CI: 0.08–0.72, and $p=0.01$).

The next intronic variant of *Axin2* gene under study was 956+16 *A>G*. The genotypic frequencies of cases and controls were compared, and it was found that there was no significant difference between the two study groups ($\chi^2=0.89$, $df=2$, and $p=0.63$). The proportion of subjects carrying the variant (*GG*) genotype was similar in both cases (0.6%) and controls (0.6%). MAF was comparable in cases (0.08) and controls (0.09). Association analysis was carried out using *AA* (wild) as a reference genotype, and it was visualized that neither of the variant genotypes, that is, *AG* and *GG*, showed any significant role in modulating the risk toward lung cancer as shown in Table 3. This same trend was also evident in subjects within every histological subtype. None of the histological subgroups showed any significant inclination toward affecting the risk associated with lung cancer in individuals carrying the combined variant (*AG+GG*) genotype for this polymorphic site.

For *Axin2* 432 *T>C*, the proportion of subjects with wild (*TT*) genotype was considerably over-represented in both cases (97.3%) and controls (94.7%). However, chi-square analysis revealed no significant difference between the genotypic frequencies among cases and controls ($\chi^2=2.07$, $df=1$, and $p=0.14$). The MAF was found to be 0.02 in controls and 0.01 in cases. We also report in this study a lack of subjects having the variant (*CC*) genotype both in cases and

controls. Taking *TT* genotype as reference, the individuals carrying the heterozygous (*TC*) genotype did not show any significant variation in their frequencies among cases and controls. Furthermore, when patients were classified on the basis of their histological subtype, none of the subgroups revealed any association with lung cancer risk. Therefore, this polymorphic site did not confer any role toward modulating lung cancer susceptibility.

The next variant under study was *Axin2* 1712+19 *G>T*. The proportion of subjects carrying the *GG* genotype was negligible in the control group and 3.6% in cases. The MAF was found to be 0.43 in controls and 0.37 in cases. The genotypic frequencies were significantly different between cases and controls ($\chi^2=33.8$, $df=2$, and $p<0.001$). When OR was calculated using *TT* genotype as a reference, a strong protective association was observed in patients having the heterozygous (*GT*) genotype which was found to be highly significant (OR=0.34, 95% CI: 0.22–0.52, and $p<0.0001$). This trend continued among even in the different histological subgroups where a decreased risk of lung cancer was observed in patients carrying the *GT* genotype within the SQCC (OR=0.29, 95% CI: 0.16–0.52, and $p<0.0001$), ADCC (OR=0.41, 95% CI: 0.23–0.75, and $p=0.004$), and SCLC (OR=0.16, 95% CI: 0.08–0.35, and $p<0.0001$).

Another variant lying in exon 7 of the *Axin2* gene, namely, 2062 *C>T* polymorphism, was also analyzed in this study. The frequency of wild (*CC*; 84.8% vs 85.5%) and heterozygous (*CT*; 15.1% vs 14.4%) was same in between cases and controls, respectively. No significant difference was observed in genotypic distribution of cases and controls ($\chi^2=0.02$, $df=1$, and $p=0.88$). MAF of 0.07 was observed in both cases and controls. The frequency of the variant (*TT*) genotype was null in both the groups. When OR was calculated using *CC* (wild) as a reference genotype, subjects carrying the heterozygous (*CT*) genotype did not exhibit any significant risk of developing lung cancer (OR=0.96, 95%CI: 0.60–1.52, $p=0.86$). When further stratification was carried out on the basis of histological subtype, this polymorphic variant did not pose any significant modulation for any of the histological subtype.

The next variant analyzed is *Axin2* 1386 *C>T*, the genotypic frequencies among cases and controls did not differ significantly ($\chi^2=0.66$, $df=2$, and $p=0.71$). The presence of subjects possessing variant (*TT*) genotype was almost similar in both cases and controls, respectively (12.2% vs 10.8%). MAF was found to be 0.40 in cases and 0.38 in controls. When logistic regression was applied by using *CC* as the reference genotype, no significant association was observed in subjects having the variant (*TT*) or heterozygous (*CT*) genotype for this polymorphic site. In the different subgroups, based on the histological subtype, only SCLC patients carrying the variant (*TT*) genotype showed a significant 2.8 fold risk of developing lung

cancer (OR=2.82, 95% CI: 1.18–6.73, and $p=0.01$). None of the other subgroups showed any such association.

Effects of smoking status and its association between *Axin2* gene polymorphisms and lung cancer risk

In order to evaluate the role of smoking in modulating the association between the *Axin2* gene polymorphisms and lung cancer risk, the subjects were classified on the basis of their smoking status into smokers and non-smokers. The association analysis was carried out for all the seven polymorphic sites of *Axin2* gene, and the genotypic frequencies along with the AOR with 95% CI within both the subgroups have been tabulated in Table 4. In case of *Axin2* 148 C>T, smokers harboring both the susceptible alleles were found to show a protective effect toward lung cancer as compared to the wild genotype patients (OR=0.50, 95% CI: 0.29–0.85, and $p=0.01$). However, no such association was observed in case of non-smokers. A similar finding was observed in case of smokers having the heterozygous GA genotype for *Axin2* 1365 G>A (OR=0.51, 95% CI: 0.28–0.91, and $p=0.02$). Also, the non-smokers having the same genotype as above exhibited a decreased risk toward lung cancer (OR=0.32, 95% CI: 0.10–1.10, and $p=0.05$). Patients in both smoker and non-smoker groups, possessing the combined variant genotype (GA+AA) reported a decreased risk (OR=0.57, 95% CI: 0.33–0.98; OR=0.33, 95% CI: 0.11–0.99) toward lung cancer with a significant p value of 0.04. In case of *Axin2* 1712+19 G>A, it was observed that the frequency of GT genotype was under-represented in the cases as compared to the control group. Furthermore, subjects who were smokers and carry single copy of variant allele (GT) were at a decreased risk of lung cancer as compared to controls. On the contrary, in case of the following genetic variants, namely, *Axin2* 432 T>C, 956+16 A>G, *Axin2* 2062 C>T, and 1386 C>T, no significant association was observed in any of the groups classified on the basis of smoking.

Combinatorial analysis of the risk associated with different polymorphic sites of *Axin2* gene

An attempt was made to analyze the SNP to SNP interaction to evaluate the combined effect of these polymorphic sites on lung cancer risk. The genotypic frequencies of these different combinations and the AOR and 95% CIs were summarized in Table 5. Various double combinations were analyzed for finding whether there exists any significant difference between the genotypic frequencies among cases and controls. We observed that subjects who carried the single copy of variant allele for the genotypic combination *Axin2* 148 C>T and 1365 G>A showed a protective effect (OR=0.42, 95% CI: 0.20–0.86, and $p=0.017$). Likewise, the patients having the heterozygous (CT+CT)

genotype for *Axin2* 148 C>T and *Axin2* 2062 C>T also exhibited a protective effect (OR=0.38, 95% CI: 0.17–0.83, and $p=0.01$). A strong protective effect was observed in patients carrying the heterozygous (CT+GT) genotype for *Axin2* 148 and *Axin2* 1712+19 G>T (OR=0.21, 95% CI: 0.09–0.48, and $p=0.003$). A similar effect was seen for the genotypic combination *Axin2* 1365 G>A and *Axin2* 1712+19 G>T in patients having heterozygous (GA+GT) genotype. We also found out that the subjects carrying a single copy of variant allele for *Axin2* 1712+19 G>T along with another single copy of variant allele of either 432 T>C (OR=0.15, 95% CI: 0.05–0.45, and $p=0.007$), 956+16 T>C (OR=0.23, 95% CI: 0.12–0.46, and $p<0.001$), and (OR=0.19, 95% CI: 0.09–0.38, and $p<0.001$) showed a decreased risk of lung cancer. However, this trend was not observed only for one genotypic combination, namely, *Axin2* 1386 C>T and 1712+19 G>T, where no significant association was assessed between the genotypic combination and lung cancer susceptibility.

Haplotype and LD analysis and their association with lung cancer susceptibility

Table 6 summarizes the frequencies for various haplotypes which were obtained on analyzing the seven polymorphic sites of the *Axin2* gene. Haplotype and LD analysis was done using sHEis software. Haplotype frequencies were classified into different haplotype blocks. Only those haplotype blocks were evaluated where the case and control frequency was more than 0.03, and rest of the haplotypes were not considered for finding their correlation with lung cancer predisposition. Out of the 70 possible haplotype combinations, only seven Hap blocks showed frequency more than 0.03 and hence could be evaluated for association. Global chi-square was also calculated for comparing the frequencies among cases and controls ($\chi^2=252.77$, df=7, and $p=1.33e^{-015}$). Hap blocks 1 and 5 showed an extremely low frequency of patients with the respective haplotypes among cases and hence conferring a strong protective effect toward lung cancer with highly significant p values. These Hap blocks have either variant allele for *Axin2* 148 C>T or 1712+19 G>T, and allele from other polymorphic sites was wild type. On the contrary, Hap3, Hap 6, and Hap7 showed an increased risk toward lung cancer as shown in Table 6. Hap block 2 consists of variant allele of *Axin2* 1386 C>T, and all other alleles were wild type. Similarly, Hap3 comprised variant alleles from both *Axin2* 1386 C>T and 1712+19 G>T and showed the maximum fourfold increased risk toward lung cancer (OR=4.05, 95% CI: 2.95–5.57, and $p=3.68e^{-019}$).

Table 7 depicts the LD (D') and r^2 values for all the seven polymorphic sites. A strong LD was observed in case of *Axin2* 432 and 1365 ($D'=0.89$ and $r^2=0.001$).

Table 4. Genotypic distribution of the Axin2 variants based on smoking status and its association with risk of lung cancer.

	Smokers				Non-smokers			
	Controls (n (%), N = 223)	Cases (n (%), N = 239)	AOR (95% CI) ^a	p ^b	Controls (n (%), N = 82)	Cases (n (%), N = 64)	AOR (95% CI) ^a	p ^b
Axin2 148 C/T								
CC	61 (27.3)	82 (34.3)	1.00 (reference)		20 (24.4)	17 (26.6)	1.00 (reference)	
CT	101 (45.2)	117 (49)	0.85 (0.55–1.32)	0.47	43 (52.4)	33 (51.6)	0.95 (0.41–2.17)	0.91
TT	61 (27.3)	40 (16.7)	0.50 (0.29–0.85)	0.01	19 (23.2)	14 (21.9)	0.93 (0.37–2.57)	0.89
CT + TT	162 (72.6)	157 (65.6)	0.71 (0.47–1.08)	0.11	62 (75.6)	57 (89.1)	0.94 (0.43–2.08)	0.89
Axin2 1365 G/A								
GG	182 (81.6)	211 (88.3)	1.00 (reference)		67 (81.7)	57 (89.1)	1.00 (reference)	
GA	37 (16.6)	23 (9.6)	0.51 (0.28–0.91)	0.02	14 (17.1)	6 (9.4)	0.32 (0.10–1.01)	0.05
AA	4 (1.8)	5 (2.1)	1.21 (0.31–4.75)	0.77	1 (1.2)	1 (1.6)	0.64 (0.03–12.22)	0.77
GA + AA	41 (18.4)	28 (11.7)	0.57 (0.33–0.98)	0.04	15 (18.3)	7 (11)	0.33 (0.11–0.99)	0.04
Axin2 956+16 A/G								
AA	182 (81.6)	201 (84.1)	1.00 (reference)		66 (80.5)	54 (84.4)	1.00 (reference)	
AG	39 (17.5)	36 (15.1)	0.82 (0.49–1.38)	0.46	16 (19.5)	10 (15.6)	0.73 (0.29–1.82)	0.50
GG	2 (0.9)	2 (0.8)	–	–	–	–	–	–
AG + GG	41 (18.4)	38 (15.9)	0.82 (0.50–1.36)	0.46	–	–	–	–
Axin2 432 T/C								
TT	213 (95.5)	233 (97.5)	1.00 (reference)		76 (92.7)	62 (96.9)	1.00 (reference)	
TC	10 (4.5)	6 (2.5)	0.63 (0.22–1.79)	0.38	6 (7.3)	2 (3.1)	0.41 (0.07–2.18)	0.30
Axin2 1712+19 G/T								
TT	16 (7.2)	65 (27.2)	1.00 (reference)		26 (31.7)	22 (34.4)	1.00 (reference)	
GT	207 (92.8)	167 (69.8)	0.18 (0.10–0.33)	<0.0001	56 (68.3)	38 (59.4)	0.89 (0.42–1.98)	0.77
GG	–	7 (2.9)	–	–	–	4 (6.2)	–	–
GT + TT	207 (92.8)	174 (72.7)	0.19 (0.10–0.34)	<0.0001	56 (68.3)	42 (65.6)	0.95 (0.45–2.05)	0.93
Axin2 2062 C/T								
CC	188 (84.3)	200 (83.7)	1.00 (reference)		73 (89.0)	57 (89.1)	1.00 (reference)	
CT	35 (15.7)	39 (16.3)	0.97 (0.58–1.63)	0.93	9 (11)	7 (10.9)	0.72 (0.23–2.20)	0.56
Axin2 1386 C/T								
CC	78 (35)	70 (29.3)	1.00 (reference)		25 (30.5)	36 (56.2)	1.00 (reference)	
CT	120 (54)	136 (56.9)	1.37 (0.89–2.08)	0.14	49 (59.7)	24 (37.5)	0.63 (0.29–1.36)	0.24
TT	25 (11)	33 (13.8)	1.65 (0.88–3.11)	0.11	8 (9.6)	4 (6.3)	0.52 (0.13–2.06)	0.35
CT + TT	145 (62)	169 (70.7)	1.42 (0.94–2.13)	0.09	57 (69.3)	28 (43.8)	0.62 (0.29–1.30)	0.21

AOR: adjusted odds ratio; CI: confidence interval.

^aAORs, 95% CIs, and their corresponding p values were calculated by unconditional logistic analysis after adjusting for age, gender, and smoking status.^bTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

Table 5. Genotypic distribution based on different genotypic combinations and their association with lung cancer risk.

Axin2 148 C/T and Axin2 1365 G/A	Cases (n (%), N = 105)	Controls (n (%), N = 98)	Adjusted OR (95% CI) ^a	p value ^b
(CC + GG) ⁰	87 (82.8)	70 (71.4)	1.00 (reference)	
(CT + GA) ¹	17 (16.2)	28 (28.6)	0.42 (0.20–0.86)	0.017
(TT + AA) ²	1 (1)	–	–	
(CT + TT + GA + AA) ³	18 (17.2)	28 (28.6)	0.44 (0.22–0.89)	0.012
Axin2 148 C/T and Axin2 2062 C/T	Cases (n (%), N = 97)	Controls (n (%), N = 104)	Adjusted OR (95% CI) ^a	p value ^b
(CC + CC) ⁰	75 (77.3)	71 (68.3)	1.00 (reference)	
(CT + CT) ¹	12 (12.3)	27 (26)	0.38 (0.17–0.83)	0.01
(CT + TT + CT + TT) ³	22 (10.4)	34 (32.7)	0.57 (0.30–1.09)	0.08
Axin2 148 C/T and Axin2 432 T/C	Cases (n (%), N = 97)	Controls (n (%), N = 85)	Adjusted OR (95% CI) ^a	p value ^b
(CC + TT) ⁰	94 (96.9)	75 (88.2)	1.00 (reference)	
(CT + TC) ¹	2 (2.1)	9 (10.6)	0.19 (0.03–0.93)	0.04
(CT + TT + TC + CC) ³	3 (3.0)	10 (11.8)	0.2 (0.06–0.97)	0.04
Axin2 148 C/T and Axin2 956 A/G	Cases (n (%), N = 113)	Controls (n (%), N = 102)	Adjusted OR (95% CI) ^a	p value ^b
(CC + AA) ⁰	82 (72.6)	63 (61.8)	1.00 (reference)	
(CT + AG) ¹	22 (19.5)	23 (22.5)	0.76 (0.38–1.55)	0.46
(TT + GG) ²	1 (0.9)	–	–	
(CT + TT + AG + GG) ³	31 (2.6)	39 (38.2)	0.63 (0.34–1.18)	0.13
Axin2 148 C/T and Axin2 1386 C/T	Cases (n (%), N = 121)	Controls (n (%), N = 109)	Adjusted OR (95% CI) ^a	p value ^b
(CC + CC) ⁰	29 (24)	24 (22)	1.00 (reference)	
(CT + CT) ¹	87 (72)	76 (69.7)	1.03 (0.52–2.01)	0.92
(TT + TT) ²	5 (4)	9 (8.3)	0.46 (0.13–1.66)	0.24
(CT + TT + CT + TT) ³	92 (76)	85 (78)	0.96 (0.50–1.86)	0.92
Axin2 148 C/T and Axin2 1712 + 19 G/T	Cases (n (%), N = 126)	Controls (n (%), N = 138)	Adjusted OR (95% CI) ^a	p value ^b
(CC + TT) ⁰	29 (23)	9 (6.5)	1.00 (reference)	
(CT + GT) ¹	97 (77)	129 (93.5)	0.21 (0.09–0.48)	0.003
Axin2 1365 G/A and Axin2 2062 C/T	Cases (n (%), N = 229)	Controls (n (%), N = 214)	Adjusted OR (95% CI) ^a	p value ^b
(GG + CC) ⁰	227 (99.1)	210 (98.1)	1.00 (reference)	
(GA + CT) ¹	2 (0.9)	4 (1.9)	0.43 (0.07–2.47)	0.34
Axin2 1365 G/A and Axin2 956 A/G	Cases (n (%), N = 235)	Controls (n (%), N = 209)	Adjusted OR (95% CI) ^a	p value ^b
(GG + AA) ⁰	228 (97)	201 (96.2)	1.00 (reference)	
(GA + AG) ¹	7 (3)	8 (3.8)	0.72 (0.25–2.06)	0.54
Axin2 1365 G/A and Axin2 1386 C/T	Cases (n (%), N = 103)	Controls (n (%), N = 113)	Adjusted OR (95% CI) ^a	p value ^b
(GG + CC) ⁰	84 (81.5)	83 (73.4)	1.00 (reference)	
(GA + CT) ¹	18 (17.5)	29 (25.7)	0.54 (0.27–1.08)	0.08
(AA + TT) ²	1 (1)	1 (0.9)	–	
(GA + AA + CT + TT) ³	19 (18.5)	30 (26.6)	0.57 (0.29–1.12)	0.10

(Continued)

Table 5. (Continued)

Axin2 1365 G/A and Axin2 1712 + 19 G/T	Cases (n (%), N=99)	Controls (n (%), N=69)	Adjusted OR (95% CI) ^a	p value ^b
(GG + TT) ⁰	78 (78.8)	30 (43.5)	1.00 (reference)	
(GA + GT) ¹	21 (21.2)	39 (56.5)	0.13 (0.06–0.29)	<0.0001
Axin2 2062 C/T and Axin2 432 T/C	Cases (n (%), N=255)	Controls (n (%), N=253)	Adjusted OR (95% CI) ^a	p value ^b
(CC + TT) ⁰	252 (98.8)	249 (98.4)	1.00 (reference)	
(CT + TC) ¹	3 (1.2)	4 (1.6)	0.72 (0.15–3.38)	0.68
Axin2 2062 C/T and Axin2 956 A/G	Cases (n (%), N=235)	Controls (n (%), N=214)	Adjusted OR (95% CI) ^a	p value ^b
(CC + AA) ⁰	222 (94.5)	209 (97.7)	1.00 (reference)	
(CT + AG) ¹	13 (5.5)	5 (2.3)	2.2 (0.77–6.49)	0.13
Axin2 432 T/C and Axin2 956 A/G	Cases (n (%), N=257)	Controls (n (%), N=240)	Adjusted OR (95% CI) ^a	p value ^b
(TT + AA) ⁰	252 (98)	236 (98.3)	1.00 (reference)	
(TC + AG) ¹	5 (2)	4 (1.7)	1.26 (0.32–4.88)	0.72
Axin2 1386 C/T and Axin2 956 A/G	Cases (n (%), N=114)	Controls (n (%), N=129)	Adjusted OR (95% CI) ^a	p value ^b
(CC + AA) ⁰	84 (73.7)	92 (71.3)	1.00 (reference)	
(CT + AG) ¹	30 (26.3)	37 (28.7)	0.84 (0.46–1.51)	0.56
Axin2 2062 C/T and Axin2 1712 + 19 G/T	Cases (n (%), N=110)	Controls (n (%), N=83)	Adjusted OR (95% CI) ^a	p value ^b
(CC + TT) ⁰	77 (70)	35 (42.2)	1.00 (reference)	
(CT + GT) ¹	33 (30)	48 (57.8)	0.19 (0.09–0.38)	<0.0001
Axin2 432 T/C and Axin2 1712 + 19 G/T	Cases (n (%), N=93)	Controls (n (%), N=58)	Adjusted OR (95% CI) ^a	p value ^b
(TT + TT) ⁰	86 (92.5)	42 (72.4)	1.00 (reference)	
(TC + GT) ¹	7 (7.5)	16 (27.6)	0.15 (0.05–0.45)	0.007
Axin2 1386 C/T and Axin2 1712 + 19 G/T	Cases (n (%), N=129)	Controls (n (%), N=164)	Adjusted OR (95% CI) ^a	p value ^b
(CC + TT) ⁰	18 (13.9)	14 (8.5)	1.00 (reference)	
(CT + TT + GT + GG) ³	111 (86.1)	150 (91.5)	0.52 (0.23–1.16)	0.11
Axin2 956 G/A and Axin2 1712 + 19 G/T	Cases (n (%), N=106)	Controls (n (%), N=83)	Adjusted OR (95% CI) ^a	p value ^b
(GG + TT) ⁰	74 (69.8)	35 (42.2)	1.00 (reference)	
(GA + GT) ¹	32 (30.2)	48 (57.8)	0.23 (0.12–0.46)	<0.0001

OR: odds ratio; CI: confidence interval; 0: wild genotype; 1: heterozygote genotype; 2: mutant genotype; 3: combined hetero and mutant genotype.

^aAdjusted OR, 95% CI, and their corresponding p values were calculated by unconditional logistic analysis after adjusting for age, gender, and smoking status.^bTwo-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

Table 6. Analyzed haplotype frequencies and risk toward lung cancer due to *Axin2* haplotypes.

Haplotype	Haplotype construction (148-1365-1386-1712-432- 2062-956)	Case frequency (n (%))	Control frequency (n (%))	OR (95% CI) ^a	p value ^b
Hap1	C G C G T C A	19.74 (0.033)	84.22 (0.138)	0.18 (0.11–0.30)	9.47e–013
Hap2	C G T G T C A	47.48 (0.078)	17.33 (0.028)	2.69 (1.53–4.74)	0.0003
Hap3	C G T T T C A	200.05 (0.330)	66.06 (0.108)	4.05 (2.95–5.57)	3.68e–019
Hap4	T G C G T C A	38.30 (0.063)	32.47 (0.053)	1.09 (0.67–1.78)	0.70
Hap5	T G C T T C A	13.52 (0.022)	124.80 (0.205)	0.07 (0.04–0.13)	5.35e–027
Hap6	T G T G T C A	68.48 (0.113)	38.10 (0.062)	1.76 (1.16–2.68)	0.007
Hap7	T G T T T C A	87.99 (0.145)	44.01 (0.072)	2.03 (1.38–2.99)	0.00027

OR: odds ratios; CI: confidence interval.

^aOdds ratios, 95% CI, and their corresponding p values were calculated by logistic regression analysis.

^bTwo-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls.

Table 7. Showing pairwise D' (r^2) values for the seven polymorphic sites among cases and controls.

Axin2 SNP	Axin2 1365	Axin2 1386	Axin2 1712	Axin2 432	Axin2 2062	Axin2 956
<i>Axin2 148</i>	0.067 (0.00)	0.110 (0.009)	0.028 (0.00)	0.628 (0.007)	0.233 (0.004)	0.068 (0.00)
<i>Axin2 1365</i>	–	0.130 (0.002)	0.005 (0.007)	0.892 (0.001)	0.277 (0.001)	0.007 (0.00)
<i>Axin2 1386</i>	–	–	0.213 (0.04)	0.297 (0.003)	0.081 (0.001)	0.194 (0.002)
<i>Axin2 1712</i>	–	–	–	0.795 (0.019)	0.354 (0.015)	0.088 (0.001)
<i>Axin2 432</i>	–	–	–	–	0.171 (0.007)	0.243 (0.012)
<i>Axin2 2062</i>	–	–	–	–	–	0.027 (0.001)

Also, the LD was eminently strong in case of *Axin2 148* and 432 ($D=0.628$ and $r^2=0.007$) and *Axin2 956* and 432 ($D=0.243$ and $r^2=0.012$). The pairwise LD is illustrated in Figure 1. The blocks represent the D' values. The red color shows the maximum LD between the two variants and white shows the least LD values.

MDR analysis

MDR uses a computational algorithm approach to investigate all the possible interactions between these susceptibility contributing factors which might include different SNPs, other genes, and some environmental determinants. The results obtained from MDR analysis have been tabulated in Table 8. The interpretation and selection of the best model are done on the basis of the CVC and average prediction error (1-testing accuracy). The best model can be determined if it has maximum CVC and minimum prediction error or maximum testing accuracy. In this study, the best interaction model (*Axin2 148*, *Axin2 2062*, and *Axin2 1712+19*) has CVC 10/10 and prediction error of 0.38 with a permutation p value of 0.02. Out of the three models given by the software, this interaction model gave the minimum prediction error, maximum CVC, and significant p value. It is clear from these results that these three SNPs are evidently contributing to disease susceptibility. However, in this study, the data have not been adjusted for covariates. So, further advanced analysis can reveal some interesting findings.

The entropy dendrogram obtained from this analysis is illustrated in Figure 2. Figure 2 suggests the degree and type of interaction taking place in between the SNPs which contribute in modulating the lung cancer susceptibility. As the variant *Axin2 1712+19 G>T* is linked with the other SNPs with a green-colored line, it shows that it has a redundant or very weak interaction with the other two SNPs, namely, *Axin2 148 C>T* and *2062 C>T*. However, these two variants were found to be closely related as they were lined with a red-colored line which shows synergy between the two. Also, the length of the line connecting these two SNPs is shorter, and hence, it is evident that they have a strong interaction.

CART analysis

CART analysis was used to find out the complex interactive patterns of all the polymorphic sites in this study. This analysis did reveal all the possible high-order interactions in between these variants of *Axin2* gene. For this interactive and complex algorithm, the mutant and heterozygous genotypes were combined as a single-variant genotype for the polymorphic sites. The decision tree gave 13 terminal nodes as shown in Figure 3. The splitting of the tree started from the *Axin2 1712+19 G>T*, a similar result was also observed in case of MDR result where *Axin2 1712+19 G>T* came out to be a single-risk factor and was the initiator of the best interaction model. Terminal node 13 (*Axin2 956+16 (W)/Axin2 432 (M)/Axin2 1712+19 (M)*) was

selected as the reference as it has lowest case rate of 14.2. All the other genotypes were compared with the genotype of this terminal node so as to obtain the OR associated with them. The OR and 95% CI associated with all the terminal nodes have been enlisted in Table 9. The subjects having the genotype as mentioned in terminal node 9 were found to possess maximum risk of developing lung cancer (OR=24, 95% CI: 1.68–34.08, and $p=0.01$). Certain other interactive combinations reported in terminal node 1 (OR=15.60, 95% CI: 3.29–73.84, and $p=0.0005$), terminal node 5 (OR=19.50, 95% CI: 3.00–126.48, and $p=0.001$), and terminal node 7 (OR=21, 95% CI:

2.39–183.94, and $p=0.005$) also exhibited a very high risk of acquiring lung cancer with highly significant p values as illustrated in Table 9. The OR obtained from this highly complex regression analysis was not adjusted for the various confounders such as age, sex, and smoking.

Discussion

In this study, the role of seven polymorphic sites, which includes five variants located in different exons of the *Axin2* gene, namely, Pro⁵⁰Ser in exon1, Ile¹⁴⁴Ile in exon 1, Pro⁴⁵⁵Pro in exon5, Pro⁴⁶²Pro in exon 5, and Leu⁶⁸⁸Leu in exon 7, and two other intronic variants *Axin2* 1712+19 and 956+16, has been evaluated in association toward lung cancer predisposition. The loss of heterozygosity in the region where genomic locus of *Axin2* lies has been reported in various cancers including ovarian, hepatocellular, and colorectal.¹⁵ There have been reports in ovarian and colon cancers which have cited that sequence variations in the functional domains of *Axin2* gene may affect its interaction with β -catenin and APC.³¹ The experiments involving the truncation of APC where the *Axin2* binding sites have been removed have reported its role in different human cancers.³² There are various domains present in *Axin2* protein, and the APC binding domain has significant homology with regulators of G-protein signaling (RGS); hence, it is referred as RGS domain.³³ In one of the most commonly studied SNP, *Axin2* 148 C>T (Pro⁵⁰Ser in exon1), there is substitution of proline with serine at codon 50 which lies in close proximity to RGS domain. Its SNP might play a role in modulating *Axin2* binding efficiency.²¹

It was keen to observe a protective effect in subjects who carried the variant (TT) genotype for *Axin2* 148 C>T ($p=0.01$). Many studies have evidently reported the same

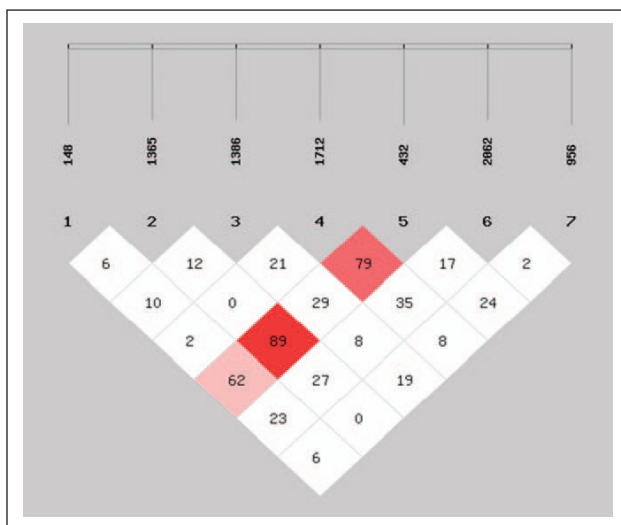


Figure 1. Showing pairwise linkage disequilibrium (D') between the seven variants of *Axin2* gene. red shows the maximum LD and white depicts the minimum LD between the polymorphic sites.

Table 8. Multifactor dimensionality reduction (MDR) analysis showing association of high-order interactions with lung cancer.

No. of risk factors	Best interaction model	Cross-validation consistency (CVC)	Prediction error	p value
1	<i>Axin2</i> 1712+19	10/10	0.41	0.08
2	<i>Axin2</i> 1712+19, <i>Axin2</i> 148	9/10	0.40	0.05
3	<i>Axin2</i> 148, <i>Axin2</i> 1712+19, <i>Axin2</i> 2062	10/10	0.38	0.02

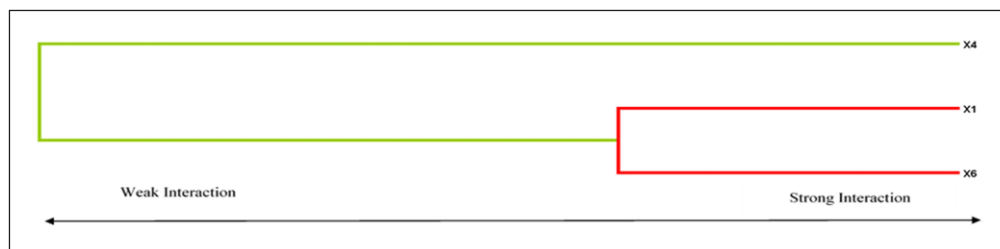


Figure 2. The multifactor dimensionality reduction (MDR) interaction dendrogram. X1: *Axin2* 148 C>T; X4: *Axin2* 1712+19 G>T; X6: *Axin2* 2062 C>T.

(Continued)

Table 9. (Continued)

Terminal node	Genotype for each terminal node	Control (n)	Case (n)	Case rate ^a	OR (95% CI)	p value ^b
Node 9	<i>Axin2</i> 2062 (M)/ <i>Axin2</i> 148 (W)/ <i>Axin2</i> 1365 (M)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1712 (M)	1	4	80	24.00 (1.68–34.08)	0.01
Node 10	<i>Axin2</i> 2062 (W)/ <i>Axin2</i> 148 (W)/ <i>Axin2</i> 1365 (M)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1712 (M)	8	5	38.4	3.75 (0.57–24.27)	0.16
Node 11	<i>Axin2</i> 148 (M)/ <i>Axin2</i> 1365 (M)/ <i>Axin2</i> 432 (W)/ <i>Axin2</i> 1712 (M)	33	17	34	3.09 (0.61–15.41)	0.16
Node 12	<i>Axin2</i> 956 (M)/ <i>Axin2</i> 432 (M)/ <i>Axin2</i> 1712 (M)	4	5	55.5	7.50 (1.02–54.98)	0.04
Node 13	<i>Axin2</i> 956 (W)/ <i>Axin2</i> 432 (M)/ <i>Axin2</i> 1712 (M)	12	2	14.2	1.0 ^c	

OR: odds ratio; CI: confidence interval.

^aCase rate is percentage LC cases among all individuals in each terminal node {case/(case + control) × 100}.

^b $p < 0.05$ is considered statistically significant.

^cReference group.

findings in case of various cancers. A study carried in Turkish population in lung cancer patients also gave the same insight about this polymorphic site where in the TT genotype conferred a decreased risk of lung cancer.²⁰ The similar results have been reported in Chinese²³ and Japanese lung cancer patients.²¹ Also, similar reports are available in case of other type of cancers including breast,³³ prostate,²⁴ astrocytoma,²⁵ and ovarian.²² This suggests that being a potent tumor suppressor gene, the variations in the genetic sequence might compensate on its function of regulating Wnt cascade as it is an indispensable part of the control regime of Wnt signalling.³⁴ Another variant lying in exon 5, *Axin2* 1365 $G > A$, has revealed a protective effect in subjects who were carrying a single copy of the susceptible allele. However, no other study has reported such an association in case of lung cancer, perhaps only one study has evaluated this SNP in lung cancer where no association was reported.²⁰ The reasons behind this difference might pertain to very low sample size and also the differences in ethnicity. Although, this variant is also studied in case of Turkish astrocytoma patients where no association was observed for this polymorphic site in regard of disease susceptibility. Another polymorphic site which was found to have a very pronounced protective effect in our study was *Axin2* 1712+19 $G > T$, which is located in intron 5 of the *Axin2* gene. *Axin2* 1386 and 1712+19 showed a certain deviation from HWE. The reasons behind this slight deviation might be population stratification, random mating, and sample bias for these two variants. The intronic variants may have been examined to identify their role in influencing the m-RNA formation. These variations may lead to differential splicing pattern, and henceforth, it affects the primary structure of the protein.³⁵ However, the studies done in Turkish lung cancer²⁰ and astrocytoma patients²⁵ are not concordant with the findings of our study.

In case of two polymorphic sites, namely, *Axin2* 956+16 $A > G$ and 2062 $C > T$, no effect was observed on lung cancer risk in this study. However, these two variants are correlated with the tooth agenesis in Brazilian population.³⁶

Genetic variant 956+16 $A > G$ lies in intron2 and has been reported to influence the splicing process at exon2, although this interaction is weak.³⁷ Similarly, 2062 $C > T$, which is although a synonymous substitution of leucine with leucine at 688 codon has been proposed to effect the splicing process by disrupting the enhancer sequences present just before the splice site in exon7.³⁷ In line with our findings, these genetic variants did not confer any modulation in Turkish lung cancer patients.²⁰

Furthermore, in this study, lung cancer risk was also evaluated in various subgroups made on the basis of histological subtype. Being a diverse type of carcinoma, these subtypes of lung cancer have been predicted to have different disease etiology and pathogenesis.³⁸ In our study, the *Axin2* 148 $C > T$ was found to harbor a protective effect in SQCC patients ($p = 0.03$). These results are concordant with those observed in Turkish lung cancer patients.²⁰ In case of 1365 $G > A$, the Turkish population did not exhibit any significant association in any of the subgroups,²⁰ and our study also revealed similar results for SQCC patients, but a slight association was observed in SCLC patients. For all other five variants, the results from the study carried on Turkish lung cancer population suggest that they do not contribute in modulating the risk of the disease particularly in any histological subgroup.²⁰ Our findings are in line with this study except in the case of *Axin2* 1712+19 $G > T$ where a decreased risk was observed in all the three subgroups.

Smoking has always been considered as one of the major factors contributing toward the risk of developing lung cancer. This factor may interact with other genetic variabilities differently to modulate the individual's predisposition toward the oncogenesis of lung tissue. This study evaluated the risk associated with these polymorphic sites in smokers and non-smokers separately. A previous study done in Turkey has reported a lower risk of lung cancer in case of smokers having mutant (TT) genotype of *Axin2* 148 $C > T$. Our findings are also in accordance to the results of this study. This previous study has examined the

risk of all the seven polymorphic sites under study in smokers and non-smokers. No effect has been observed for any of the category in case of *Axin2* 432 T>C, 956+16 A>G, 2062 C>T, 1386 C>T, and 1712+19 G>A.²⁰ Our study also revealed no significant correlation in smokers and non-smokers for above-mentioned sites except for 1712+19 G>T, where smokers possessing GT genotype showed a protective association.

Haplotype and LD analysis has revealed the status of the linkage of these seven genetic variants present on *Axin2* gene. This is probably the first attempt to find out the haplotype blocks comprising the seven alleles which significantly contribute to lung cancer risk in North Indians. The Hap block which showed the maximum risk toward lung cancer comprised the variant allele from both *Axin2* 1386 C>T and 1712+19 G>T, and all other alleles were wild type. Recent study done in Chinese NSCLC patients has examined LD and risk associated with the different haplotypes of *Axin2* gene (rs2240308, rs22040307) and MMP7 (rs4791169).³⁹ A similar kind of study was also carried out in ovarian cancer patients of Poland, where haplotype analysis was done for four different SNPs of *Axin2* gene.²²

MDR approach was also applied to find the various non-parametric interactions taking place in between these polymorphic variants. None of the previous studies have analyzed such high-order interactions for all the seven polymorphic sites of *Axin2* gene. This analysis revealed the best interaction model (*Axin2* 148, *Axin2* 2062, and *Axin2* 1712+19) with maximum CVC (10/10) and minimum prediction error (0.38) along with significant permutation *p* value. Very few reports are available where MDR was employed to find out these interactions; one such recent study was done in Chinese population where *Axin2* gene (rs2240308 and rs22040307) and MMP7 (rs4791169) and smoking were taken as a parameter to find out the best interaction model contributing toward lung cancer risk.³⁹ Another study also used this exhaustive MDR method in ovarian cancer to find out the interactions between variants of *Axin2*, β -catenin, and APC.²²

So far, this is the first study where high-order regression analysis was carried out to obtain the decision tree and odds associated with combinations of various genotypes using CART analysis. This was the first attempt to look into the SNP–SNP interaction taking place between the *Axin2* genetic variants and their contribution in modulating lung cancer predisposition as *Axin2* is a potent tumor suppressor gene. This was evident from the results of this study that *Axin2* 148 C>T and *Axin2* 1712+19 G>T are the major role-playing SNPs in the interactive models obtained both from MDR and CART. These kinds of novel innovative findings can help in defining the role of these sequence variants within the tumor suppressor genes in the process of carcinogenesis. This study has given an insight about the genetic variants of *Axin2* protein, which is an

indispensable participant of the destruction complex maintaining β -catenin levels and hence regulating Wnt pathway. However, the functional significance of these variations is yet unclear. Studies in near future with larger sample size and in different ethnicities can support these findings.

Conclusion

Evidences from this study have exhibited an interesting correlation of these seven genetic variants with lung cancer susceptibility. The data obtained from this study reveals a strong association of these SNPs with lung cancer risk in North Indians. This study has employed certain innovative high-order computational algorithms such as MDR and CART to find the interactions between these polymorphic sites and their contribution in lung cancer predisposition.

Acknowledgements

We would like to express our gratitude to all the subjects who participated in this current study. Dr Digamber Behra and Dr Navneet Singh have provided access to data, helped in the approval of study, and guided in the analysis of data. Dr Siddharth Sharma has designed the study, supervised, and edited the manuscript. Charu Bahl has carried out all the experimental work and statistical analysis and drafted the manuscript under the esteemed supervision of Dr Siddharth Sharma, who is the corresponding author for this research work and also the principal investigator.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by grant from the Indian Council of Medical Research, New Delhi, India (Grant No. 5/13/126/2011/NCD-III).

References

1. Yardy GW and Brewster SF. Wnt signalling and prostate cancer. *Prostate Cancer Prostatic Dis* 2005; 1: 119–126.
2. Luo J, Chen J, Deng Z-L, et al. Wnt signaling and human diseases: what are the therapeutic implications? *Lab Invest* 2007; 87: 97–103.
3. Akiyama T. Wnt/ β -catenin signaling. *Cytokine Growth Factor Rev* 2000; 11: 273–282.
4. Stewart DJ. Wnt signaling pathway in non-small cell lung cancer. *J Natl Cancer Inst* 2014; 106: 1–11.
5. Kikuchi A. Modulation of Wnt signalling by Axin and Axil. *Cytokine Growth Factor Rev* 1999; 10: 255–265.
6. Dong X, Seelan RS, Qian C, et al. Genomic structure, chromosome mapping and expression analysis of the human AXIN2 gene. *Cytogenet Cell Genet* 2001; 28: 26–28.
7. Yang Y. Wnt signaling in development and disease. *Cell Biosci* 2012; 2: 1

8. Tetsu O and McCormick F. β -catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature* 1999; 398: 422–426.
9. Barker N and Clevers H. Mining the Wnt pathway for cancer therapeutics. *Nature* 2007; 5: 997–1015.
10. Polakis P. Wnt signalling and cancer. *Genes Dev* 2000; 14: 1837–1851.
11. Winn RA, Bremnes RM, Bemis L, et al. γ -Catenin expression is reduced or absent in a subset of human lung cancers and re-expression inhibits transformed cell growth. *Oncogene* 2002; 21: 7497–7506.
12. Mazieres J, He B, You L, et al. Wnt signaling in lung cancer. *Cancer Lett* 2005; 222: 1–10.
13. Huang H and He X. Wnt/ β -catenin signaling: new (and old) players and new insights. *Curr Opin Cell Biol* 2008; 20: 119–125.
14. Davoodi B, Kh O and Heydari Nasrabadi M. Genetic polymorphism of the Axin2 gene in Wnt pathway in patients with ovarian cancer. *J Ardabil Univ Med Sci* 2015; 15: 220–227.
15. Phelan CM, Borg A, Cuny M, et al. Consortium study on 1280 breast carcinoma; allelic loss on chromosome 17 targets subregions associated with family history and clinical parameters. *Cancer Res* 2000; 58: 1004–1012.
16. Gong J, Jiang Y, Hao N, et al. Quantitative assessment of the association between AXIN2 rs2240308 polymorphism and cancer risk. *Sci Rep* 2015; 5: 1–7.
17. Tseng RC, Lin RK, Wen CK, et al. Epigenetic silencing of AXIN2/ β -catenin and deregulation of p53-mediated control lead to wild-type β -catenin nuclear accumulation in lung tumorigenesis. *Oncogene* 2008; 27: 4488–4496.
18. Marshall AL and Christiani DC. Genetic susceptibility to lung cancer—light at the end of the tunnel?. *Carcinogenesis* 2013; 34: 487–502.
19. Wu Z, Sun Y, Tang S, et al. AXIN2 rs2240308 polymorphism contributes to increased cancer risk: evidence based on a meta-analysis. *Cancer Cell Int* 2015; 15: 68.
20. Gunes EG, Pinarbasi E, Pinarbasi H, et al. Strong association between lung cancer and the AXIN2 polymorphism. *Mol Med Rep* 2009; 2: 1029–1035.
21. Kanzaki H, Ouchida M and Hanafusa H. Single nucleotide polymorphism of the AXIN2 gene is preferentially associated with human lung cancer risk in a Japanese population. *Int J Mol Med* 2006; 18: 279–284.
22. Mostowska A, Pawlik P and Sajdak S. An analysis of polymorphisms within the Wnt signaling pathway in relation to ovarian cancer risk in a Polish population. *Mol Diagn Ther* 2014; 18: 85–91.
23. Liu D, Li L, Yang Y, et al. The Axin2 rs2240308 polymorphism and susceptibility to lung cancer in a Chinese population. *Tumor Biol* 2014; 35: 10987–10991.
24. Ma C, Liu C, Huang P, et al. Significant association between the Axin2 rs2240308 single nucleotide polymorphism and the incidence of prostate cancer. *Oncol Lett* 2014; 8: 789–794.
25. Gunes EG, Pinarbasi E and Pinarbasi H. AXIN2 polymorphism and its association with astrocytoma in a Turkish population. *Mol Med Rep* 2010; 3: 705–709.
26. Sobti RC, Sharma S, Joshi A, et al. Genetic polymorphism of the *CYP1A1*, *CYP 2E1*, *GSTM1* & *GSTT1* genes and lung cancer susceptibility in a north Indian population. *Mol Cell Biochem* 2004; 266: 1–9.
27. Páez D, Gerger A, Zhang W, et al. Association of common gene variants in the WNT/ β -catenin pathway with colon cancer recurrence. *Pharmacogenomics J* 2014; 14: 142–150.
28. Shi YY and He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res* 2005; 15: 97–98.
29. Ritchie MD and Motsinger AA. Multifactor dimensionality reduction for detecting gene–gene and gene–environment interactions in pharmacogenomics studies. *Future Med* 2005; 6: 823–834.
30. Srivastava A, Sharma KL, Srivastava N, et al. Significant role of estrogen and progesterone receptor sequence variants in gallbladder cancer predisposition: a multi-analytical strategy. *PLoS One* 2012; 7: e40162.
31. Rubinfeld B, Robbins P, El-Gamil M, et al. Stabilization of β -catenin by genetic defects in melanoma cell lines. *Science* 1997; 275: 1790–1792.
32. Aristizabal-Pachon AF, Carvalho TI, Humberto H, et al. AXIN2 polymorphisms, the β -catenin destruction complex expression profile and breast cancer susceptibility. *Asian Pac J Cancer P* 2015; 16: 7277–7284.
33. Ikeda S, Kishida S, Yamamoto H, et al. Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3 β and β -catenin and promotes GSK-3 β -dependent phosphorylation of β -catenin. *EMBO J* 1998; 17: 1371–1384.
34. Willert K, Shibamoto S and Nusse R. Wnt-induced dephosphorylation of axin releases β -catenin from the axin complex. *Genes Dev* 1999; 13: 1768–1773.
35. Shen LX, Basilion JP and Stanton VP. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc Natl Acad Sci USA* 1999; 96: 7871–7876.
36. Mostowska A, Biedziak B and Jagodzinski PP. Axis inhibition protein 2 (AXIN 2) polymorphisms may be a risk factor for selective tooth agenesis. *J Hum Genet* 2006; 51: 262–266.
37. Callahan N, Modesto A, Meira R, et al. Axis inhibition Protein2 (AXIN2) and tooth agenesis. *Arch Oral Biol* 2009; 54: 45–49.
38. Sodhi KK, Bahl C, Singh N, et al. Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians. *Future Oncol* 2015; 11: 2159–2173.
39. Han S, Lv L, Wang X, et al. Association of AXIN2 and MMP7 polymorphisms with non-small cell lung cancer in Chinese Han population. *Int J Clin Exp Pathol* 2016; 9: 2253–2258.

RESEARCH ARTICLE

For reprint orders, please contact: reprints@futuremedicine.com

Association of polymorphisms in *Dickopf* (*DKK*) gene towards modulating risk for lung cancer in north Indians

Charu Bahl¹, Navneet Singh², Digambar Behera² & Siddharth Sharma^{*1}

Aim: To investigate the association between the genetic variants of *DKK4* (rs2073664), *DKK3* (rs2291599, rs3206824 and rs7391689) and *DKK2* (rs447372, rs419558 and rs17037102) and lung cancer predisposition in north Indians. **Materials & methods:** A total of 600 subjects were genotyped using PCR restriction fragment length polymorphism technique. Association analysis was carried out using logistic regression approach. **Results:** *DKK3* rs7396187 showed a protective effect ($p = 0.01$). Subjects with heterozygous genotype of *DKK2* rs17037102 and rs419558 showed an increased risk. The variant genotype of the genotypic combination, *DKK3* rs3206824 and *DKK2* rs419558 showed a twofold increased risk ($p = 0.008$). Lung cancer risk increased four-times in subjects with variant genotype of all the three *DKK2* variants. **Conclusion:** *DKK2* is certainly playing a crucial role in modulating cancer susceptibility.

First draft submitted: 9 March 2016; Accepted for publication: 31 August 2016; Published online: 19 September 2016

Lung cancer is characterized by complex molecular perturbations which are a result of dysregulated expression of genes involved in the process of cell proliferation, differentiation and apoptosis [1]. Beside smoking and other environmental exposures, the genetic, epigenetic alterations and variations in genes encoding proteins involved in signaling pathways also lead to molecular abnormalities which cater to the process of lung carcinogenesis [2]. Certain signaling cascades have shown some promising evidences which would help to understand the disease pathology to a better extent. One of these cascades is the Wnt/ β -catenin signaling pathway that has been reported to play an important role in the process of oncogenesis and particularly lung cancer development [3]. Wnt signaling pathway is involved in various processes such as cell proliferation, survival, polarity, adhesion and development of various tissues by altering the transcription of downstream targets such as *cyclin D1*, *VEGF-A*, *c-myc* among others [4]. Categorically, Wnt pathway is broadly classified into canonical and noncanonical pathway. Focusing on the canonical pathway, which is centrally mediated by β -catenin levels in the cytoplasm, the activation of this pathway is augmented by binding of the Wnt ligands to the frizzled receptors and LRP5/6 complex, therefore these interactions lead to inhibition of phosphorylation of β -catenin by GSK-3 β within a large cytoplasmic complex including dishevelled, adenomatous polyposis coli and Axin [5]. This leads to the destruction of this degradation complex and hence the ubiquitination of the β -catenin is prevented and it is translocated to the nucleus where it interacts with the transcription factors *TCF/LEF* which activate target genes such as matrix metalloproteins, *Cox-2*, *cyclin D1*, *VEGF*, *c-myc* among others. The aberrant activation of Wnt signaling pathway has been related to human cancers such as colorectal, melanoma,

KEYWORDS

• DKK • lung cancer • Wnt pathway

¹Department of Biotechnology, Thapar University, Patiala, Punjab-147002, India

²Department of Pulmonary Medicine, Post Graduate Institute of Education & Medical Research (PGIMER), Sector 14, Chandigarh, India

*Author for correspondence: siddharthsharma.phd@thapar.edu

gastric, hepatocellular, prostate, ovarian and breast carcinoma [6]. Overexpression of Wnt target genes have been confirmed in various tumor tissues [7]. The internalization of the β -catenin to the nucleus is prohibited by the regulation of the Wnt signaling by the Wnt antagonists, which block the signaling cascade by allowing the Axin protein to promote the ubiquitination of β -catenin [8]. Wnt antagonists include the Dickkopf (DKK) family of proteins, that are secreted extracellular glycoproteins and antagonize the Wnt signaling by binding to co-receptors LRP5/6 thereby preventing the formation of active Wnt–Frizzled–LRP5/6 complexes [9]. DKK family includes four proteins DKK1–DKK4 which are identified by two cysteine rich domains separated by a nonconservative linker region where in the location of ten cysteine residues is highly conserved within the DKK family and it also possess a N-terminal signal peptide [10]. The *DKK* genes are located on different chromosomes, namely *DKK2* located at 4q25 and contains four exons whereas *DKK3* is located at 11p15 and is comprised of seven exonic regions. *DKK4* is located at 8p11 and has seven exons [11]. Recent studies have demonstrated that *DKK3* is downregulated in case of breast, cervical, non-small-cell lung cancer, prostate and renal cancer [12]. Also, the epigenetic hypermethylation of the *DKK3* promoter has been observed in high percentage of lung tumors [13]. The expression of *DKK2* gene are reported to decrease in melanoma and gastric cancer [14,15]. There are evidences supporting the oncogenic role of *DKK3* gene in process of lung cancer progression which is frequently inactivated in lung cancer [16]. Certain epigenetic alterations have also been seen in renal carcinoma in context of *DKK2* gene [17].

Single point variations in the gene sequence may confer changes in the mRNA structure/function and are also responsible for certain phenotypic expressions such as disease susceptibility and drug resistance [18]. Several SNPs of the *DKK3* gene such as rs2291599, rs3206824 and rs7391689 have been screened in various cancers such as renal [19], prostate cancer [20], breast [21] and also in lung cancer [22]. Similarly, certain nonsynonymous substitutions in the *DKK2* gene have been found to be implicated in patients with renal [19], colorectal [23] and breast cancer [21]. Also, the polymorphic site rs2073664 (C/T) in *DKK4* gene has been studied in renal carcinoma patients [19] and another

SNP (rs3763511) located within the *DKK4* gene has been investigated in breast cancer subjects of Saudi Arabia [21]. The constitutive activation of the Wnt pathway may be attributed to the sequence variation in the genes coding the Wnt signal regulating proteins namely Dickkopf proteins which prevent the nuclear translocation of β -catenin.

Considering all this information, we thought it would be interesting to screen some of the genetic variability within the Wnt pathway for any evidence to show whether the polymorphic variants within the *DKK* genes plays a role in altering the predisposition of north Indians toward lung cancer. As far our knowledge is concerned, this is the first attempt to investigate the role of polymorphic sites of *DKK3* (rs2291599, rs3206824 and rs7391689), *DKK2* (rs447372, rs419558 and rs17037102) and *DKK4* (rs2073664) genes toward modulating the lung cancer susceptibility of north Indians. In this study, attempt was made to find the susceptible subgroups within the sample population by stratifying the population on the basis of their histology, clinic-pathological features and smoking habits. Furthermore, we also tested whether haplotypes of *DKK* gene contribute to host susceptibility to lung cancer.

Materials & methods

• Study population

Three hundred lung cancer cases and 300 controls were enrolled in the present study, the sample collection was carried out from the Department of Pulmonary Medicine, Post Graduate Institute of Medical Education & Research (PGIMER), Chandigarh. A written consent was obtained from each subject and also a questionnaire was filled by a trained personnel to obtain the various epidemiological parameters such as age, gender, smoking history, etc. Ethical clearance was obtained from the Ethics Committee of PGIMER to carry out this study. A 5-ml blood sample was collected in ethylenediaminetetraacetic acid (EDTA) vials from each subject. Only the newly diagnosed lung cancer patients were included in the study with no age, gender, stage, histology and TNM restrictions. Patients having history of any other malignancy were excluded from the study. The controls were selected in a manner so as to pairwise match the age (± 10 years), gender and smoking habits with the cases. This was done in order to avoid any sampling bias. Pack-years were obtained using

this equation: [(cigarettes or beedis (type of Indian cigarette) per day/20) × number of years smoked). The clinical details of the patients such as histology, TNM staging, metastasis and other important information was collected from the medical records of the patients from the hospital.

DNA isolation

The 5-ml blood collected from the subjects was used to isolate the genomic DNA. Genomic DNA was isolated using the phenol chloroform extraction procedure as described by Sodhi *et al.* [24].

Genotyping of DKK variants using PCR restriction fragment length polymorphism

Genotyping analysis was performed using PCR restriction fragment length polymorphism technique using the primers and restriction enzymes as previously reported by Hirata *et al.* [19], which are tabulated in **Table 1**. A nested PCR approach was used for all the polymorphic sites. The annealing temperature used in the first round was 52°C and the amplicon of first was amplified further in the second round using 58°C annealing temperature. The PCR mixture of 25 µl consists of 1× PCR buffer, 1.5 mM MgCl₂, with 0.5 µM of both forward and reverse primer, 200 µM of each dNTPs, 100 µg/ml bovine serum albumin and 1 U Taq polymerase (DNAzyme II DNA Polymerase, ThermoScientific, MA, USA) and approximately 200 ng DNA. The PCR product was digested with appropriate restriction enzyme as shown in **Table 1** and then run on agarose gel for analysis. Scoring of gels was done by two independent observers so as to avoid any bias. This scoring done was according to the restriction pattern for every SNP which was based on the generation or removal of the restriction site in the presence or absence of the variant allele. The genotyping analysis was carried out without knowing the status of the sample whether it is a case or control. 15% samples were randomly genotyped twice to confirm the reproducibility of the results which was found to be 100%. The electrophoretic map has been illustrated in **Figure 1**.

Statistical analysis

Demographic variables such as age, gender, smoking status and pack years were classified into categorical and continuous variables. The continuous variables such as age and pack years were compared among cases and controls using

t-test statistics and categorical variables such as gender and smoking status were analyzed to find the difference between cases and controls using χ^2 test. The Hardy–Weinberg equilibrium (HWE) theory ($p^2 + 2pq + q^2 = 1$; where *p* is the frequency of the wild-type allele and *q* is the frequency of the variant allele) was used for controls to calculate the genotype frequencies of all the polymorphisms using χ^2 test. Pearson's χ^2 test was used to compare the allelic and genotypic frequencies between cases and controls to find whether significant difference occurs between the two groups. The association analysis was carried out using unconditional logistic regression to find the odds ratio (OR) along with 95% CIs adjusted for age, gender and smoking as they can act as confounding factors. Further, to find out the high risk subgroups, the sample population was stratified into different categories based on their histology, smoking status and also clinic-pathological factors [24]. Haplotype analysis for the both SNPs was performed and the *D'* value and *r*² was calculated with the SHEsis software [25] [26]. All *p*-values were two sided, and a *p*-value of <0.05 was considered statistically significant. All the statistical analyses were performed with MedCalc Statistical Software version 15.11.4 (MedCalc Software bvba, Ostend, Belgium; 2015) [27].

Results

• Demographic characteristics of the sample population

The sample population can be characterized by various epidemiological and demographic parameters including age, gender, smoking status and the clinical specifications of the cases as tabulated in **Table 2**. The mean age came out to be 57.3 ± 10.74 years and 53.23 ± 10.44 years in cases and controls, respectively. The study was performed on 300 cases and 300 controls including 258 (86%) males in cases and 253 (84.3%) males in controls whereas 42 (14%) and 47 (15.7%) were females both in case and control groups, respectively. No significant difference was found between the distribution of males and females in cases and controls (*p* = 0.64), hence there is no sampling bias based on gender. The number of smokers and nonsmokers were also compared between the two study groups. The case group comprised 78% smokers whereas controls had only 71% smokers, a significant difference was observed in the proportion of smokers and nonsmokers (*p* = 0.01). The attempt to

Table 1. List of primers and restriction enzymes used.

Gene, SNP ID	Substitution	Primer sequence	Product size (bp)	Restriction enzyme	Restriction pattern
DKK4, rs2073664 C/T	Synonymous	First: F5-GCC ATG GCA TTA CTG CTT TT-3; R5-ATT GCT GGT CAA TTG GCT TC-3	292	EcoNI	CC (223, 69 bp), CT (292, 223, 69 bp), TT (292 bp)
	Exon 4 V ¹⁶⁹ V	Second: F5-CTG CGT GCT GTG TCT GTT TT-3; R5-AAC GCT GGA AGA TTT CTG GA-3			
DKK3, rs2291599 C/T	Intron 4 C/T	First: F5-CAG AGG ACA TGG GGT GGA T-3; R5-CTG CTG CTC TGC TCT CCA TT-3	250	AvaI	CC (120, 130 bp), CT (250,120,130 bp), TT (250 bp)
		Second: F5-CTT GTC CTC CAG GAG TCA GC-3; R5-CAT CAT CGA CGA GGA CTG TG-3			
DKK3, rs3206824, G/A	Nonsynonymous	First: F5-GAG GTC CCC GAT GAG TAT GA-3; R5-TAG GAA GAA GCC TGG TCA GC-3	209	DdeI	GG (209 bp), AG (209, 114, 97 bp), AA (114, 95 bp)
	Exon7 Arg ³³⁵ Gly	Second: F5-GGT CCC CGA TGA GTA TGA AG-3; R5-AGC ACA CAC CTG GGG AAA TA-3			
DKK3, rs7396187 G/C	Intron 4 G/C	First: F5-TTC CTT AGG TCC CTA GGT CCA-3; R5-AGG GCA AAG GAG ACT CTT CA-3	245	Fnu4HI	GG (245 bp), GC (170, 75 bp), CC (170, 75 bp)
		Second: F5-ACA GGG CAT GGC AGT TAG AG-3; R5-CTC TTC ACC CAA CAG GCA TT-3			
DKK2, rs447372 A/G	Intron 1	First: F5-TGC TCA AAA AGC AGC ATC TC-3; R5-AGC TAA CGG TCA CTC AAC CAA-3	249	EcoRV	AA (249 bp), AG (249, 119, 130 bp), GG (119, 130 bp)
		Second: F5-TCA AAA AGC AGC ATC TCA GG-3; R5-CTC CAA ATT TTG CCC TAC CA-3			
DKK2, rs419558 C/T	3'-UTR	First: F5- TACGGACACAGGACCTCACA -3; R5-TCC AAT ATA TGT GGG AAA AGA GC-3	251	AflII	CC (251 bp), TT (121, 130 bp), CT (251, 121, 130 bp)
		Second: F5-TCC CCT GGT TTC AAA GAT GA-3; R5-TGG GAA AAG AGC TAA CAG AGA GA-3			
DKK2, rs17037102 G/A	Nonsynonymous	First: F5-TGG CTT CAT ATT TCA CAT CAA GA-3; R5-TGT GTG GTC TTC CTA GAT TCT GC-3	132	DdeI	GG (76, 66 bp), GA (132, 76, 66 bp), AA (76, 66 bp)
	Exon 3 Arg ¹⁴⁶ Glu	Second: F5-TGA TCA TCT CCA GGC ATC TG-3; R5-ATT CTG CCA TCC CAA GTC AT-3			

adequately match the number of smokers and nonsmokers was made but as established smoking is a significant parameter responsible for occurrence of lung, same was observed in this sample population also. Smoking was quantified in terms of pack years ([cigarettes or beedis per day/20] × years smoked), the means of pack years were also compared among cases and controls and a significant difference was observed (p = 0.003). The parameters such as age, gender and smoking were adjusted for genotype analysis using unconditional logistic regression as they can act as confounding factors. Further, the different clinical specifications such as histology, clinical stage, tumor size extensions, lymph node invasion and distant metastasis were also recorded from medical reports. Out of the 300 cases, 128 (42.8%) were histologically identified as squamous cell carcinoma (SQCC), 97 (32.3%) as adenocarcinoma (ADCC) and 70 (23.3%) as small-cell-lung carcinoma (SCLC). The TNM data were available for 270 cases and when classified on the basis of clinical stage, only

few cases were reported with stage I and II and most of them were categorized into stage III and IV (I: 3 II: 12 III: 140 IV: 123) as shown in Table 2.

Genotypic & allelic frequencies of DKK polymorphic variants among cases & controls & their role in modifying the susceptibility toward lung cancer

Table 3 summarizes the allelic and genotypic frequencies for all the seven variants of DKK gene including DKK2, DKK3 and DKK4. In case of DKK4 rs2073664 (C/T) no significant difference between the distribution of genotypic frequencies was found between cases and controls ($\chi^2 = 3.7$, df = 2, p = 0.15). The percentage of subjects carrying heterozygous CT genotype were over represented in controls (47.7%) as compared with cases (45.6%). The minor allele frequency (MAF) was found out to be 0.28 and 0.26 in controls and cases, respectively. When the odds were calculated using wild genotype (CC) as reference, no significant association was

seen in subjects carrying one susceptible allele (CT) genotype (OR: 0.95; 95% CI: 0.68–1.34; $p = 0.80$) however, a protective effect was observed in subjects with both susceptible alleles (TT) genotype (OR: 0.37; 95% CI: 0.14–0.95; $p = 0.03$).

Various polymorphic sites of *DKK3* gene including rs2291599, rs3206824 and rs7391689 were also investigated for their role in modulating lung cancer susceptibility. For *DKK3* rs2291599 (C/T), the genotypic and allelic frequency of control subjects was found to follow HWE ($\chi^2 = 1.85$, $df = 2$, $p = 0.17$). The proportion of subjects with wild (CC) genotype was found to be similar in both cases and controls (65%) however, the percentage of subjects carrying heterozygous (CT) genotype were found to be over represented in controls than cases (29.7 vs 27%).

The difference between the genotypic frequencies between cases and controls was not found to be significant ($\chi^2 = 1.97$, $df = 2$, $p = 0.37$). MAF was 0.21 in cases and 0.20 in controls. The association analysis revealed that subjects with TT genotype were at a slightly higher risk (OR: 1.46; 95% CI: 0.74–2.88; $p = 0.26$) of acquiring lung cancer but it was not statistically significant. The next variant of *DKK3* gene which was analyzed is rs3206824 (G/A). The genotypic frequencies for *DKK3* rs3206824 (G/A) in control group were following the HWE ($\chi^2 = 0.72$, $df = 2$, $p = 0.39$). No significant difference was observed in genotypic frequencies among cases and controls ($\chi^2 = 0.80$, $df = 2$, $p = 0.66$). The MAF was found to be 0.17 and 0.15 in cases and controls, respectively. The logistic regression approach revealed that the subjects carrying

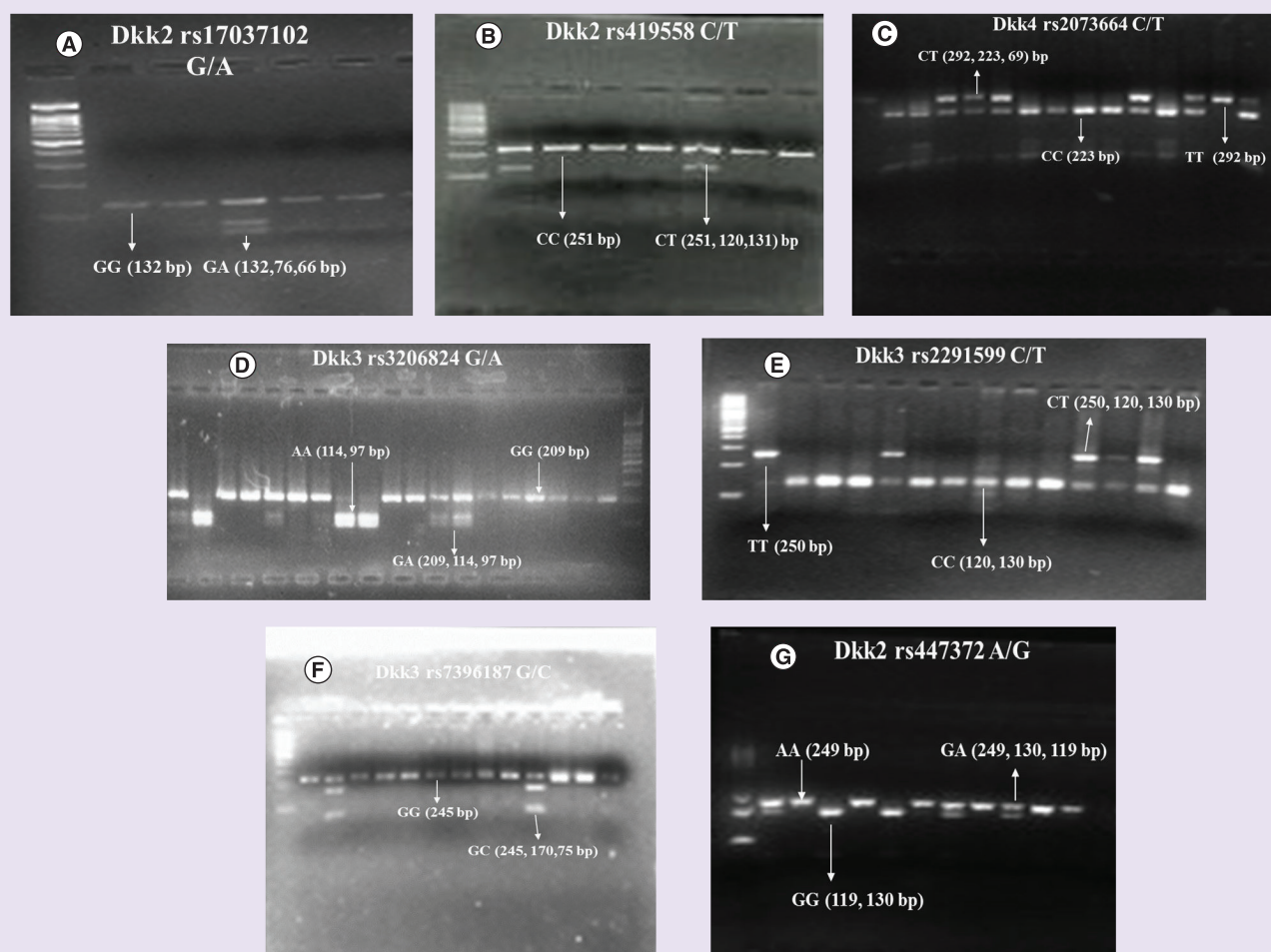


Figure 1. Representative example of Restriction patterns of DKK variants. (A) DKK2 rs17037102. (B) DKK2 rs419558. (C) DKK3 rs2073664. (D) DKK3 rs3206824. (E) DKK3 rs2291599. (F) DKK3 rs7396187. (G) DKK2 rs447372.

Table 2. Demographic characteristics among cases and controls.

Variable	Cases (n = 300), n (%)	Controls (n = 300), n (%)	p-value [†]
Age (years):			p < 0.0001
– Mean ± SD	57.38 ± 10.74	53.23 ± 10.44	
Gender:			0.645
– Male	258 (86)	253 (84.3)	
– Female	42 (14)	47 (15.7)	
Smoking status:			0.01
– Smokers	235 (78.3)	208 (71.4)	
– Nonsmokers	65 (21.7)	92 (28.6)	
Pack years:			0.0003
– Mean ± SD	28.09 ± 34.09	19.4 ± 19.56	
Histological types:			
– SQCC	128 (42.8)		
– ADCC	97 (32.3)		
– SCLC	70 (23.3)		
– Others	4 (1.3)		
– Unknown	1 (0.3)		
TNM staging:			
– I	3 (1)		
– II	12 (4)		
– III	140 (46.7)		
– IV	123 (41)		
– Unclassified	22 (7.3)		
Tumor size:			
– Tx	10 (3.3)		
– T1	14 (4.7)		
– T2	39 (13)		
– T3	79 (26.3)		
– T4	128 (42.7)		
– Unknown	30 (10)		
Lymph node:			
– N0	48 (16)		
– N1	36 (12)		
– N2	112 (37.3)		
– N3	72 (24)		
– Unknown	30 (10)		
Metastasis:			
– M0	154 (51.3)		
– M1	116 (38.7)		
– Unknown	30 (10)		
Student's t-test was used for age. All p-values are two-sided. Bold p < 0.05 was considered statistically significant.			
[†] p-values were derived from Pearson χ^2 test, except age.			
ADCC: Adenocarcinoma; SD: Standard deviation; SCLC: Small-cell-lung carcinoma; SQCC: Squamous cell carcinoma.			

mutant (AA) genotype did not have any significant risk of developing lung cancer (OR: 1.29; 95% CI: 0.38–4.39; p = 0.68). The genotypic frequencies for the last variant of *DKK3* gene rs7391689 (G/C) were also tabulated as shown in Table 3. Since the frequency of the mutant (CC) genotype was less represented (0.3%), we combined the heterozygous and mutant as a single genotype (GC + CC). The proportion of subjects with variant rs7391689 alleles (GC + CC) were found to be slightly higher in the control group as compared with cases (36.6 vs 27.3%).

The distribution of the genotypic frequencies did show a significant difference between cases and controls ($\chi^2 = 5.1$, df = 1, p = 0.02). When compared with GG genotype (reference), the GC genotype of the variant *DKK3* rs7396187 conferred a protective effect as this genotype showed a lower frequency in patients (OR: 0.63; 95% CI: 0.44–0.91; p = 0.01).

Further, the variants of *DKK2* gene were also analyzed in order to investigate their role in modifying the predisposition toward lung cancer. The first polymorphic site which was

Table 3. Genotypic distribution of the DKK genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histology.

Genotype	Controls (n = 300), n (%)	Overall lung cancer			SQCC			ADCC			SCLC		
		Cases (n = 300), n (%)	AOR (95% CI) [†]	p-value [*]	Cases (n = 128), n (%)	AOR (95% CI) [†]	p-value [*]	Cases (n = 97), n (%)	AOR (95% CI) [†]	p-value [*]	Cases (n = 70), n (%)	AOR (95% CI) [†]	p-value [*]
DKK4, rs2073664 C/T													
CC	147 (49)	150 (50)	1.00 (Reference)		65 (50.8)	1.00 (Reference)		52 (53.6)	1.00 (Reference)		31 (44.3)	1.00 (Reference)	
CT	137 (45.7)	143 (47.7)	0.95 (0.68–1.34)	0.80	60 (46.9)	0.91 (0.59–1.41)	0.69	42 (43.3)	0.92 (0.57–1.49)	0.74	38 (54.3)	1.14 (0.66–1.98)	0.62
TT	16 (5.3)	7 (2.3)	0.37 (0.14–0.95)	0.03	3 (2.3)	0.34 (0.09–1.24)	0.10	3 (3.1)	0.55 (0.15–2.02)	0.37	1 (1.4)	0.14 (0.013–1.63)	0.11
CT + TT	143 (51)	150 (50)	0.89 (0.64–1.24)	0.50	63 (49.2)	0.84 (0.54–1.29)	0.44	45 (46.4)	0.88 (0.15–1.41)	0.60	39 (55.7)	1.02 (0.59–1.76)	0.93
C (allele)	431	443											
T (allele)	169	157											
MAF	0.28	0.26											
HWE	$\chi^2 = 3.7$, df = 2			0.15									
DKK3, rs2291599 C/T													
CC	195 (65)	195 (65)	1.00 (Reference)		79 (61.7)	1.00 (Reference)		63 (65)	1.00 (Reference)		50 (71.4)	1.00 (Reference)	
CT	89 (29.7)	81 (27)	0.88 (0.61–1.28)	0.53	39 (30.5)	1.18 (0.73–1.91)	0.47	28 (28.9)	0.81 (0.48–1.39)	0.46	13 (18.6)	0.61 (0.30–1.22)	0.16
TT	16 (5.3)	24 (8)	1.46 (0.74–2.88)	0.26	10 (7.8)	1.50 (0.64–3.54)	0.34	6 (6.1)	1.19 (0.44–3.25)	0.72	7 (10)	1.83 (0.65–5.19)	0.25
CT + TT	105 (35)	105 (35)	0.97 (0.68–1.37)	0.87	49 (38.3)	1.22 (0.78–1.91)	0.36	34 (35)	0.85 (0.52–1.41)	0.55	20 (28.6)	0.78 (0.43–1.43)	0.43
C (allele)	479	471											
T (allele)	121	129											
MAF	0.20	0.21											
HWE	$\chi^2 = 1.85$, df = 2			0.17									
DKK3, rs3206824 G/A													
GG	214 (71.3)	204 (68)	1.00 (Reference)		79 (61.7)	1.00 (Reference)		70 (72.2)	1.00 (Reference)		52 (74.3)	1.00 (Reference)	
GA	81 (27)	90 (30)	1.14 (0.79–1.64)	0.47	45 (35.2)	1.42 (0.89–2.26)	0.13	26 (26.8)	0.95 (0.56–1.62)	0.86	17 (24.3)	0.81 (0.43–1.52)	0.52
AA	5 (1.7)	6 (3)	1.29 (0.38–4.39)	0.68	4 (3.1)	2.04 (0.51–8.05)	0.30	1 (1)	–	–	1 (1.4)	–	–

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.[‡]Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

ADCC: Adenocarcinoma; AOR: Adjusted odds ratio; SCLC: Small-cell lung carcinoma; SQCC: Squamous cell carcinoma.

Table 3. Genotypic distribution of the DKK genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histology (cont.).

Genotype	Controls (n = 300), n (%)	Overall lung cancer				SQCC		ADCC		SCLC			
		Cases (n = 300), n (%)	AOR (95% CI) [†]	p-value [‡]	Cases (n = 128), n (%)	AOR (95% CI) [†]	p-value [‡]	Cases (n = 97), n (%)	AOR (95% CI) [†]	p-value [‡]	Cases (n = 70), n (%)	AOR (95% CI) [†]	p-value [‡]
DKK3, rs3206824 G/A (cont.)													
GA + AA	86 (28.7)	96 (32)	1.15 (0.80–1.64)	0.42	49 (38.3)	1.46 (0.93–2.29)	0.09	27 (27.8)	0.95 (0.56–1.60)	0.85	18 (25.7)	0.81 (0.44–1.50)	0.51
G (allele)	509	498											
A (allele)	91	102											
MAF	0.15	0.17											
HWE	χ ² = 0.72, df = 2		0.39										
DKK3, rs7396187 G/C													
GG	190 (63.3)	217 (72.3)	1.00 (Reference)		87 (68)	1.00 (Reference)		74 (76.3)	1.00 (Reference)		53 (75.7)	1.00 (Reference)	
GC	110 (36.7)	82 (27.4)	0.63 (0.44–0.91)	0.01	41 (32)	0.89 (0.56–1.42)	0.64	23 (23.7)	0.45 (0.26–0.78)	0.004	16 (22.9)	0.56 (0.30–1.07)	0.08
CC	–	1 (0.3)	–	–	–	–	–	–	–	–	1 (1.4)	–	–
GC + CC	110 (36.7)	83 (27.7)	0.63 (0.44–0.91)	0.01	41 (32)	–	–	23 (23.7)	–	–	17 (24.3)	0.60 (0.32–1.12)	0.11
G (allele)	490	516											
C (allele)	110	84											
MAF	0.18	0.14											
HWE	χ ² =15.11, df = 2		0.001										
DKK2, rs447372 A/G													
AA	143 (48)	150 (50)	1.00 (Reference)		74 (57.8)	1.00 (Reference)		45 (46.4)	1.00 (Reference)		29 (41.4)	1.00 (Reference)	
AG	140 (47)	123 (41)	0.85 (0.60–1.20)	0.37	43 (33)	0.60 (0.38–0.95)	0.03	40 (41.2)	0.90 (0.55–1.49)	0.70	37 (52.9)	1.50 (0.85–2.65)	0.15
GG	17 (5)	27 (9)	1.52 (0.78–2.95)	0.20	11 (9)	1.31 (0.56–3.05)	0.52	12 (12.3)	2.25 (0.99–5.09)	0.05	4 (5.7)	1.61 (0.47–5.51)	0.44
AG + GG	157 (52)	150 (50)	0.92 (0.67–1.28)	0.66	54 (42)	0.67 (0.43–1.04)	0.07	52 (53.5)	1.06 (0.66–1.07)	0.78	41 (58.6)	1.50 (0.86–2.61)	0.14
A (allele)	426	423											
G (allele)	174	177											
MAF	0.29	0.295											
HWE	χ ² = 5.32, df = 2		0.02										

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.
[‡]Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.
ADCC: Adenocarcinoma; AOR: Adjusted odds ratio; SCLC: Small-cell lung carcinoma; SQCC: Squamous cell carcinoma.

Table 3. Genotypic distribution of the <i>DKK</i> genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histology (cont.).										
Genotype	Controls (n = 300), n (%)	Overall lung cancer			SQCC		ADCC		SCLC	
		Cases (n = 300), n (%)	AOR (95% CI) [†]	p-value [‡]	Cases (n = 128), n (%)	AOR (95% CI) [†]	p-value [‡]	Cases (n = 97), n (%)	AOR (95% CI) [†]	p-value [‡]
DKK2, rs419558 C/T										
CC	238 (79.3)	203 (67.7)	1.00 (Reference)		86 (67.2)	1.00 (Reference)		67 (69.1)	1.00 (Reference)	
CT	62 (20.7)	97 (32.3)	1.77 (1.21–2.58)	0.002	42 (32.8)	1.71 (1.05–2.76)	0.02	30 (30.9)	1.87 (1.10–3.18)	0.02
C (allele)	538	503								
T (allele)	62	97								
MAF	0.10	0.16								
HWE	χ ² = 3.78, df = 1		0.05							
DKK2, rs17037102 G/A										
GG	245 (81.7)	215 (71.7)	1.00 (Reference)		85 (66.4)	1.00 (Reference)		71 (73.2)	1.00 (Reference)	
GA	55 (18.3)	82 (27.3)	1.62 (1.09–2.40)	0.01	41 (32)	2.13 (1.29–3.49)	0.0028	25 (25.8)	1.54 (0.87–2.69)	0.13
AA	–	3 (1)	–		2 (1.6)	–	–	1 (1)	–	–
GA + AA	55 (18.3)	85 (28.3)	1.69 (1.14–2.50)	0.008	43 (33.6)	2.28 (1.40–3.72)	0.0009	26 (26.8)	1.60 (0.92–2.79)	0.09
G (allele)	545	512								
A (allele)	55	88								
MAF	0.09	0.14								
HWE	χ ² = 3.01, df = 2		0.08							
†AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status. ‡Two-sided χ ² test for either genotype distribution or allelic frequencies between the cases and controls. ADCC: Adenocarcinoma; AOR: Adjusted odds ratio; SCLC: Small-cell lung carcinoma; SQCC: Squamous cell carcinoma										

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

^{††}Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

ADCC: Adenocarcinoma; AOR: Adjusted odds ratio; SCLC: Small-cell lung carcinoma; SQCC: Squamous cell carcinoma.

investigated was *DKK2* rs447372 (A/G). The genotypic frequencies of the control population showed a slight fluctuation from the HWE ($\chi^2 = 5.32$, $df = 1$, $p = 0.02$). The percentage of subjects having heterozygous AG genotype were over represented in controls in comparison to cases (47 vs 41% as shown in Table 3. However, the greater proportion of subjects were found to have mutant (GG) genotype in cases than in controls (9 vs 5%). The difference in the genotypic frequencies was not found to be significant ($\chi^2 = 3.5$, $df = 2$, $p = 0.17$). The MAF in both cases and controls was found to be 0.29. The odds of occurrence of lung cancer were calculated taking AA genotype as a reference and it was observed that the subjects carrying GG genotype were at moderate risk of developing lung cancer, however, it was not significant (OR: 1.52; 95% CI: 0.78–2.95; $p = 0.20$).

The allelic and genotypic frequencies of the second polymorphic site of *DKK2* (rs419558 C/T) were in accordance to HWE ($\chi^2 = 3.78$, $df = 1$, $p = 0.05$). The percentages of subjects with heterozygous CT genotype were overrepresented in cases as compared with controls (32.3 vs 20.6%). There was a significant difference in the genotypic frequencies between cases and controls ($\chi^2 = 9.8$, $df = 1$, $p = 0.001$). The MAF was found out to be 0.10 in controls and 0.16 in cases. The subjects with the heterozygous CT genotype were found to be at higher risk of developing lung cancer and found to be statistically significant (OR: 1.77; 95% CI: 1.21–2.58; $p = 0.002$) when compared with subjects carrying wild (CC) genotype. The third single nucleotide variant of *DKK2* gene analyzed was rs17037102 (G/A). The genotypic and allelic frequencies were in accordance with HWE ($\chi^2 = 3.01$, $df = 1$, $p = 0.08$). The proportion of subjects with heterozygous (GA) genotype were higher in cases as compared with controls (28.3 vs 18.3%). There was significant difference among the genotypic frequencies of cases and controls ($\chi^2 = 7.8$, $df = 1$, $p = 0.005$). The MAF was found to be 0.09 in controls and 0.14 in cases. The association analysis showed that the individuals with the combined variant (GA + AA) were at higher risk of developing lung cancer with OR: 1.69; 95% CI: 1.14–2.50; $p = 0.008$, which was found to be statistically significant. The heterozygous (GA) and mutant (AA) genotype were combined as single genotype since the percentage of subjects with mutant genotype was very low.

Distribution of genotypic combinations of the *DKK2* & *DKK3* gene & risk of lung cancer on basis of histological subtypes

In order to segregate the population into high-risk subgroups, the patients were divided into categories based upon their histological subtypes. The comparison of the genotypic frequency of the *DKK* variants was carried to find out the prevalence of the high-risk subgroups from all the categories as shown in Table 3. In case of *DKK4* rs2073664 (C/T), the genotypic frequencies for different histological types were found to be similar to the control group. Hence, none of the histological subgroup showed prominent effect on the risk of developing lung cancer.

For *DKK3* rs2291599 (C/T), an increased frequency of subjects with mutant (TT) genotype was observed in SQCC (7.8%) and SCLC (10%) patients as compared with controls (5.3%). When compared with the wild (CC) genotype a risk was found to be associated both in SQCC and SCLC but the risk was not significant, however a visible trend was observed. In case of *DKK3* rs3206824 (G/A) it was observed that the frequency of mutant (AA) genotype was present in higher frequency in ADCC patients as compared with controls but no significant association was seen. In the case of *DKK3* rs7396187 (G/C), the ADCC subjects carrying one susceptible allele (GC) were found to be less represented as compared with controls (23.7 vs 36%), therefore there was a decreased risk of developing lung cancer which was found to be significant (OR: 0.45, 95% CI: 0.26–0.78; $p = 0.0045$). The stratified analysis carried in case of *DKK2* rs447372 (A/G) variant, showed that ADCC patients carrying mutant (GG) genotype were at higher risk of developing lung cancer (OR: 2.25; 95% CI: 0.99–5.09; $p = 0.05$).

The other *DKK2* rs419558 (C/T) polymorphic site did confer a modulation in the predisposition toward lung cancer in all histological subtypes, where a similar increase in risk toward lung cancer was observed in all the three histological subgroups, where patients with heterozygous CT genotype were at 1.8-times higher risk of developing lung cancer (OR: 1.77; 95% CI: 1.05–2.76; $p = 0.02$). On the other hand, in case of *DKK2* rs17370102 (G/A), a significant increase in frequency of GA + AA genotype patients was observed in SQCC group (OR: 2.28; 95% CI: 1.40–3.72; $p = 0.009$). The

other two histological groups did not show a significantly increased frequency of patient with GA genotype, hence no significant association was seen in SCLC and ADCC group of patients.

Effects of smoking status on the association between *DKK* gene polymorphisms & lung cancer risk

In order to study gene–environment interaction, patients were subdivided into smokers and non-smokers based upon their smoking status. The genotypic frequencies are summarized in **Table 4** along with the associated odds ratio. *DKK4* rs2073664 (C/T) variant did not confer any contribution in increasing the lung cancer risk neither in smokers nor in nonsmokers. However, a decreased frequency of patients with mutant (TT) genotype was observed among smokers (OR: 0.32; 95% CI: 0.11–0.94; $p = 0.03$). In case of *DKK3* rs3206824, subjects who were nonsmokers and carried heterozygous (GA) genotype exhibited a twofold risk of developing lung cancer (OR: 1.91; 95% CI: 0.93–3.84; $p = 0.04$). However, no such association was seen in case of smokers. In case of *DKK3* rs7396187 (G/C), decreased frequency of patients with heterozygous (GC) genotype was observed in nonsmoker group which was found to be significant (OR: 0.44; 95% CI: 0.22–0.89; $p = 0.02$). For *DKK2* rs447372 (A/G), neither smokers nor nonsmokers showed any significant association with risk of lung cancer. In case of *DKK2* rs419558 (C/T) polymorphic site, patients who were smokers and had heterozygous (CT) genotype were found to have a twofold risk for developing lung cancer (OR: 1.94; 95% CI: 1.60–2.99; $p = 0.002$). However, no such significant risk was observed in nonsmoker category. On the contrary, the nonsmokers with variant (GA + AA) genotype of *DKK2* rs17370102 (G/A) were found at almost 2.7-times higher risk of developing lung cancer (OR: 2.69; 95% CI: 1.16–6.24; $p = 0.02$).

Association between lung cancer & different combinations of *DKK2* & *DKK3* genotypes

Further, in order to investigate gene–gene interaction the genotypic combinations of the various *DKK* variant genotypes were analyzed to find out whether there is a significant difference in frequency between cases and controls, the data are tabulated in **Table 5**. Out of all the double combinations, a twofold significant risk of lung cancer was observed in patients with

variant (GA + AA + GC) genotype of *DKK3* rs3206824 (G/A) and *DKK2* rs419558 (C/T; OR: 2.32; 95% CI: 1.23–4.36; $p = 0.008$) as there were no subjects reported with CC genotype of *DKK2* rs419558 (C/T), so they were not included in the analysis. In case of genotypic combinations of *DKK3* rs7396187 (G/C) and *DKK2* rs17370102 (G/A), the subjects with variant genotype were overrepresented in cases in respect to controls (15% vs 9%) and exhibited a marginal risk of acquiring lung cancer (OR: 1.82; 95% CI: 0.92–3.60; $p = 0.08$) but it was not significant. The two other genotypic combinations namely *DKK3* rs7396187 (G/C) and *DKK2* rs447372 (A/G) and *DKK3* rs7396187 (G/C) and *DKK4* rs2291599 (C/T) showed a significantly decreased in the frequency of variant genotype subjects in cases as compared with controls. The triple combination of *DKK* polymorphic sites were also analyzed in order to find out the combined effect of these polymorphic sites in modulating lung cancer susceptibility. For both the *DKK2* and *DKK3* polymorphic sites we combined both the heterozygous and mutant genotypes as a single group and compared it with the wild-type genotype. It was observed that the triple combination of variant *DKK3* gene (rs2291599 (C/T), rs3206824 (G/A) and rs7396187 (G/C)) did not confer any modulation in the lung cancer risk in subjects with variant genotype of all these polymorphic sites (OR: 1.08; 95% CI: 0.62–1.86; $p = 0.78$). In case of *DKK2* polymorphism (rs447372 (A/G), rs17037102 G/A, rs419558 C/T) it was observed that 12.6% of the patients carried the three susceptible alleles of all the three variants as compared with 3% of controls. When analyzed a 4.5-fold risk for lung cancer was observed in patients with combined variant genotype of all the three *DKK2* SNPs as compared with the wild-type genotype (OR: 4.62; 95% CI: 1.24–17.27; $p = 0.02$).

Relation between *DKK* genotype distribution & clinicopathologic characteristics

Table 6 shows the effect of *DKK* variant SNPs on the various clinic-pathological features such as stage, primary tumor extension, lymph node invasion and distant metastasis. The patients were stratified according to the clinical stage into stage III and stage IV patients. The subjects were categorized on the basis of tumor extension into TX + T1 + T2 and T3 + T4. On the basis of

Table 4. Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer.

Genotype	Smokers				Nonsmokers			
	Controls (n = 208), n (%)	Cases (n = 235), n (%)	AOR (95% CI) [†]	p-value [‡]	Controls (n = 92), n (%)	Cases (n = 65), n (%)	AOR (95% CI) [†]	p-value [‡]
DKK4, rs2073664 C/T								
CC	94 (45.2)	113 (48.1)	1.00 (Reference)		53 (57.1)	37 (57)	1.00 (Reference)	
CT	101 (48.6)	116 (49.4)	0.94 (0.63–1.39)	0.77	36 (39.6)	27 (41.5)	1.03 (0.52–2.03)	0.91
TT	13 (6.2)	6 (2.5)	0.32 (0.11–0.94)	0.03	3 (3.3)	1 (1.5)	0.53 (0.04–5.72)	0.60
CT + TT	114 (54.8)	122 (51.9)	0.87 (0.59–1.27)	0.48	39 (42.8)	28 (43)	0.99 (0.51–1.93)	0.99
DKK3, rs2291599 C/T								
CC	141 (67.8)	158 (67.2)	1.00 (Reference)		54 (59.3)	37 (56.9)	1.00 (Reference)	
CT	56 (26.9)	57 (24.3)	0.88 (0.56–1.37)	0.58	33 (36.2)	24 (36.9)	0.94 (0.46–1.94)	0.86
TT	11 (5.3)	20 (8.5)	1.59 (0.72–3.48)	0.24	5 (5.5)	4 (6.2)	1.00 (0.24–4.15)	0.99
CT+TT	67 (32.2)	77 (32.8)	1.00 (0.66–1.50)	0.99	38 (41.7)	28 (43.1)	0.91 (0.46–1.79)	0.80
DKK3, rs3206824 G/A								
GG	144 (69.2)	163 (69.4)	1.00 (Reference)		70 (76.9)	41 (63.1)	1.00 (Reference)	
GA	59 (28.4)	66 (28.1)	0.98 (0.64–1.50)	0.94	22 (24.1)	24 (36.9)	1.91 (0.93–3.89)	0.04
AA	5 (2.4)	6 (2.5)	1.35 (0.39–4.60)	0.63	–	–	–	
GA+AA	64 (30.8)	72 (30.6)	1.01 (0.66–1.52)	0.96	22 (24.1)	24 (36.9)	–	
DKK2, rs447372 A/G								
AA	99 (47.6)	121 (51.5)	1.00 (Reference)		44 (47.8)	29 (44.6)	1.00 (Reference)	
AG	99 (47.6)	92 (39.1)	0.79 (0.53–1.16)	0.26	41 (44.6)	31 (47.7)	1.01 (0.50–2.00)	0.97
GG	10 (4.8)	22 (9.4)	1.79 (0.80–3.99)	0.15	7 (7.6)	5 (7.7)	1.02 (0.28–3.68)	0.97
AG+GG	109 (52.4)	114 (48.5)	0.89 (0.69–1.30)	0.56	48 (52.2)	36 (55.4)	1.02 (0.53–1.98)	0.93
DKK3, rs7396187 G/C								
GG	141 (67.8)	170 (72.3)	1.00 (Reference)		49 (53.8)	47 (72.3)	1.00 (Reference)	
GC	67 (32.2)	64 (27.2)	0.76 (0.50–1.16)	0.21	43 (46.2)	18 (27.7)	0.44 (0.22–0.89)	0.02
CC	...	1 (0.5)			–	–	–	
GC+CC	67 (32.2)	65 (27.7)	0.77 (0.52–1.18)	0.23	–	–	–	
DKK2, rs17037102 G/A								
GG	165 (79.3)	169 (71.9)	1.00 (Reference)		80 (86.8)	46 (70.8)	1.00 (Reference)	
GA	43 (20.7)	65 (27.6)	1.43 (0.91–2.24)	0.11	12 (13.2)	17 (26.2)	2.32 (0.98–5.49)	0.05
GA+AA	43 (20.7)	66 (28.1)	1.45 (0.92–2.27)	0.10	12 (13.2)	19 (29.2)	2.69 (1.16–6.24)	0.02
DKK2, rs419558 C/T								
CC	161 (77.4)	151 (64.2)	1.00 (Reference)		77 (83.5)	50 (76.9)	1.00 (Reference)	
CT	47 (22.6)	82 (34.8)	1.94 (1.6–2.99)	0.0024	15 (16.5)	15 (23.1)	1.48 (0.65–3.38)	0.34

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.
[‡]Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.
AOR: Adjusted odds ratio.

lymph node invasion, patients were divided into N0 and N1+N2+N3 groups. Further to classify the patients on the distant metastatic status, they were stratified into M0 and M1 + M2 groups. Regarding clinical stage, tumor size, lymph node invasion and distant metastasis, no effect of SNP's was observed in patients as shown in Table 6. However, only a higher frequency of CT genotype of DKK3 rs2291599 (C/T) was seen in patients who were positive for lymph node invasion (OR: 2.36; 95% CI: 0.97–5.74; p = 0.05).

Linkage disequilibrium & haplotype analysis in DKK2 & DKK3 gene polymorphisms
Linkage disequilibrium and haplotype frequencies were calculated and tabulated as shown in Table 7 for all the three possible combination of SNPs of DKK2 gene and DKK3 gene. The haplotype with the frequencies less than 0.03 were not included in the analysis. DKK3 rs2291599 (C/T) and DKK3 rs3206824 (G/A) were in linkage disequilibrium ($D' = 0.36$, $r^2 = 0.096$, global $\chi^2 = 2.90$, $df = 3$, $p = 0.40$). However,

Table 5. Genotypic distribution based on different genotypic combinations and their association with lung cancer risk.

Genotype	Cases, n (%)	Controls, n (%)	AOR (95% CI) [†]	p-value [‡]
<i>DKK3</i> , rs2291599 C/T and <i>DKK3</i> , rs3206824 G/A:	n = 202	n = 191		
– 0	156 (77.2)	154 (80.6)	1.00 (Reference)	
– 1	44 (21.8)	35 (18.3)	1.24 (0.74–2.08)	0.40
– 2	2 (1)	2 (1.1)	0.94 (0.12–6.94)	0.95
– 3	46 (22.8)	37 (19.4)	1.22 (0.74–2.03)	0.42
<i>DKK3</i> , rs2291599 C/T and <i>DKK3</i> , rs7396187 G/C:	n = 221	n = 203		
– 0	172 (77.8)	148 (72.9)	1.00 (Reference)	
– 1	49 (22.2)	55 (27.1)	0.73(0.46–1.16)	0.18
<i>DKK3</i> , rs2291599 C/T and <i>DKK2</i> , rs447372 A/G:	n = 129	n = 124		
– 0	93 (72.1)	88 (71)	1.00 (Reference)	
– 1	35 (27.1)	36 (29)	0.92 (0.53–1.61)	0.78
– 2	1 (0.8)	–	–	
– 3	36 (27.9)	36 (29)	0.95 (0.54–1.65)	0.85
<i>DKK3</i> , rs2291599 C/T & <i>DKK2</i> , rs419558 C/T:	n = 155	n = 172		
– 0	130 (83.9)	154 (89.5)	1.00 (Reference)	
– 1	25 (16.1)	18 (10.5)	1.56 (0.80–3.02)	0.18
<i>DKK3</i> , rs2291599 C/T & <i>DKK2</i> , rs17037102 G/A:	n = 161	n = 181		
– 0	139 (86.3)	162 (89.5)	1.00 (Reference)	
– 1	22 (13.7)	19 (10.5)	1.31 (0.67–2.58)	0.42
<i>DKK3</i> , rs2291599 C/T and <i>DKK4</i> , rs2073664 C/T:	n = 150	n = 131		
– 0	103 (68.7)	94 (71.7)	1.00 (Reference)	
– 1	46 (30.7)	36 (27.5)	1.12 (0.65–1.19)	0.67
– 2	1(0.6)	1 (0.8)	0.67 (0.04–11.02)	0.77
– 3	47 (31.3)	37 (28.3)	1.10 (0.64–1.87)	0.78
<i>DKK3</i> , rs3206824 G/A and <i>DKK3</i> , rs7396187 G/C:	n = 216	n = 190		
– 0	170 (78.7)	147 (77.4)	1.00 (Reference)	
– 1	45 (20.8)	41 (21.6)	0.94 (0.57–1.54)	0.82
– 3	46 (21.3)	43 (22.6)	0.92 (0.57–1.50)	0.76
<i>DKK3</i> , rs3206824 G/A and <i>DKK2</i> , rs447372 A/G:	n = 141	n = 139		
0	101 (71.6)	101 (72.7)	1.00 (Reference)	
1	40 (28.4)	38 (27.3)	1.08 (0.64–1.84)	0.75
<i>DKK3</i> , rs3206824 G/A and <i>DKK2</i> , rs419558 C/T:	n = 179	n = 186		
– 0	143 (79.9)	169 (90.9)	1.00 (Reference)	
– 3	36 (20.1)	17 (9.1)	2.32 (1.23–4.36)	0.008
<i>DKK3</i> , rs3206824 G/A and <i>DKK2</i> , rs17037102 G/A:	n = 167	n = 154		
– 0	143 (79.9)	136 (88.3)	1.00 (Reference)	
– 3	24 (20.1)	18 (11.7)	1.28 (0.65–2.52) [†]	0.46
<i>DKK3</i> , rs3206824 G/A and <i>DKK4</i> , rs2073664 C/T:	n = 160	n = 97		
– 0	109 (68.1)	74 (76.3)	1.00 (Reference)	
– 1	50 (31.1)	23 (23.7)	1.40 (0.77–2.55)	0.26
– 3	51 (31.9)	23 (23.7)		
<i>DKK2</i> , rs447372 A/G and <i>DKK2</i> , rs419558 C/T:	n = 183	n = 175		
– 0	118 (64.5)	128 (73.1)	1.00 (Reference)	
– 1	50 (27.3)	40 (22.8)	1.32 (0.81–2.17)	0.25
– 3	65 (35.3)	47 (26.9)	1.46 (0.92–2.32)	0.10
<i>DKK2</i> , rs447372 A/G and <i>DKK2</i> , rs17037102 G/A:	n = 120	n = 132		
– 0	93 (77.5)	110 (83.3)	1.00 (Reference)	
– 1	22 (18.3)	19 (14.4)	1.36 (0.69–2.70)	0.36

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

[‡]Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

0: Wild genotype, 1: Heterozygote genotype, 2: Mutant genotype, 3: Combined hetero and mutant genotype.

AOR: Adjusted odds ratio.

Table 5. Genotypic distribution based on different genotypic combinations and their association with lung cancer risk (cont.).

Genotype	Cases, n (%)	Controls, n (%)	AOR (95% CI) [†]	p-value [‡]
<i>DDK2</i> , rs447372 A/G and <i>DDK2</i> , rs17037102 G/A (cont.): – 3	27 (22.5)	22 (16.7)	1.41 (0.75–2.67)	0.28
<i>DDK2</i> , rs447372 A/G and <i>DDK4</i> , rs2073664 C/T: – 0 – 1 – 3	n = 128 70 (54.7) 57 (44.5) 58 (45.3)	n = 132 68 (51.5) 63 (47.7) 64 (48.5)	1.00 (Reference) 0.84 (0.51–1.39) 0.84 (0.51–1.39)	0.51 0.50
<i>DDK2</i> , rs17037102 G/A and <i>DDK4</i> , rs2073664 C/T: – 0 – 3	n = 157 111 (70.7) 46 (29.3)	n = 150 121 (80.7) 29 (19.3)	1.00 (Reference) 1.52 (0.87–2.67)	0.13
<i>DDK3</i> , rs7396187 G/C and <i>DDK2</i> , rs447372 A/G: – 0 – 3	n = 141 104 (73.8) 37 (26.2)	n = 143 88 (61.5) 55 (38.5)	1.00 (Reference) 0.59 (0.35–0.99)	0.04
<i>DDK3</i> , rs7396187 G/C and <i>DDK2</i> , rs419558 C/T: – 0 – 3	n = 164 142 (86.6) 22 (13.4)	n = 182 155 (85.2) 27 (14.8)	1.00 (Reference) 0.81 (0.43–1.52)	0.51
<i>DDK3</i> , rs7396187 G/C and <i>DDK2</i> , rs17037102 G/A: – 0 – 3	n = 191 162 (84.8) 29 (15.2)	n = 165 150 (90.9) 15 (9.1)	1.00 (Reference) 1.82 (0.92–3.60)	0.08
<i>DDK3</i> , rs7396187 G/C and <i>DDK4</i> , rs2073664 C/T: – 0 – 3	n = 150 109 (72.7) 41 (27.3)	n = 151 94 (62.3) 57 (37.7)	1.00 (Reference) 0.58 (0.35–0.96)	0.03
<i>DDK3</i> , rs2291599 C/T, <i>DDK3</i> , rs3206824 G/A and <i>DDK3</i> , rs7396187 G/C: – 0 – 1 – 3	n = 182 142 (78) 31 (17) 40 (22)	n = 150 119 (79.3) 25 (16.7) 31 (20.7)	1.00 (Reference) 1.02 (0.56–1.88) 1.08 (0.62–1.86)	0.93 0.78
<i>DDK2</i> , rs447372 A/G, <i>DDK2</i> , rs17037102 G/A and <i>DDK2</i> , rs419558 C/T: – 0 – 1 – 3	n = 82 70 (85.4) 8 (9.76) 12 (14.6)	n = 99 96 (97) 2 (2) 3 (3)	1.00 (Reference) 4.35 (0.88–21.44) 4.62 (1.24–17.27)	0.07 0.02

[†]AORs (95% CI) and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

[‡]Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls

0: Wild genotype, 1: Heterozygote genotype, 2: Mutant genotype, 3: Combined hetero and mutant genotype.

AOR: Adjusted odds ratio.

none of the haplotypes showed any significant difference between the frequencies in cases and controls.

DDK3 rs2291599 (C/T) and *DDK3* rs7396187 (G/C) were found to be in linkage disequilibrium ($D' = 0.46$, $r^2 = 0.16$, the global $\chi^2 = 9.37$, $df = 3$, $p = 0.02$). The C-C haplotype of this combination was found to be under represented in cases as compared with controls (OR: 0.48; 95% CI: 0.30–0.78; $p = 0.02$). *DDK3* rs3206824 (G/A) and *DDK3* rs7396187 (G/C) gave a linkage disequilibrium ($D' = 0.29$, $r^2 = 0.08$, global $\chi^2 = 14.19$, $df = 3$, $p = 0.002$). The G-C haplotype was found to decrease the risk of lung cancer in patients (OR: 0.48; 95% CI: 0.22–0.72; $p = 0.002$).

The *DDK2* haplotypes were also analyzed to find out significant difference in frequency of various haplotype in cases and controls. *DDK2*

rs419558 (C/T) and *DDK2* rs17037102 (G/A) were found to be in linkage disequilibrium ($D' = 0.68$, $r^2 = 0.01$, global $\chi^2 = 14.8$, $df = 3$, $p = 0.0006$). Both haplotypes T-G (OR: 1.56; 95% CI: 1.10–2.21; $p = 0.01$) and C-A (OR: 1.58; 95% CI: 1.10–2.27; $p = 0.01$) showed a significant association with lung cancer risk. The next haplotype combination *DDK2* rs419558 (C/T) and rs4437372 (A/G) were also in linkage disequilibrium ($D' = 0.43$, $r^2 = 0.07$, global $\chi^2 = 10.43$, $df = 2$, $p = 0.01$). The T-A haplotype patients were found at twofold risk of developing lung cancer (OR: 2.10; 95% CI: 1.22–3.60; $p = 0.006$). The last haplotypic combination of *DDK2* rs17037102 (G/A) and *DDK2* rs447372 (A/G) gave a linkage disequilibrium of $D' = 0.53$, $r^2 = 0.016$, global $\chi^2 = 9.21$, $df = 2$, $p = 0.01$. The frequency of A-A haplotypes was more represented in cases as compared with controls (0.12% vs 0.07%). The

Genotype	Clinical stage	AOR (95% CI) [†]	p-value [‡]	Primary tumor extension		AOR (95% CI) [†]	p-value [‡]	Lymph node invasion		AOR (95% CI) [†]	p-value [‡]	Metastasis		AOR (95% CI) [†]	p-value [‡]
				Tx + T1 + T2	T3 + T4			N0	N1 + N2 + N3			No	Yes		
DKK4, rs2073664 C/T:															
- CC	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- CT	72 60	Reference		25	110	Reference		27	106	Reference		78	58	Reference	
- CT	65 61	1.18 (0.71–1.96)	0.51	37	91	0.53 (0.29–0.97)	0.04	19	110	1.38 (0.71–2.65)	0.33	72	55	1.11 (0.67–1.85)	0.67
- CT + TT	68 63	1.15 (0.70–1.91)	0.56	38	97	0.55 (0.30–0.99)	0.04	21	114	1.31 (0.69–2.49)	0.39	76	58	1.11 (0.67–1.84)	0.67
DKK3, rs2291599 C/T:															
- CC	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- CT	92 78	Reference		40	138	Reference		35	142	Reference		106	72	Reference	
- CT	33 38	1.32 (0.74–2.38)	0.33	19	51	0.82 (0.42–1.57)	0.55	7	62	2.36 (0.97–5.74)	0.05	33	37	1.67 (0.93–3.02)	0.08
- CT + TT	48 45	1.04 (0.61–1.76)	0.87	23	69	0.92 (0.50–1.69)	0.80	13	78	1.53 (0.75–3.10)	0.23	47	44	1.32 (0.77–2.24)	0.30
DKK3, rs3206824 G/A:															
- GG	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- GG	96 84	Reference		40	145	Reference		33	150	Reference		105	80	Reference	
- GA	38 39	1.15 (0.66–2.01)	0.60	22	57	0.73 (0.39–1.35)	0.31	14	65	1.04 (0.51–2.09)	0.91	43	36	1.08 (0.62–1.89)	0.76
- GA + AA	44 39	1.00 (0.58–1.72)	0.99	23	62	0.79 (0.43–1.45)	0.46	15	70	1.03 (0.52–2.04)	0.92	49	36	0.96 (0.55–1.65)	0.88
DKK3, rs7396187 G/C:															
- GG	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- GG	101 90	Reference		45	154	Reference		37	161	Reference		114	85	Reference	
- GC	38 33	1.00 (0.57–1.76)	0.98	18	52	0.82 (0.43–1.55)	0.54	11	58	1.21 (0.57–2.54)	0.60	39	31	1.13 (0.64–2.00)	0.66
- GC + CC	39 33	0.98 (0.56–1.72)	0.95	18	53	0.83 (0.44–1.58)	0.58	11	59	1.23 (0.58–2.58)	0.57	39	31	1.10 (0.62–1.95)	0.72
DKK2, rs447372 A/G:															
- AA	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- AA	66 64	Reference		29	104	Reference		20	111	Reference		73	60	Reference	
- AG	62 47	0.69 (0.40–1.18)	0.17	27	85	0.89 (0.48–1.65)	0.72	24	88	0.68 (0.35–1.33)	0.26	67	45	0.76 (0.44–1.30)	0.32
- AG + GG	74 59	0.73 (0.44–1.22)	0.23	34	103	0.86 (0.48–1.54)	0.63	28	109	0.72 (0.38–1.36)	0.31	80	56	0.79 (0.47–1.31)	0.36
DKK2, rs419558 C/T:															
- CC	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- CC	94 83	Reference		41	139	Reference		33	145	Reference		102	78	Reference	
- CT	46 40	1.08 (0.63–1.85)	0.75	22	68	0.84 (0.46–1.54)	0.58	15	75	1.07 (0.54–2.12)	0.83	52	38	1.06 (0.62–1.80)	0.82
DKK2, rs17037102 G/A:															
- GG	140 123	Reference		63	207	Reference		48	220	Reference		154	116	Reference	
- GG	96 86	Reference		45	146	Reference		33	156	Reference		109	83	Reference	
- GA	44 35	0.86 (0.49–1.49)	0.60	17	59	1.06 (0.56–2.03)	0.83	15	61	0.85 (0.43–1.70)	0.66	43	32	0.95 (0.54–1.67)	0.87
- GA + AA	44 37	0.90 (0.52–1.55)	0.71	18	61	1.04 (0.55–1.96)	0.90	15	64	0.90 (0.46–1.79)	0.78,	44	33	0.91 (0.52–1.59)	0.75

[†]AORs, 95% CIs and their corresponding [‡]p-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking.
[‡]ORs: Adjusted odds ratio.

[†]AORs, 95% CIs and their corresponding [‡]p-values were calculated by unconditional logistic analysis after adjusting for age, gender and smoking.
AOR: Adjusted odds ratio.

Table 7. Haplotype frequencies and lung cancer risk associated with DKK2 and DKK3 haplotypes.				
Haplotype	Case frequency (%)	Control frequency (%)	OR (95% CI) [†]	p-value [‡]
Haplotype frequencies & lung cancer risk associated with the DKK3 haplotypes				
<i>DKK3, rs2291599 C/T and DKK3, rs3206824 G/A:</i>				
– C-G	424.3 (0.70)	428.3 (0.71)	0.96 (0.75–1.24)	0.80
– C-A	46.6 (0.07)	50.71 (0.08)	0.91 (0.60–1.38)	0.66
– G-T	73.61 (0.12)	80.71 (0.13)	0.90 (0.64–1.26)	0.53
– A-T	55.4 (0.09)	40.3 (0.06)	1.41 (0.92–2.15)	0.10
<i>DKK3, rs2291599 C/T and DKK3, rs7396187 G/C:</i>				
– C-G	443.3 (0.73)	424.8 (0.70)	1.16 (0.90–1.50)	0.23
– C-C	27.6 (0.04)	54.1 (0.09)	0.48 (0.30–0.78)	0.002
– G-T	72.6 (0.12)	65.1 (0.10)	1.13 (0.79–1.61)	0.49
– T-C	56.3 (0.09)	55.8 (0.09)	1.00 (0.68–1.48)	0.96
<i>DKK3, rs3206824 G/A and DKK3, rs7396187 G/C:</i>				
– G-G	453.5 (0.70)	432.4 (0.72)	1.25 (0.96–1.02)	0.08
– G-C	39.5 (0.06)	76.61 (0.12)	0.48 (0.22–0.72)	0.002
– A-G	57.5 (0.09)	57.6 (0.04)	0.99 (0.68–1.46)	0.99
– A-C	44.4 (0.07)	33.4 (0.05)	1.35 (0.85–2.16)	0.19
Haplotype frequencies & lung cancer risk associated with the DKK2 haplotypes				
<i>DKK2, rs419558 C/T and DKK2, rs17037102 G/A:</i>				
C-G	421.27 (0.70)	483.03 (0.80)	0.59 (0.45–0.77)	0.0001
T-G	90.73 (0.15)	4.98 (0.10)	1.56 (1.10–2.21)	0.01
C-A	81.73 (0.136)	54.98 (0.09)	1.58 (1.10–2.27)	0.01
T-A	6.27 (0.01)	0.03 (0.00)	–	...
<i>DKK2, rs419558 C/T and DKK2, rs447372 A/G:</i>				
C-A	381.29 (0.63)	405.39 (0.67)	0.83 (0.65–1.06)	0.14
C-G	121.71 (0.20)	132.61 (0.22)	0.89 (0.68–1.18)	0.44
T-A	41.71 (0.07)	20.61 (0.034)	2.10 (1.22–3.60)	0.006
T-G	55.29 (0.09)	41.39 (0.06)	1.37 (0.90–2.08)	0.14
<i>DKK2, rs17037102 G/A and DKK2, rs447372 A/G:</i>				
G-A	346.10 (0.57)	379.79 (0.63)	0.79 (0.62–1.00)	0.05
G-G	165.9 (0.27)	165.21 (0.27)	1.01 (0.78–1.30)	0.93
A-A	76.9 (0.12)	46.2 (0.07)	1.77 (1.20–2.60)	0.003
A-G	11.10 (0.01)	8.79 (0.01)	–	–
Bold values refer to statistically significant values.				
[†] ORs, 95% CIs and their corresponding p-values were calculated by logistic regression analysis.				
[‡] Two- sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.				
OR: Odds ratio.				

odds of developing lung cancer was found to be increased as compared with the GA haplotype (OR: 1.77; 95% CI: 1.20–2.60; p = 0.003).

Discussion

Pathway based approach is increasingly becoming a critical way to understand the process of disease initiation and progression as it is evident that it involves multitude of signaling networks and targeting a single marker will not enable us to predict the prognosis and treatment efficacy with high accuracy [28]. Targeting the highly complex intermeshing pathways such as Wnt signaling cascade can help us to understand the pathogenesis of lung cancer that lung cancer

is a highly heterogeneous and complex genetic disorder. Abnormalities in the Wnt signaling cascade have been recently linked to the process of oncogenesis in various human cancers [3]. Wnt signaling cascade is regulated by the secreted glycoproteins DKKs (Dickkopf) which antagonize the signal cascade by binding the LRP co-receptor, in doing so it disrupts the interactive complex of Wnt ligand/ Frizzled receptor/LRP. These interactions may affect the regulation of the Wnt cascade in lung tumorigenesis [29]. The present study is focusing on the various genetic variations in the DKK genes, which in turn can alter the interaction between the DKK protein and LRP5/6 thus

dysregulating the Wnt signaling and playing a role in modulating the lung cancer susceptibility. To our knowledge this is the first study to utilize such an approach to link SNPs in Wnt pathway genes with predisposition to lung cancer in north Indian population.

When effect of *DKK4* (rs2073664) polymorphisms on lung cancer were investigated in this study, we detected that *DKK4* (rs2073664) polymorphism was not associated with lung cancer development and thus shows a protective effect. Our data are consistent with a study done in lung cancer of Turkish patients [22] and renal cancer patients of North America [19]. Similarly, another polymorphic site of the *DKK4* gene (rs3763511) was analyzed in breast cancer females of Saudi Arabian by Alanzi *et al.* [21]. They too reported that this SNP in the *DKK4* gene did not show any association in overall breast cancer.

The investigations in the present study showed there was no significant difference in the frequency of the subjects with variant genotype for both the *DKK3* genes including rs3206824 (Arg³³⁵Gly) and *DKK3* rs2291599 (Intronic variation). However, for the *DKK3* rs7396187 (G/C) genotype a protective effect for lung cancer was observed (OR: 0.63; $p = 0.01$) and this protective effect was more pronounced in ADCC subjects ($p = 0.004$). Our results were concordant with the study carried by Kobayashi *et al.* [30], where they screened lung cancer patients for codon 335 mutation in the *DKK3* gene rs3206824 (Arg³³⁵Gly) and found no significant difference between cases and controls. On the contrary, a case-control study carried out in Turkish lung cancer and renal carcinoma patients have reported that the *DKK3* rs3206824 (Arg³³⁵Gly) was found to confer a protective effect [19,22]. However, these two studies did not find any association for the other *DKK3* variants. Investigations carried in breast cancer females to find out the association between the SNP of *DKK3* gene (rs6485350) and *DKK4* (rs3763511) gene showed a decrease in frequency of the patients with variant genotype for both the polymorphic sites [21].

Screening of colon cancer patients for *DKK3* rs3206824 and *DKK2* rs17037102 variants did not show any positive correlation [23]. Some other reports showed association of various *DKK3* polymorphisms with prostate cancer and factors related to it [20]. The functional role of *DKK3* protein as a tumor suppressor has been greatly supported by various reports [31,32]. Unlike, other members of the DKK family *DKK3* do not have

the affinity for the Wnt receptor complex, however its function yet remains unclear [33] but recently it has been seen that *DKK3* is a negative regulator of β -catenin [34]. Moreover, lung cancer was the first primary tumor where the dysregulated expression pattern of *DKK3* was asserted [16]. Recent evidences have shown the role of *DKK3* protein as a key factor in lung cancer pathogenesis [12]. This can be clearly understood by a study which states that *DKK3* protein is antagonizing β -catenin driven Wnt pathway in lung cancer cells, which in turn leads to the lowering of expression of Wnt target genes such as c-myc and cyclin D1 specifically in non-small-cell lung cancer cell lines [12]. Also, the variation in levels of *DKK3* m-RNA have been confirmed in case of human renal carcinoma cells [32]. Another interesting finding of the present study included the association of various *DKK3* haplotypes with lung cancer, the G-C haplotype of *DKK3*, rs3206824 and *DKK3*, rs7396187 showed a significantly decreased frequency in cases when compared with controls ($p = 0.0002$) thus this haplotype block confers a decreased risk toward lung cancer. However a similar study conducted in Turkish lung cancer patients did not show any significant difference in the frequency among the cases and controls for this haplotype [22]. Another study in renal cancer patients reported that the A-C instead of G-C haplotype was significantly lower in cases [19].

Coming to the *DKK2* genetic variants, our study showed that individuals with one susceptible allele for either Arg¹⁴⁶Glu variant (OR: 1.69; $p = 0.008$) or rs419558 (3'-UTR region; OR: 1.77; $p = 0.002$) did confer a positive association toward lung cancer. It is evident that genotypic frequencies of the control sample set in case of rs447372 variant were inconsistent with HWE and also the frequencies for rs419558 are on the line of deviating HWE, this flaw might have some effect on the important findings of *DKK2*. However, these deviations are attributable to the fact that HWE which is based on various assumptions such as random mating, genetic drift among others, which might be violated in this sample set taken from the entire population. The reasons include population stratification, sample size and nonrandom mating. So, the trends observed during this study for *DKK2* are evidently showing its implication in the process of lung carcinogenesis. Variations in the 3'-UTR region have important implication in transcriptional regulation as it is

involved in the stability of the mRNA and its translation [21]. *DKK2* Arg¹⁴⁶Glu was also evaluated in Turkish lung cancer patients to find out whether it confers any effect on lung cancer susceptibility; however, no significant effect was observed [22]. This discrepancy might insight us toward the difference in the population due to their ethnicities and other factors [21] and also the sample size was very less in that study. A study conducted in renal carcinoma patients found that none of the polymorphic sites of *DKK2* gene which was studied there had any effect toward renal cancer. However, the *DKK2* Arg¹⁴⁶Glu variant was found to have a positive association with the survival of renal cancer patients [19]. Similarly their haplotype analysis for *DKK2* variants did not show any significant difference in the frequencies among cases and controls [19].

But in the present study the haplotypes for three variants of *DKK2* gene were analyzed. In case of *DKK2* rs419558 and *DKK2* rs447372 the TA haplotype was overrepresented in cases than controls, hence a twofold increased risk for lung cancer was observed ($p = 0.006$). Also the AA haplotype block of *DKK2* rs447372 and *DKK2* rs17037102 showed a higher risk of acquiring lung cancer among cases ($p = 0.003$). The two haplotypes (CA and TG) of *DKK2* rs419558 and rs17037102 showed almost 1.5-fold risk of developing lung cancer, whereas decreased risk for lung cancer was observed in cases having CG haplotype (OR: 0.59; $p = 0.0001$).

Gene–gene interaction plays a significant role in altering the susceptibility of carcinogenesis. The gene–gene interactions were analyzed for the *DKK* gene variants. Out of the 21 possible combinations among all the *DKK3* variants, our study reveals for the first time that combined variant genotypes of *DKK3* rs3206824 and *DKK2* rs419558 showed a twofold (OR: 2.32; $p = 0.008$) increased risk in lung cancer patients. Similar gene–gene analyses of the *DKK* variants carried out by Hirata *et al.* showed no significant correlation among the different genotypic combination; however, they did report a strong association toward decreased risk of renal carcinoma in patients with variant genotype of *DKK3* rs3206824 and *sFRP4* rs1802074 [19]. The triple combination of three polymorphic sites of *DKK2* showed an increased frequency among the patients with variant genotype of these sites as compared with controls. This is probably the first report which is giving us the insight about the lung cancer patients having the variant genotype of all the three investigated

DKK2 SNPs. A very few reports are available on these variants of *DKK2* to compare our findings with. *DKK2* protein has been categorized as a one of the important antagonist of Wnt pathway and binds to LRP/Wnt receptor complex [35], also recent studies have reported about the aberrant activation of Wnt pathway which might play an imminent role in the process of lung tumorigenesis [3]. Therefore, the involvement of *DKK2* is inevitable in etiology of this disease. In addition, our data were in accordance to that of Hirata *et al.* [19] and Yilmaz *et al.* [22], where both reported that polymorphisms within the *DKK* gene were not related to clinic-pathological features. Since Wnt signaling is a highly diverse and complex pathway with different proteins involved and many functional dimensions are still unexplored, so it may not be possible to correlate clinic-pathological factors of the disease with known functions of the Wnt pathway [36]. Future epidemiological and functional studies are warranted to confirm these findings.

Conclusion

In conclusion, this is the first study to highlight the genetic variations of *DKK3*, *DKK2* and *DKK4* gene and their effect on lung cancer risk in north Indians with a larger sample size and also their association with clinic-pathological factors. An attempt was made to find out the high-risk subgroups, by stratifying the subjects based on smoking status also. However, we also did not find any effect of smoking and polymorphism within the *DKK* gene. Patients with variant genotype of all the three variants of *DKK2* gene were found to be positively correlated with lung cancer predisposition. None of the SNPs studied in the *DKK* gene had any effect on clinic-pathological parameters, including tumor size, stage, lymph node status, or metastasis among both non-small-cell-lung cancer and SCLC groups. These certain interesting findings along with studies in gene expression with larger sample size and stratified based onset of the disease might help in establishing the influence of genetic variants in the Wnt signaling pathway genes. Certain limitations of the study include the significant difference in the mean age of cases and controls; however, a sincere attempt was made to match the age of cases and controls with a (± 10 years) difference, but still the due to various epidemiological reasons and also being a hospital-based case–control study exact matching was not achieved. So, in order to combat that repercussions of age bias which could clearly act as

a confounding factor in determining the effect of the genetic variability in *DKK* gene on lung cancer predisposition we used a broad multivariate logistic regression approach to carry out all the statistical comparisons. Therefore, it can be clearly said that the results were already controlled by adjusting the OR with age as one of the confounding factor. The other limitation is the small sample size in the subgroups made on the basis of histology, which is a bit of a constraint but then this will give an insight about whether future studies when conducted or carried out can also be toward this direction. Similar studies done would validate or negate our findings. The stratified analysis done with larger sample size can aid in distinguishing the high-risk subgroups from population to help in the early diagnosis of the disease. There have been very few studies which have actually evaluated the role of these polymorphisms within the *DKK* gene in modulating lung cancer susceptibility, so to first establish its implication in the process of lung carcinogenesis it is important to study these polymorphisms for lung cancer.

Future perspective

The molecular abnormalities involved in the process of lung carcinogenesis are highly complex and involves the intermeshing cell signaling and proliferation pathways such as Wnt pathway. Dickkopf (*DKK*) family of proteins are involved in the regulation of the Wnt signaling cascade. However, the genetic variabilities within the genes coding this extracellular glycoprotein may confer a role in altering its functionalities, hence

confer a hand in the aberrant activation of Wnt pathway which has been recently known to play a significant role in process of cancer initiation and progression. The study of these single nucleotide variations present in the *DKK* protein can greatly help in looking for novel therapeutic modalities, as they can serve as a target. Evidences have been increasing nowadays, which show that *DKK* protein can be a major player in tumorigenesis. Further studies can also help in establishing the *DKK* as a biomarker for certain particular subgroups. These kind of studies help in giving us the insight about lung carcinogenesis and these tumor suppressor proteins are strong candidature for a prognostic marker also.

Financial & competing interests disclosure

This work was supported by grant from the Indian Council of Medical Research, New Delhi, India (Grant No. 5/13/126/2011/NCD-III). The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

The authors state that they have obtained appropriate institutional review board approval or have followed the principles outlined in the Declaration of Helsinki for all human or animal experimental investigations. In addition, for investigations involving human subjects, informed consent has been obtained from the participants involved.

EXECUTIVE SUMMARY

- The present study investigates the role of the various SNP in *DKK* protein in modifying the lung cancer susceptibility in north Indians.
- *DKK4* (rs2073664) did not confer any effect on lung cancer susceptibility in subjects having the variant allele for this SNP.
- Two variants of *DKK3* namely rs3206824 and rs2291599 didn't show any association except rs7396187 which showed a protective effect (odds ratio: 0.63; $p = 0.01$) in patients carrying the variant allele for this polymorphic site.
- The subjects with heterozygous genotype of *DKK2* rs17037102 and rs419558 were found to have a higher risk of developing lung cancer.
- The variant genotype of the genotypic combination, *DKK3* rs3206824 and *DKK2* rs419558 showed a twofold increased risk of developing lung cancer ($p = 0.008$).
- Lung cancer risk increased four-times in subjects with variant genotype of all the three *DKK2* variants (rs447372, rs419558 and rs17037102). None of the polymorphic site was found to be associated with the clinic-pathological features.
- Limitations of the present study is a smaller sample size, the screening of larger sample size in different populations can pave a way for *DKK* protein to be established as a biomarker.

References

Papers of special note have been highlighted as:

• of interest; •• of considerable interest

- 1 Winn RA, Bremnes RM, Bemis L *et al.* γ -Catenin expression is reduced or absent in a subset of human lung cancers and re-expression inhibits transformed cell growth. *Oncogene* 21, 7497–7506 (2002).
- 2 Yokota JI, Shiraishi K, Kohno T. Genetic basis for susceptibility to lung cancer: recent progress and future directions. *Adv. Cancer Res.* 109, 51–72 (2010).
- 3 Mazieres J, He B, You L, Xu Z, Jablons DM. Wnt signaling in lung cancer. *Cancer Lett.* 222, 1–10 (2005).
- Discusses the role of Wnt pathway in the process of lung carcinogenesis; gives an insight about the importance of Wnt pathway in lung cancer.
- 4 Yang Y. Wnt signaling in development and disease. *Cell Biosci.* 2, 1 (2012).
- Discusses the role of Wnt signalling in various metabolic pathways and cellular processes.
- 5 Akiyama T. Wnt/ β -catenin signaling. *Cytokine Growth Factor Rev.* 11, 273–282 (2000).
- 6 Giles RH, van Es JH, Clevers H. Caught up in a Wnt storm: Wnt signaling in cancer. *Biochim. Biophys. Acta* 1653, 1–24 (2003).
- 7 Polakis P. Wnt signalling and cancer. *Genes Dev.* 14, 1837–1851 (2000).
- 8 Luo J, Chen J, Deng Z-L *et al.* Wnt signaling and human diseases: what are the therapeutic implications? *Lab. Invest.* 87, 97–103 (2007).
- 9 Zorn AM. Wnt signalling: antagonistic Dickkopfs. *Curr Biol.* 11, 592–595 (2001).
- Explains the role of DKK proteins in controlling the constitutive activation of Wnt pathway.
- 10 Krupnik VE, Sharp JD, Jiang C *et al.* Functional and structural diversity of the human *Dickkopf* gene family. *Gene* 238, 301–313 (1999).
- 11 Niehrs C. Function and biological roles of the *Dickkopf* family of Wnt modulators. *Oncogene* 25, 7469–7481 (2006).
- 12 Veeck J, Dahl E. Targeting the Wnt pathway in cancer: the emerging role of Dickkopf-3. *Biochim. Biophys. Acta* 18–28 (2012) (1825).
- Gives us an insight about the role of DKK3 protein in the process of carcinogenesis.
- 13 Yue W, Sun Q, Dacic S *et al.* Downregulation of Dkk3 activates β -catenin/TCF-4 signaling in lung cancer. *Carcinogenesis* 29, 84–92 (2008).
- 14 Sato H, Suzuki H, Toyota M *et al.* Frequent epigenetic inactivation of DICKKOPF family genes in human gastrointestinal tumors. *Carcinogenesis* 28, 2459–2466 (2007).
- 15 Kuphal S, Lodermeier S, Bataille F, Schuierer M, Hoang BH, Bosserhoff AK. Expression of *Dickkopf* genes is strongly reduced in malignant melanoma. *Oncogene* 25, 5027–5036 (2006).
- 16 Nozaki I, Tsuji T, Iijima O, Ohmura Y, Andou A, Miyazaki M, Shimizu N, Namba M. Reduced expression of REIC/Dkk-3 gene in non-small cell lung cancer. *Int. J. Oncol.* 19, 117–121 (2001).
- Evaluated the role of DKK3 protein in non-small-cell lung cancer tumors.
- 17 Hirata H, Hinoda Y, Nakajima K *et al.* Wnt antagonist gene *DKK2* is epigenetically silenced and inhibits renal cancer progression through apoptotic and cell cycle pathways. *Clin. Cancer Res.* 15, 5678–5688 (2009).
- 18 Shen LX, Basilion JP, Stanton VP. Single-nucleotide polymorphisms can cause different structural folds of mRNA. *Proc. Natl Acad. Sci. USA* 96, 7871–7876 (1999).
- 19 Hirata H, Hinoda Y, Nakajima K *et al.* Wnt antagonist gene polymorphisms and renal cancer. *Cancer* 115, 4488–4503 (2009).
- Examined the role of all the SNPs studied in the present study in renal cancer; gives the necessary information about the SNPs under study.
- 20 Kim MS, Lee HN, Kim HJ, Myung SC. Single nucleotide polymorphisms in *DKK3* gene are associated with prostate cancer risk and progression. *Int. Braz. J. Urol.* 41, 869–897 (2015).
- 21 Alanazi MS, Parine NR, Shaik JP, Alabdulkarim H a., Ajaj SA, Khan Z. Association of single nucleotide polymorphisms in Wnt signaling pathway genes with breast cancer in Saudi patients. *PLoS ONE* 8, e59555 (2013).
- Gives the data about the effect of the polymorphisms in Wnt antagonists gene in breast cancer risk.
- 22 Yilmaz M, Donmez G, Kacan T *et al.* Significant association between polymorphisms of Wnt antagonist genes and lung cancer. *J. Investig. Med.* 63, 935–942 (2015).
- Recent study done on Turkish population where the various SNPs of the Wnt antagonist have been investigated for the first time in lung cancer patients.
- 23 Páez D, Gerger A, Zhang W *et al.* Association of common gene variants in the WNT/ β -catenin pathway with colon cancer recurrence. *Pharmacogenomics J.* 14, 142–150 (2014).
- 24 Sodhi KK, Bahl C, Singh N, Behera D, Sharma S. Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in north Indians. *Future Oncol.* 11, 2159–2173 (2015).
- 25 Shi YY, He L. SHEsis, a powerful software platform for analyses of linkage disequilibrium, haplotype construction, and genetic association at polymorphism loci. *Cell Res.* 15, 97–98 (2005).
- 26 SHEsis. <http://analysis.bio-x.cn/myAnalysis.php>
- 27 MedCalc®. www.medcalc.org
- 28 Szkandera J, Pichler M, Absenger G *et al.* A functional germline variant in GLI1 implicates hedgehog signaling in clinical outcome of stage II and III colon carcinoma patients. *Clin. Cancer Res.* 20, 1687–1697 (2014).
- 29 Kawano Y, Kypta R. Secreted antagonists of the Wnt signalling pathway. *J. Cell Sci.* 116, 2627–2634 (2003).
- 30 Kobayashi K, Ouchida M, Tsuji T *et al.* Reduced expression of the REIC/Dkk-3 gene by promoter-hypermethylation in human tumor cells. *Gene* 282, 151–158 (2002).
- 31 Hsieh SY, Hsieh PS, Chiu CT, Chen WY. Dickkopf-3/REIC functions as a suppressor gene of tumor growth. *Oncogene* 23, 9183–9189 (2004).
- 32 Kurose K, Sakaguchi M, Nasu Y *et al.* Decreased expression of REIC/Dkk-3 in human renal clear cell carcinoma. *J. Urol.* 171, 1314–1318 (2004).
- 33 Brott B.K., Sokol S.Y. Regulation of Wnt/LRP signaling by distinct domains of Dickkopf proteins. *Mol. Cell Biol.* 22, 6100–6110 (2002).
- 34 Lee EJ, Jo M, Rho SB *et al.* Dkk3, downregulated in cervical cancer, functions as a negative regulator of β -catenin. *Int. J. Cancer* 124, 287–297 (2009).
- 35 Mao B, Niehrs C. Kremen2 modulates Dickkopf 2 activity during Wnt/LRP6 signaling. *Gene* 302, 179–183 (2003).
- Tells about the DKK2 protein and its implication in Wnt signaling.
- 36 Beachy PA, Karhadkar SS, Berman DM. Tissue repair and stem cell renewal in carcinogenesis. *Nature* 432, 324–331 (2004).



Single nucleotide variations in the Wnt antagonist *sFRP3* (rs7775 & rs288326) and *sFRP4* (rs1802073 & 1802074) genes and their association with lung cancer risk in North Indians

Charu Bahl^a, Siddharth Sharma^{a,*}, Navneet Singh^b, Digamber Behera^b

^a Department of Biotechnology, Thapar University, Patiala, Punjab 147002, India

^b Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER), Sector 14, Chandigarh, India

ARTICLE INFO

Keywords:

Lung cancer
sFRP
Wnt pathway
SNP

ABSTRACT

Purpose: To evaluate the role of genetic variants or single nucleotide polymorphic variants (rs288326, rs7775, rs1802073 and rs1802074) in Secreted frizzled related protein in modulating lung cancer susceptibility.

Materials and methods: A total of 610 subjects were genotyped using PCR-RFLP technique for each polymorphic site including 300 cases and 310 controls. Further association analysis was carried out using logistic regression approach to obtain adjusted odds ratio and statistical significance.

Results: The two genetic variants of *sFRP3* (rs288326 and rs7775) didn't show any significant role in modulating the susceptibility towards lung cancer whereas the two *sFRP4* variants (rs1802073 and rs1802074) did show a significant impact and were found to be significantly associated with risk of developing lung cancer. Individuals with the mutant genotype for *sFRP4* Pro³²⁰Thr variant were found to have a three times higher risk of developing lung cancer (OR = 3.21, 95% CI: 1.88–5.4 $p < 0.0001$). Adenocarcinoma patients were found to be more susceptible as compared to other histological subtypes, in case of all the four genetic variants.

Conclusion: Hence, the four genetic variants of the Secreted frizzled related protein did confer an impact on the susceptibility of North Indians towards lung cancer.

1. Introduction

Lung cancer is a complex heterogeneous molecular disease which arises due to multiple alterations either at genetic or epigenetic level and these alterations either target activation of growth promoting or inhibit tumor suppressing pathways. Thus, proper understanding of multiple signalling pathways which play an important role in the molecular pathogenesis of lung cancer is highly imperative for the development of diagnostic/prognostic markers and also for treatment strategies that can target molecular aberrations and their downstream activated pathways (Surana et al., 2013). It has been established that tobacco intake is the prime initiator of lung cancer in 80% of the total cases (Wogan et al., 2004) but recently evidences from familial clustering and segregation studies have revealed that genetic susceptibility might confer an important role in occurrence of lung cancer. A distinct amalgam of polymorphic traits in each individual affects his/her

susceptibility towards lung cancer (Yokota et al., 2010). Therefore, it is important to understand the role of various cancer related genes including the dominant oncogenes, recessive tumor suppressor genes and their respective polymorphic forms in the process of lung carcinogenesis, so as to find out the populations who would be at a higher risk of acquiring the disease. Additionally, this understanding on the same shall prove to be a tool for the identification of both predictive and prognostic markers for lung carcinogenesis (Amos et al., 1999).

Among the most extensively studied, one of the less explored tumor suppressor gene is Secreted frizzled related protein (*sFRP*), which regulates an important cell differentiation and proliferation pathway, the Wnt (wingless) pathway. Categorically, the Wnt signalling cascade has been classified into β -catenin dependent (canonical) and β -catenin independent (non -canonical) pathways. Wnt pathway is known to play several crucial roles in the process of embryonic development, tissue stability, and tumor progression (Yang, 2012). Recent literature

Abbreviations: LRP, LDL-receptor related protein; GSK-3 β , glycogen synthase kinase-3 β ; Dvl, dishevelled; TCR, T-cell receptor; LEF, lymphoid enhancing factor; sFRP, Secreted frizzled related protein; VEGF-A, vascular endothelial growth factor-A; MMP7, matrix metalloprotein 7; CRD, cysteine rich domain; Wnt, wingless; NSCLC, non-small cell lung cancer; SQCC, squamous cell carcinoma; ADCC, adenocarcinoma; SCLC, small cell lung cancer; dNTP, deoxy nucleotide triphosphate; CI, confidence intervals; PCR, polymerase chain reaction; SNP, single nucleotide polymorphism; WBC, white blood cells; RBC, red blood cells

* Corresponding author.

E-mail address: siddharthsharma.phd@thapar.edu (S. Sharma).

<http://dx.doi.org/10.1016/j.mgene.2017.07.002>

Received 6 January 2017; Received in revised form 9 June 2017; Accepted 10 July 2017

Available online 12 July 2017

2214-5400/ © 2017 Elsevier B.V. All rights reserved.

suggests the crucial role of this pathway in lung carcinogenesis (Mazieres et al., 2005), specifically lung adenocarcinoma (Coscio et al., 2014). The pathway has also been implicated in various other cancers such as colorectal, breast, ovarian cancer (Alanazi et al., 2013). Chronic activation of the Wnt pathway may occur at distinct subcellular levels, ranging from the alterations in the membranous proteins/ligands to the dysregulated activity of the transducers in the cytoplasm leading to the constitutive expression of this cascade (Luo et al., 2007). As in case of canonical pathway, the cascade of signals is transduced by the Wnt ligands of which 19 have been reported in the humans which bind to the surface membrane proteins, frizzled receptor family of proteins and LRP family (LDL-receptor related protein). It is post this interaction activation of Dvl (dishevelled) proteins occurs which inhibits the GSK-3 β from phosphorylating β -catenin and hence it is stabilized in the cytoplasm and enters the nucleus to bind the transcription factors such as TCR/LEF family to form a complex which controls the expression of target genes of this cascade, for e.g. *c-myc*, *cyclin D1*, *VEGF-A* and *MMP7* (Akiyama, 2000). All these target genes have been implicated in cancer as they all are involved in the cell proliferation pathways (Jones and Jomary, 2002). Basically, sFRP proteins antagonise the canonical pathway by binding to the Wnt ligands as they have homologous CRD (cysteine rich domain) at their N-terminal which is similar to the cysteine rich domain of frizzled receptors (Shi et al., 2007) but they do not possess any transmembrane domain (Akiyama, 2000) and hence the signal transduction is disabled, leading to the formation of destruction complex which phosphorylates the β -catenin so that it is ubiquitinated and undergoes proteolytic degradation.

Secreted frizzled related proteins belong to the family of extracellular glycoproteins which include five different members, sFRP1-sFRP5. Out of the five members, sFRP1, sFRP2 and sFRP5 are closely related and are coded by three exons located on chromosome 8p12-p11.1, 4q31.3 and 10q24.1 respectively. On the other hand, sFRP3 and sFRP4 form another group and both are encoded by six exons located on chromosome 2q31-q33 and 7p14-p13, respectively (Jones and Jomary, 2002). All the five members have been confirmed to antagonise the Wnt pathway (Shi et al., 2007), that recently have been found to play a major role in lung cancer pathogenesis (Mazieres et al., 2005). Alongside, recent studies support that an altered expression of these proteins have significantly contributed towards various cancers such as colorectal, ovarian, prostate cancer etc. (Surana et al., 2013). Allelic loss of the genomic region where sFRP3 is located has been observed in various cancers such as colorectal, prostate, gastric, papillary bladder, head and neck cancer, and non-small cell lung carcinomas (Shanmugam et al., 2007). Malignant mesothelium and osteogenic sarcoma when experimentally examined reported a lower expression of sFRP3 protein. sFRP3 gene has a distinct mechanism of regulation which is still unclear as it lacks the CpG-island in the promoter region unlike the other members of the sFRP family (Shi et al., 2007). Also, the epigenetic hyper methylation induced silencing of sFRP1 gene promoter has been established in many cancers such as hepatic (Shih et al., 2006), ovarian (Takada et al., 2004), breast cancer (Veeck et al., 2006). Also this epigenetic dysregulation of sFRP1 is reported in non-small cell lung cancer cell lines (Taguchi et al., 2016, Licchesi et al., 2008, Shi et al., 2007, Fukui et al., 2005) and NSCLC tissue samples (Zhang et al., 2011, Zhang et al., 2009).

Certain germline genetic polymorphisms of the sFRP gene are influential in modulating the susceptibility of occurrence and progression of different cancers. One such polymorphic non-synonymous variant sFRP4 Pro³²⁰Thr has been associated with early stage colorectal cancer (Frank et al., 2010) and has also been evaluated for its role in renal cancer (Hirata et al., 2009). Also, another less explored polymorphic site is rs1802074 Arg³⁴⁰Lys which has been explored for its role in process of renal carcinogenesis (Hirata et al., 2009). These two genetic variants of the sFRP4 gene have been genotyped and associated with the bone mineral density and body composition in Danish men (Boudin et al., 2012). The polymorphic sites rs288326 and rs7775 located on

sFRP3 gene have been studied for evaluating their association with colorectal cancer (Shanmugam et al., 2007). Additionally, these SNPs have been extensively studied to decipher their role in the process of osteoarthritis (Loughlin et al., 2004) and osteolysis (Mothersill et al., 1991). Henceforth, it is evident from the above cited literature that the polymorphic forms of the sFRP3 and sFRP4 genes may have certain impact on the functional aspect of these proteins which may lead to the aberrant changes in the expression of the highly complex intermeshing Wnt pathway which is been implicated in the process of oncogenesis. The primary drive behind evaluating these genetic variants of sFRP3 and sFRP4 and their cumulative role lies in the fact that there are no or a very few reports available so far for these SNPs with regard to lung cancer and the above mentioned findings state the they might be important in explaining the aberrant activation of Wnt cascade in lung cancer pathogenesis which is majorly regulated by Wnt antagonists such as sFRP. Taking into consideration, the current literature remains scanty in terms of evidences supporting the role of sFRP3 and sFRP4 genes in lung carcinogenesis, the present study, evaluated the role of four different SNPs namely sFRP3 Arg²⁰⁰Trp and sFRP3 Arg³²⁴Gly along with sFRP4 Pro³²⁰Thr and sFRP4 Arg³⁴⁰Lys in modulating the susceptibility towards Lung cancer in north Indian patients. This study will help in forecasting the role sFRP protein as a predictive and prognostic marker for lung carcinogenesis.

2. Materials and methods

2.1. Study population

The case-control study comprised 300 cases and 311 controls, hence a total of 611 subjects were recruited in the study. The cases were obtained from the Department of Pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER) Chandigarh, India. The Institutional Ethics committee of PGIMER has approved and reviewed the study. A written consent of the subjects was obtained by informing them about the details of the study and getting their approval. 5 ml blood was collected from each subject. Newly diagnosed primary lung cancer cases were included in the study with no age, gender, smoking, histological and TNM restrictions. The only exclusion criteria was to exclude those subjects who had a past history of any carcinoma. The histological classification of the cases was done into NSCLC and SCLC. Simultaneously, the control subjects were recruited from the hospital, these were the patients who came for normal health check-ups. The other important factor which was taken into consideration was sampling bias. In order to overcome that, age (± 10 years), gender and smoking habits were matched between the two study groups. However, it was certainly not possible to exactly match the smoking parameter. A detailed questionnaire was filled for each case and control by a trained interviewer. All the information regarding the epidemiological factors and the smoking habits of the subjects was obtained with the help of a questionnaire form filled by a trained personnel. In order to obtain a clear idea about the total smoking exposure, pack-years were obtained using this equation: [(cigarettes or beedis per day/20) X number of years smoked]. The clinical details of the patients such as histology, TNM staging, metastasis and other important information was collected from the medical records of the patients from the hospital.

2.2. DNA extraction

The extraction of the genomic DNA was done from the peripheral blood using the protocol by Sodhi et al., 2015. Firstly, the blood (3 ml) was thrice treated with washing buffer containing 320 mM sucrose, 1.5% Triton X-100, 20 mM Tris-HCl (pH 8.0) and 5 mM MgCl₂ in order to separate the red blood cells from the whole blood. Further, this RBC free pellet as obtained was lysed overnight using the WBC lysis buffer having 400 mM Tris-HCl (pH 8.0), 200 mM NaCl, 80 mM EDTA, 1%

SDS, and 100 µg/ml proteinase K and incubated at 45 °C overnight. Removal of proteins was done using phenol/chloroform/Isoamyl alcohol (25:24:1) solution. The mixture was centrifuged at 8000 rpm for 7 min, to separate the aqueous phase into a new vial. DNA was finally precipitated using the equal volume of chilled isopropanol which was obtained after centrifugation at 10,000 rpm for 5 min and washed with ice-cold 70% ethanol, and finally resuspended in 200 µl of sterile Tris-EDTA buffer. Qualitative and quantitative estimation of DNA was done using agarose gel electrophoresis and spectrophotometric estimation using Nanodrop (NanoDrop ND-1000 spectrophotometer). The stock DNA was diluted to 100 ng/µl and was stored at – 20 °C for further use as template.

2.3. Genotyping of the *sFRP3* (rs288326 and rs7775) and *sFRP4* (rs1802073 and rs1802074) using polymerase chain reaction-restriction fragment length polymorphism analysis (PCR-RFLP)

Two SNPs of the *sFRP3* gene (rs288326 and rs7775) were genotyped using the PCR-RFLP technique which was reported earlier (Shanmugam et al., 2007). The other two SNPs of the *sFRP4* gene, the primer sequences and restriction enzymes were used in accordance to those which were previously detailed by Hirata et al., 2009 where a nested-PCR approach was used to amplify the desired fragment.

The PCR reaction was performed in 25 µl of reaction mixture containing 1 × PCR buffer, 1.5 mM MgCl₂, with 0.5 µM of both primer, 200 µM of each dNTP's, 100 µg/ml bovine serum albumin (BSA) and 1 U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 300 ng DNA. The PCR conditions included an initial melting step of 94 °C for 5 min and followed by 30 cycles of 94 °C for 45 s, variable annealing temperature for 30s and 72 °C for 45 s for all the four polymorphic sites. PCR amplification was checked by using 1.5% agarose gel electrophoresis. Further, the amplified products were digested with 5 unit of each restriction enzyme procured from New England Biolabs, Ipswich, MA, USA. For *sFRP3* rs288326 C > T, C allele generates two fragments of 130 bp and 40 bp, T allele remains uncut when treated with *TaqI*. Similarly for *sFRP3* rs7775 C > G, C allele generates two fragments of 154 bp and 39 bp, G allele remains uncut when digested by the restriction enzyme *BstU1*. In case of *sFRP4* rs1802073 C > G, the wild genotype CC was uncut and the mutant genotype GG gave two bands of 165,81 bp and the heterozygous had three bands of 246, 165 and 81 bp when treated with *HinfI*. In the similar manner, the genotyping for *sFRP4* rs1802074 G > A was done with *EarI*, where presence of heterozygote was identified by the presence of three bands i.e. 246, 135 and 121 bp. The reproducibility of the genotyping results was very important and was checked by repeating it for 15% random samples without knowing the status of sample of being a case or a control.

2.4. Statistical analysis

The statistical variation within the demographic variables of the two study groups was done using the chi-square tests (χ^2 test) for the categorical data and student *t*-test for continuous variables. To find out the statistical difference between the allelic and genotypic frequencies of cases and controls was done by Pearson's χ^2 test. To check whether the genotypic frequencies of controls of four SNPs were in Hardy-Weinberg equilibrium (HWE) ($p^2 + 2pq + q^2 = 1$), the χ^2 test was used. To evaluate the association of lung cancer susceptibility and *sFRP* polymorphisms, the logistic regression analysis was used to calculate adjusted odds ratio (AORs) along with 95% Confidence Intervals (CI) with adjustment for possible confounders (age and pack-years of smoking as continuous variables (Sodhi et al., 2015)). Further to find out the high risk subgroups, stratified analysis was carried out to identify the modulation in susceptibility for various stratified categories made on the basis of smoking habits, histological and clinico-pathological features. Medcalc version 15.6.1 (MedCalc Software bvba, Ostend,

Table 1
Demographic characteristics among cases and controls.

Variable	Cases, n (%) N = 300	Controls, n (%) N = 311	p-Value
Age (years)			
Mean ± SD	57.3 ± 10.86	53.11 ± 10.42	< 0.0001
Range			
Gender			
Male	258 (86)	268 (86.2)	0.956
Female	42 (14)	43 (13.8)	
Smoking status			
Smokers	235 (78.3)	222 (71.4)	0.059
Non-smokers	65 (21.7)	89 (28.6)	
Pack years			
Mean ± SD	28.09 ± 34.09	19.4 ± 19.56	0.0003
Histological types			
SQCC	128 (42.7)		
ADCC	96 (32)		
SCLC	71 (23.7)		
Others	4 (1.3)		
Unknown	1 (0.3)		
TNM staging			
I	3 (1)		
II	12 (4)		
III	138 (46)		
IV	125 (41.7)		
Unclassified	22 (7.3)		
Tumor size			
Tx	10 (3.3)		
T1	14 (4.7)		
T2	39 (13)		
T3	78 (26)		
T4	127 (42.3)		
Unknown	32 (10.7)		
Lymph node involvement			
N0	49 (16.3)		
N1	36 (12)		
N2	111 (37)		
N3	72 (24)		
Unknown	32 (10.7)		
Metastasis			
M0	152 (50.6)		
M1	116 (38.7)		
Unknown	32 (10.7)		

Abbreviations: SD = standard deviation, n = total number of case patients or control subjects. *p-Values were derived from Pearson Chi-square test except age; Student *t*-test was used for age. All p-values are two-sided. *p* < 0.05 was considered statistically significant. The significant p-values are highlighted in bold.

Belgium) and SPSS Version 20.0. (Chicago, IL, USA) were the two software tools used to carry out all the statistical analysis. The haplotype analysis was carried out using the SHEis program (<http://analysis.bio-x.cn>). Non-parametric interaction analysis was carried out using multifactor dimensionality reduction approach. Genotypes are pooled into two groups high and low risk, thus reducing the genotype predictors from n dimensions to one. The model having the highest cross-validation consistency (CVC), testing accuracy and significant permutation *p*-value were taken as the best interaction model. MDR tests were performed using the version 0.5.1 of the open source MDR software package that is available online (<http://www.epistasis.org>).

3. Results

3.1. Demographic characteristics

Demographic characteristics of the study population including age, gender, smoking status, pack years and clinico-pathological factors have been summarized in Table 1. The mean age of cases and controls was found to be 57.3 ± 10.86 and 53.11 ± 10.42 respectively. However, the proportion of males and females were almost comparable in cases and controls as the percentage of males and females in both the

study groups was found to be approximately same and no significant difference was observed ($p = 0.956$). The numbers of smokers were found to be over-represented in cases (78.3%) as compared to controls (71.4%) ($p = 0.06$). The mean pack years were also calculated for both cases (28.9) and controls (19.4), and there was significant difference between the pack years of both the study groups ($p = 0.0003$). Hence, smoking came out to be significant factor in assessing the risk for lung cancer. Furthermore, the clinico-pathological factors including histology, staging and TNM status of the lung cancer subjects has also been documented. Out of 300 subjects, 42.7% of the cases were confirmed with SQCC, 32% were found to have ADCC and 23.7% were classified as SCLC along with 1.3% as others and 0.3% as unknown. Further, the classification was done on the basis of stage, where the percentage of subjects falling in different categories were as follows: Stage I: 1%, Stage 2: 4%, Stage 3: 46% and Stage 4: 41.7% rest are unknown. The proportion of subjects possessing the clinical parameters namely, primary tumor extension (Tx, T1, T2, T3 and T4), lymph node involvement (N0, N1, N2 and N3) and metastasis (M0, M1) are also listed in Table 1.

3.2. Genotypic distribution of the four genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histological subtypes

The genotypic frequency distribution for all the four genetic variants is given in Table 2. The control group of all the four polymorphic sites showed some deviations from Hardy Weinberg equilibrium. The possible reasons for these deviations are attributable to the fact that HWE which is based on various assumptions such as random mating, genetic drift etc. which might be violated in this sample set taken from the entire population. The reasons include population stratification, sample size and non-random mating. In case of *sFRP3* Arg²⁰⁰Trp C > T, the proportion of subjects with the wild type genotype (CC) were 84.8% and 87.8% in both cases and controls respectively. Also, no major difference was found in the frequency of subjects with CT genotype in both the study groups ($\chi^2 = 1.68$, df = 2, $p = 0.42$). The minor allele (T) frequency was 0.081 and 0.10 in controls and cases respectively. With CC genotype as a reference, the association analysis for estimating the susceptibility towards lung cancer was carried out using multivariate logistic regression after adjusting for age, gender and smoking parameter. No significant odds for developing lung cancer in patients carrying the mutant TT genotype for *sFRP3* Arg²⁰⁰Trp was observed (OR = 1.47, 95% CI: 0.70–3.10, $p = 0.30$). Genotypic frequencies for *sFRP3* Arg³²⁴Gly C > G in cases and controls respectively is as follows CC, CG, and GG: 77.3% vs 78.8%, 18% vs 15% and 4.7% vs 5.5% ($\chi^2 = 0.69$, df = 2, $p = 0.70$). The MAF for both cases and controls was found to be 0.13 and 0.13 however, the association analysis showed that subjects with GG genotype or variant (CG + GG) genotype were found not to have any significant predisposition towards lung cancer. For *sFRP4* Pro³²⁰Thr C > A, the distribution of the genotypic frequencies for both cases and controls was found to be significantly different as the proportion of subjects with mutant AA genotype were found to be over represented in cases as compared to controls (29.7% vs 14.5%) $\chi^2 = 23.10$, df = 2, $p < 0.0001$. The minor allele frequency (A) was found to be 0.45 and 0.57 in controls and cases respectively. Taking CC genotype as a reference, the susceptibility towards lung cancer was estimated. It was observed that subjects with AA genotype were found to have a three-fold increased risk for developing lung cancer (OR = 3.2, 95% CI: 1.88–5.46, $p < 0.0001$). Similarly, when both the mutant and heterozygous genotype were combined as a single genotype i.e. CA + AA, it showed approximately two-fold higher risk of developing lung cancer (OR = 1.71, 95% CI: 1.13–2.61, $p = 0.01$). In case of *sFRP4* rs1802074 G > A, the genotypic frequencies for genotype GG, GA and AA genotype were as follows in cases and controls (12% vs 18.3%, 82.2% vs 73.3%, 5.7% vs 8.4) ($\chi^2 = 7.19$, df = 2, $p = 0.02$). The minor allele frequency was found to be 0.46 and 0.45 in cases and controls respectively. The subjects with heterozygous (GA)

genotype showed a marginal risk towards lung cancer, OR = 1.66, 95% CI: 1.04–2.64, $p = 0.03$. The combined variant (GA + AA) genotype also showed a marginal significant risk associated with lung cancer (OR = 1.61, 95% CI: 1.01–2.57, $p = 0.04$). Additionally, the subjects were stratified on the basis of gender, in case of *sFRP3* Arg²⁰⁰Trp female subject with combined (CT + TT) genotype were found to have a pronounced risk towards lung cancer (OR = 7.25, 95% CI: 1.36–38.64, $p = 0.02$). A similar trend was observed in case of female subjects with combined (CG + GG) genotype of *sFRP3* Arg³²⁴Gly, where the risk was found to be almost three times with OR = 2.86, 95% CI: 0.9–9.09, $p = 0.07$. Also, the female subjects with the combined variant (C-A + AA) genotype of *sFRP4* Pro³²⁰Thr were found to have five times higher risk of developing lung cancer, however this trend was not followed in the case of female subjects with the variant genotype of *sFRP4* Arg³⁴⁰Lys (data not shown).

Further, we also evaluated the risk for developing lung cancer after stratification on basis of histological subtypes as shown in Table 2. For *sFRP3* Arg²⁰⁰Trp C > T, the stratified analysis revealed that the adenocarcinoma patients with heterozygous (CT) genotype showed a two-fold higher risk towards lung cancer with OR = 2.32, 95% CI: 1.14–4.73, $p = 0.02$ which was found to be statistically significant. Similarly, for the combined CT + TT genotype a significant two-fold risk was observed (OR = 2.28 95% CI: 1.26–4.14 $p = 0.006$). However, the subjects having SQCC and SCLC didn't show any such association. In regard of *sFRP3* Arg³²⁴Gly C > G, no significant risk was found to be associated in case of all the three histological subtypes. However, very marginal trend was observed in case of ADCC subjects with combined variant genotype i.e. GG + GC (OR = 1.46; 95% CI: 0.84–2.52, $p = 0.17$) but it was not significant. On the contrary, in case of *sFRP4* Pro³²⁰Thr C > A, subjects with mutant (AA) genotype, exhibited almost four-fold risk in case of ADCC (OR = 3.67, 95% CI: 1.33–6.78, $p = 0.001$). A similar trend was also observed in case of SCLC (OR = 3.01, 95% CI: 1.33–6.78, $p = 0.007$). Also, in case of subjects with combined variant (CA + AA) genotype, a trend of increased susceptibility was observed, which was more pronounced and significant in ADCC subjects as compared to SQCC and SCLC (OR = 2.03, 95% CI: 1.05–3.94 $p = 0.03$). In case of *sFRP4* rs1802074 G > A, the subjects with heterozygous genotype (GA) exhibited a two-fold significant risk of developing lung cancer in ADCC (OR = 2.23, 95% CI: 1.06–4.69, $p = 0.03$) whereas a marginal association was observed in SCLC patients (OR = 2.39, 95% CI: 0.88–6.55, $p = 0.08$), however no such trend was observed in case of SQCC patients. The risk was almost twice in case of ADCC patients carrying the mutant (AA) genotype as compared to subjects with heterozygous (GA) genotype but was found to be not statistically significant (OR = 4.82, 95% CI: 0.75–30.7, $p = 0.09$).

3.3. Haplotype and linkage disequilibrium analysis

SHEis program platform was used as the tool for carrying the haplotype analysis, the results of analysis carried out for *sFRP3* Arg²⁰⁰Trp and *sFRP3* Arg³²⁴Gly is tabulated in Table 3a and the other two variants of *sFRP4* is summarized in Table 3b where the frequencies of the four possible haplotypes has been enlisted for both cases and controls. However, the haplotype frequencies < 0.03 among cases and controls have been excluded from the analysis. The haplotype C–C was taken as the reference for *sFRP3* haplotypes, similarly for *sFRP4* C–G was considered as the reference haplotype as it carried the wild alleles. The odds ratio for all the haplotypes are given in the tables along with the p -value. For *sFRP3* gene, significant association of T–C haplotype with lung cancer risk was observed (OR = 1.54, 95% CI: 1.01–2.34, $p = 0.03$). The global chi square test gave $\chi^2 = 4.74$, df = 2, $p = 0.09$ ($D = 0.009$). In case of *sFRP4* gene, lung cancer risk was found to be higher for A–A haplotype (OR = 1.71, 95% CI: 1.28–2.28, $p = 0.0001$). The global chi square test gave $\chi^2 = 21.31$, df = 3 for the comparison of the haplotypes among cases and controls for *sFRP4* ($D = 0.128$, $r^2 = 0.015$) suggesting that the two SNPs were not in linkage

Table 2

Genotypic distribution of the four genetic variants and their association with risk of lung cancer along with the stratified association analysis based on histological subtypes.

	Controls n (%) N = 311	Overall			SQCC			ADCC			SCLC		
		Cases n (%) N = 300	AOR (95% CI) ^b	p ^a	Cases n (%) N = 128	AOR (95% CI) ^b	p ^a	Cases n (%) N = 96	AOR (95% CI) ^b	p ^a	Cases n (%) N = 71	AOR (95% CI) ^b	p ^a
sFRP3 Arg 200 Trp													
CC	273 (87.8)	254 (84.7)	1.00 (reference)		113 (88.3)	1.00 (reference)		73 (76.1)	1.00 (reference)		64 (90.1)	1.00 (reference)	
CT	25 (8)	27 (9)	1.33 (0.73–2.41)	0.34	9 (7)	1.03 (0.45–2.38)	0.92	15 (15.6)	2.32 (1.14–4.73)	0.02	2 (2.8)	0.45 (0.10–2.01)	0.29
TT	13 (4.2)	19 (6.3)	1.47 (0.70–3.10)	0.30	6 (4.7)	0.83 (0.30–2.30)	0.73	8(8.3)	2.34 (0.91–5.96)	0.07	5 (7.1)	1.09 (0.36–3.31)	0.86
CT + TT	38 (12.2)	46 (15.3)	1.36 (0.84–2.19)	0.19	15 (11.7)	0.94 (0.48–1.82)	0.86	23 (23.9)	2.28 (1.26–4.14)	0.006	7 (9.8)	0.76 (0.31–1.85)	0.55
C	285	267											
T	25	32											
MAF	0.081	0.10											
sFRP3 Arg 324 Gly													
CC	245 (78.8)	232 (77.3)	1.00 (reference)		101 (78.9)	1.00 (reference)		70 (72.9)	1.00 (reference)		58 (81.7)	1.00 (reference)	
CG	49 (15.7)	54 (18)	1.12 (0.72–1.74)	0.59	22 (17.1)	1.06 (0.60–1.88)	0.82	22 (22.9)	1.68 (0.93–3.04)	0.08	8 (11.3)	0.66 (0.29–1.52)	0.33
GG	17 (5.5)	14 (4.7)	0.8231 (0.38–1.74)	0.61	5 (4)	0.63 (0.22–1.81)	0.39	4 (4.2)	0.87 (0.27–2.74)	0.81	5 (7)	1.05 (0.36–3.06)	0.92
GG + CG	66 (21.2)	68 (22.7)	1.0497 (0.70–1.55)	0.80	27 (21.1)	0.94 (0.56–1.59)	0.84	26 (27.1)	1.46 (0.84–2.52)	0.17	13 (18.3)	0.77 (0.39–1.53)	0.46
C	269	259											
G	41	41											
MAF	0.13	0.13											
sFRP4 Pro 320 Thr													
CC	75 (24.1)	46 (15.3)	1.00 (reference)		20 (15.6)	1.00 (reference)		13 (13.5)	1.00 (reference)		13 (18.3)	1.00 (reference)	
CA	191 (61.4)	165 (55)	1.37 (0.88–2.11)	0.15	75 (58.6)	1.29 (0.72–2.3)	0.38	54 (56.3)	1.63 (0.82–3.23)	0.16	33 (46.5)	0.99 (0.47–2.06)	0.98
AA	45 (14.5)	89 (29.7)	3.21 (1.88–5.46)	< 0.0001	33 (25.8)	2.49 (1.25–4.95)	0.009	29 (30.2)	3.67 (1.69–7.96)	0.001	25 (35.2)	3.01 (1.33–6.78)	0.007
CA + AA	236 (75.9)	254 (84.7)	1.71 (1.13–2.61)	0.01	108 (84.3)	1.52 (0.87–2.67)	0.13	83 (86.4)	2.03 (1.05–3.94)	0.03	58 (81.7)	0.80 (0.38–1.66)	0.55
C	170	128											
A	140	171											
MAF	0.45	0.57											
sFRP4 rs1802074													
GG	57 (18.3)	36 (12)	1.00 (reference)		20 (15.6)	1.00 (reference)		10 (10.4)	1.00 (reference)		5 (7.1)	1.00 (reference)	
GA	228 (73.3)	247 (82.3)	1.66 (1.04–2.64)	0.03	99 (77.3)	1.07 (0.60–1.93)	0.79	81 (84.4)	2.23 (1.06–4.69)	0.03	63 (88.7)	2.39 (0.88–6.55)	0.08
AA	26 (8.4)	17 (5.7)	1.7477 (0.70–4.30)	0.22	9 (7.1)	1.14 (0.40–3.28)	0.79	5 (5.2)	4.82 (0.75–30.7)	0.09	3 (4.2)	1.74 (0.31–9.5)	0.52
GA + AA	254 (81.7)	264 (88)	1.61 (1.01–2.57)	0.04	108 (84.3)	0.96 (0.54–1.74)	0.91	86 (89.5)	2.15 (1.02–4.55)	0.04	66 (93)	2.28 (0.85–6.13)	0.1
G	171	159											
A	140	140											
MAF	0.46	0.45											

Abbreviations: ADCC, adenocarcinoma; SQCC, squamous cell carcinoma; SCLC, small cell lung carcinoma. The significant *p*-values are highlighted in bold.^a Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.^b Adjusted odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

disequilibrium.

3.4. Genotypic distribution of the four genetic variants based on smoking status and its association with risk of developing lung cancer

Since it has been established that smoking plays an important role in determining the risk associated with lung cancer, we wanted to evaluate whether *sFRP* genetic variants played a role in modifying the susceptibility towards lung cancer. The study subjects were classified as smokers and non-smokers on basis of their smoking habits. Table 4 shows the genotypic distributions for both smokers and non-smokers for all the four *sFRP* SNP's. In case of *sFRP3* Arg²⁰⁰Trp C > T, no significant association was seen in smokers with variant (CT and TT) genotype of

this polymorphic site. However, it was observed that the non-smoker subjects with the combined variant CT + TT genotype showed an approximately three-fold increased risk (OR = 2.82, 95% CI: 1.16–6.83, *p* = 0.02) of developing lung cancer. When the analysis was carried out for *sFRP3* Arg³²⁴Gly C > G, neither the smokers nor non-smokers showed any significant association between the variant genotype with the susceptibility towards lung cancer, although a trend of increased risk was seen in non-smokers (OR = 1.73, 95% CI: 0.32–9.22, *p* = 0.51). However, in regard of *sFRP4* Pro³²⁰Thr C > A, smokers with mutant (AA) genotype showed 2.5 fold risk of developing lung cancer which was found to be significant (OR = 2.55, 95% CI: 1.39–4.70, *p* = 0.002). This effect was more pronounced in non-smokers with mutant (AA) genotype (OR = 5.88, 95% CI: 1.8–18.7, *p* = 0.002). Also,

Table 3a
Haplotype frequencies and lung cancer risk associated with the sFRP3 haplotypes.

sFRP3 Arg200Trp C/T & sFRP3 Arg324Gly C/G	Case frequency, n (%)	Control frequency, n (%)	OR (95% CI) ^b	p-Value ^a
C-C	0.76 (459)	0.80 (498)	0.77 (0.58–1.02)	0.07
C-G	0.12 (75)	0.11 (72)	1.08 (0.770–1.53)	0.63
T-C	0.09 (58)	0.06 (40)	1.54 (1.01–2.34)	0.03
T-G	0.01 (6)	0.01 (10)	–	–

The significant *p*-values are highlighted in bold.

^a Two-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls.

^b Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis.

Table 3b
Haplotype frequencies and lung cancer risk associated with the sFRP4 haplotypes.

sFRP4 Pro320Thr C/A & sFRP3 Arg340Lys G/A	Case frequency, n (%)	Control frequency, n (%)	OR (95% CI) ^b	p-Value ^a
A-A	0.24 (145)	0.15 (98)	1.71 (1.28–2.28)	0.0001
A-G	0.32 (197)	0.29 (182)	1.17 (0.92–1.49)	0.191
C-A	0.22 (135)	0.29 (181)	0.70 (0.54–0.91)	0.007
C-G	0.20 (121)	0.25 (159)	0.74 (0.56–0.96)	0.028

The significant *p*-values are highlighted in bold.

^a Two-sided χ^2 test for either haplotypic distribution or frequencies between the cases and controls.

^b Odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by logistic regression analysis.

Table 4
Genotypic distribution of the four genetic variants based on smoking status and its association with risk of lung cancer.

	Smokers				Non-smokers			
	Controls n (%) N = 222	Cases n (%) N = 235	AOR (95% CI) ^b	<i>p</i> ^a	Controls n (%) N = 89	Cases n (%) N = 65	AOR (95% CI) ^b	<i>p</i> ^a
sFRP3 Arg 200 Trp								
CC	195 (87.9)	205 (87.2)	1.00 (reference)		78 (87.6)	49 (75.4)	1.00 (reference)	
CT	14 (6.3)	19 (8.1)	1.28 (0.61–2.66)	0.50	11 (12.4)	8 (12.3)	1.30 (0.47–3.60)	0.61
TT	13 (5.8)	11 (4.7)	0.67 (0.28–1.56)	0.35		8 (12.3)		
CT + TT	27 (12.2)	30 (12.8)	0.97 (0.55–1.72)	0.93	11 (12.4)	16(24.6)	2.82(1.16–6.83)	0.02
sFRP3 Arg 324 Gly								
CC	170 (76.6)	181 (77)	1.00 (reference)		75 (84.3)	51 (78.5)	1.00 (reference)	
CG	38 (17.1)	43 (18.3)	1.09 (0.67–1.79)	0.71	11 (12.3)	11 (16.9)	1.39 (0.57–3.57)	0.48
GG	14 (6.3)	11 (4.7)	0.68 (0.29–1.60)	0.38	3 (3.4)	3 (14.6)	1.73 (0.32–9.22)	0.51
CG + GG	52 (23.4)	54 (23)	0.98 (0.63–1.53)	0.94	14 (15.7)	14 (21.5)	1.46 (0.62–3.42)	0.37
sFRP4 Pro 320 Thr								
CC	50 (22.5)	39 (16.6)	1.00 (reference)		25 (28)	71 (10.8)	1.00 (reference)	
CA	139 (62.6)	127 (54)	1.12 (0.68–1.84)	0.63	52 (58.4)	38 (58.5)	2.59 (0.94–7.13)	0.06
AA	33 (14.9)	69 (29.4)	2.55 (1.39–4.70)	0.002	12 (13.6)	20 (30.7)	5.88 (1.8–18.7)	0.002
CA + AA	172 (77.5)	196 (83.4)	1.39 (0.86–2.25)	0.16	64 (71.9)	58 (89.2)	3.23 (1.23–8.48)	0.01
sFRP4 rs1802074								
GG	32 (14.4)	28 (12)	1.00 (reference)		25 (28.1)	8 (12.3)	1.00 (reference)	
GA	169 (76.1)	192 (82)	1.37 (0.78–2.40)	0.27	59 (66.3)	55 (84.6)	2.81 (1.14–6.92)	0.02
AA	21 (9.5)	15 (6)	1.33 (0.51–3.46)	0.54	5 (5.6)	2 (3.1)	8.54 (0.5–122.8)	0.11
GA + AA	190 (85.5)	207 (88.1)	1.33 (0.76–2.34)	0.30	64 (72)	57 (87.7)	2.70 (1.10–6.61)	0.02

The significant *p*-values are highlighted in bold.

^a Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls.

^b Adjusted odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

the non-smokers with combined mutant and heterozygote (AA + CA) genotype exhibited a three-fold significantly increased risk of developing lung cancer (OR = 3.23, 95% CI: 1.23–8.48 *p* = 0.01). However, in case of sFRP4 Arg³⁴⁰Lys G > A there was no association found for the subjects who were smokers and carried the variant genotype for this polymorphic site. However, in case of non-smokers the subjects with heterozygote (GA) genotype showed higher risk of acquiring lung cancer (OR = 2.81, 95% CI: 1.14–6.92, *p* = 0.02). The subjects having the combined variant (GA + AA) genotype also showed a significant 2.7 fold risk of developing lung cancer (OR = 2.70, 95% CI: 1.10–6.61, *p* = 0.02). Furthermore, to find out the high risk subgroups the non-smokers were categorized according to their histology, where a high prevalence of risk was observed in case of ADCC subjects with the combined variant (CT + TT) genotype of sFRP3 Arg²⁰⁰Trp (OR = 4.74, 95% CI: 1.73–13.01, *p* = 0.0024). Similarly, the non-smoker ADCC subjects carrying mutant (AA) genotype for sFRP4 Pro³²⁰Thr showed a very high risk of developing lung cancer (OR = 12.56, 95% CI: 2.39–66.08, *p* = 0.0028) (data not shown).

3.5. Lung cancer risk associated with the different genotypic combinations of sFRP gene

The genotypic frequency distribution and lung cancer of all the possible genotypic combinations of the four SNP's of sFRP gene is listed in Table 5. The genotypic combination sFRP3 Arg²⁰⁰Trp C > T and sFRP3 Arg³²⁴Gly C > G, didn't show any association between the variant genotypes of both the polymorphic sites and lung cancer risk. Genotypic combination of sFRP3 Arg³²⁴Gly C > G and sFRP4 Pro³²⁰Thr C > A, showed that subjects with the mutant genotype (TT and AA) of both the genetic variants are at 2.5 fold risk of developing lung cancer, however it was not statistically significant. In the next genotypic combination comprising sFRP3 Arg²⁰⁰Trp C > T and sFRP4 Pro 320Thr C > A, the subjects with the heterozygous genotype (CT and CA) of both the polymorphic sites exhibited an increased risk of acquiring lung cancer with OR = 2.64, 95% CI: 1.05–6.67, *p* = 0.038. Also the subjects with combined variant genotype (CT + TT and CA + AA) of both

Table 5
Genotypic distribution based on different genotypic combinations and their association with lung cancer risk.

Genotypes	Cases, n (%) N = 194	Controls, n (%) N = 223	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP3 Arg200Trp & sFRP3 Arg324Gly				
0 (CC + CC)	192 (99)	216 (96.9)	1.00 (reference)	
1 (CT + CG)	1 (0.5)	4 (1.8)	0.33 (0.03–3.11)	0.33
2 (TT + GG)	1 (0.5)	3 (1.3)	0.19 (0.01–2.05)	0.17
1–2	2 (1)	7 (3.1)	0.25 (0.04–1.31)	0.10
Genotypes	Cases, n (%) N = 69	Controls, n (%) N = 93	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP3 Arg324Gly & sFRP4 Pro320Thr				
0 (CC + CC)	34 (49.3)	58 (62.4)	1.00 (reference)	
1 (CG + CA)	31 (44.9)	33 (35.5)	1.52 (0.76–3.03)	0.22
2 (GG + AA)	4 (5.8)	2 (2.1)	2.52 (0.37–16.9)	0.33
1–2	35 (50.7)	35 (37.6)	1.59 (0.81–3.11)	0.17
Genotypes	Cases, n (%) N = 68	Controls, n (%) N = 68	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP3 Arg200Trp & sFRP4 Pro320Thr				
0 (CC + CC)	40 (58.8)	56 (82.4)	1.00 (reference)	
1 (CT + CA)	18 (26.5)	12 (17.6)	2.64 (1.05–6.67)	0.03
2 (TT + AA)	10 (14.7)	–	–	–
1–2	28 (41.2)	12 (17.6)	3.94 (1.64–9.60)	0.002
Genotypes	Cases, n (%) N = 51	Controls, n (%) N = 72	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP3 Arg 200 Trp & sFRP4 rs1802074				
0 (CC + GG)	30 (58.8)	52 (72.2)	1.00 (reference)	
1 (CT + GA)	21 (41.2)	17 (23.6)	2.66 (1.12–6.09)	0.02
2 (TT + AA)	–	3 (4.2)	–	–
1–2	21 (41.2)	20 (27.8)	2.14 (0.95–4.83)	0.06
Genotypes	Cases, n (%) N = 147	Controls, n (%) N = 147	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP4 rs1802074 & sFRP4 Pro 320 Thr				
0 (CC + GG)	6 (4.1)	6 (4.1)	1.00 (reference)	
1 (CA + GA)	137 (93.2)	138 (93.9)	1.58 (0.54–4.60)	0.39
2 (AA + AA)	4 (2.7)	3 (2)	1.80 (0.16–19.37)	0.62
1–2	141 (95.9)	141 (94.1)	1.59 (0.55–4.63)	0.38
Genotypes	Cases, n (%) N = 76	Controls, n (%) N = 80	Adjusted OR (95% CI) ^b	p-Value ^a
sFRP3 Arg 324 Gly & sFRP4 rs1802074				
0 (CC + GG)	28 (36.8)	45 (56.2)	1.00 (reference)	
1 (CG + GA)	47 (61.8)	33 (41.3)	1.96 (0.93–4.13)	0.07
2 (GG + AA)	1 (1.4)	2 (2.5)	0.24 (0.005–10.35)	0.46
1–2	48 (63.2)	35 (43.5)	1.86 (0.88–3.92)	0.10

The significant p-values are highlighted in bold.

^a Two-sided χ^2 test for either genotype distribution or allelic frequencies between the cases and controls. 0: wild genotype, 1: heterozygote genotype, 2: mutant genotype, 1–2: combined hetero and mutant genotype.

^b Adjusted odds ratios, 95% confidence intervals and their corresponding p-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking status.

the polymorphic sites, showed almost a significant four –fold increased risk of developing lung cancer (OR = 3.94, 95% CI: 1.64–9.60 $p = 0.002$). Similarly a significant association was observed between the susceptibility towards lung cancer and genotypic combination *sFRP3* Arg²⁰⁰Trp C > T and *sFRP4* Arg³⁴⁰Lys G > A, where subjects with heterozygous genotype (CT and GA) of variants showed a two-fold increased Odds (OR = 2.66, 95% CI: 1.12–6.09, $p = 0.02$) of developing lung cancer. The dichotomized combined hetero and mutant group of this genotypic combination revealed a two-fold increased risk

of acquiring lung cancer (OR = 2.14, 95% CI: 0.95–4.83 $p = 0.06$). A certain trend of increased susceptibility is observed in case of the subjects with variant genotype of both *sFRP4* Pro³²⁰Thr C > A and *sFRP4* Arg³⁴⁰Lys G > A genetic variant. Also, it was observed that the subjects with the heterozygous genotype (CG and GA) of *sFRP3* Arg³²⁴Gly C > G and *sFRP4* Arg³⁴⁰Lys G > A showed almost two fold increased risk of acquiring lung cancer, but it was not statistically significant.

Table 6
Relationship of clinicopathological features and sFRP3 & sFRP4 genotypes in lung cancer patients.

Variable	Clinical stage		AOR (95% CI)	p-Value	TNM stage		AOR (95% CI) ^a	p-Value	Metastasis		AOR (95% CI)	p-Value
	III	IV			T3 + T4	T2 + T1			Positive	Negative		
<i>sFRP3</i> Arg ²⁰⁰ Trp	138	125										
CC	118		Reference		176	54	Reference		99	131	Reference	
CT	14	7	0.52 (0.19–1.4)	0.19	18	6	0.91 (0.34–2.4)	0.86	9	15	0.70 (0.28–1.74)	0.45
CT + TT	20	9	0.80 (0.34–1.62)	0.55	29	9	0.99 (0.44–2.25)	0.99	17	21	0.82 (0.45–1.91)	0.84
<i>sFRP3</i> Arg ³²⁴ Gly	138	125										
CC	116	88	Reference		160	45	Reference		83	122	Reference	
CG	18	28	2.16 (1.10–4.29)	0.024	37	12	0.85 (0.40–1.78)	0.67	26	23	1.79 (0.93–3.43)	0.076
CG + GG	22	37	2.31 (1.25–4.27)	0.007	45	18	0.69 (0.36–1.3)	0.27	31	30	1.79 (0.95–3.07)	0.071
<i>sFRP4</i> Pro ³²⁰ Thr	138	125										
CC	22	19	Reference		31	11	Reference		19	23	Reference	
CA	74	65	0.93 (0.45–1.90)	0.84	116	33	1.36 (0.60–3.07)	0.44	60	89	0.98 (0.96–1.01)	0.36
AA	42	41	1.08 (0.50–2.32)	0.84	58	19	1.04 (0.42–2.54)	0.92	37	40	1.07 (0.49–2.32)	0.85
CA + AA	116	106	0.97 (0.48–1.92)	0.93	174	52	1.22 (0.57–2.64)	0.59	97	129	0.81 (0.41–1.61)	0.56
<i>sFRP4</i> Arg ³⁴⁰ Lys	138	125										
GG	18	12	Reference		19	10	Reference		10	19	Reference	
GA	109	109	1.56 (0.69–3.52)	0.28	172	51	1.67 (0.71–3.88)	0.23	102	121	1.56 (0.67–3.6)	0.29
GA + AA	120	113	1.50 (0.66–3.38)	0.32	186	53	1.72 (0.74–4.00)	0.20	106	133	1.50 (0.65–3.48)	0.33

The significant *p*-values are highlighted in bold.

^a Adjusted odds ratios, 95% confidence intervals and their corresponding *p*-values were calculated by unconditional logistic analysis after adjusting for age, gender, smoking and histological sub-types.

3.6. Clinico-pathological parameters and their association with genotype

Various clinico-pathological parameters such as clinical staging, primary tumor extension, metastasis etc. were analysed to find out their correlation with the genotypes of various sFRP variants. This association analysis basically gives an insight about the role of the studied SNPs in tumor progression. The subjects were categorized into groups of two on the basis of their clinical stage, primary tumor extension and metastasis as summarized in Table 6. Out of all the genetic variants, association was seen in case of subjects with combined variant (C-G + GG) genotype of *sFRP3* Arg³²⁴Gly and clinical staging, with OR = 2.31, 95% CI: 1.25–4.27, *p* = 0.007. Also, the metastatic status was found to be positively correlated with the combined variant (C-G + GG) genotype of *sFRP3* Arg³²⁴Gly with OR = 1.79, 95% CI: 0.95–3.07, *p* = 0.07. However, a slight association was also observed in case of subjects with variant genotype of *sFRP4* Arg³⁴⁰Lys and three clinical parameters but the association was not found to be significant. Therefore, it can be concluded that *sFRP3* Arg³²⁴Gly variant might be playing a role in tumor progression if not in modulating the susceptibility towards lung cancer.

3.7. Multifactor dimensionality reduction (MDR) analysis

High order gene-gene interactions cannot be determined by the traditional logistic regression approach, so this non parametric method was used to obtain a clarity about the factors governing lung cancer susceptibility. The best interaction model which has CVC of 10/10, prediction error = 0.43 and *p*-value > 0.0001 was one factor model, sFRP4_{rs1802073}. The second best model was the three factor model comprising sFRP4_{rs1802073}, sFRP3_{rs288326} and sFRP3_{rs7775} with CVC of 6/10 and prediction error = 0.48. This model shows the interactive effect of two genes which are keenly involved in the regulation of Wnt signalling and are found to contribute significantly in governing lung cancer predisposition.

4. Discussion

Wnt signalling cascade is a highly complex and diverse pathway which is implicated in various biological processes such as embryonic development, cell proliferation, cell adhesion, cell polarity, tumorigenesis and apoptosis (Veeck and Dahl, 2012). This pathway is

regulated by Secreted frizzled related proteins which basically shares a structural homology at the N-terminal with the frizzled related receptors. This structural homology aids them in antagonising the complex signal cascade by competing with the frizzled related receptors for binding to the Wnt ligands which transduce the signal cascade (Pohl et al., 2014). Aberrant Wnt signalling cascade has been recently reported to have an important implication in lung cancer and its progression (Mazieres et al., 2005). In the present study, first attempt has been made to evaluate the influence of the germline genetic polymorphisms in the sFRP genes namely *sFRP3* and *sFRP4* on individual's predisposition towards lung cancer.

The present case control study revealed certain significant associations between the genetic variants of *sFRP3* and *sFRP4* gene and individual's susceptibility towards lung cancer. The two *sFRP3* polymorphic sites namely *sFRP3* Arg²⁰⁰Trp and *sFRP3* Arg³²⁴Gly were analysed to find out their effect on the lung cancer risk. Both these polymorphic sites exhibit a non-synonymous substitution due to the single base pair change and have been reported to increase the β -catenin levels in the cells which might occur due to the inability of this truncated protein to antagonise the Wnt signalling cascade (Páez et al., 2014). It has been said that arginine residues in both SNPs are conserved among the species. Both of these polymorphic sites are located near the C-terminal end of the sFRP3 protein but the Wnt ligands bind at cysteine rich N-terminal of the sFRP proteins (Shanmugam et al., 2007). In the present study these two sFRP3 genetic variants didn't show any pronounced association with lung cancer susceptibility. To best of our knowledge, there have been no reports regarding these polymorphisms particularly in lung cancer. However, a similar study performed on German population which reported a significantly increased risk of colorectal cancer associated with the variant genotype of *sFRP3* Arg³²⁴Gly. But the same study revealed the lack of association of *sFRP3* Arg²⁰⁰Trp variant genotype with risk of colorectal carcinoma, which is in accordance with results observed from the present study (Shanmugam et al., 2007). On the contrary, another study on colorectal neoplasia reported a lack of association of *sFRP3* Arg³²⁴Gly and *sFRP3* Arg²⁰⁰Trp variant genotype with risk of colorectal neoplasia (Berndt et al., 2009). There have been contradictory reports on the role of these genetic variants in modulating the susceptibility. These polymorphic sites have been extensively studied for their role in osteoarthritis (Loughlin et al., 2004) and osteolysis (Mothersill et al., 1991). The role of sFRP3 has also been examined in renal cancer cells (Hirata et al.,

2010) and androgen dependent prostate cancer cell lines (Zi et al., 2005). These studies explained that sFRP3 expression plays an important role in cell division and metastasis of renal and prostate cancer cells.

The other two genetic variants which are evaluated in the present study includes the sFRP4 genetic variants located in exon 6 namely *sFRP4* Pro³²⁰Thr C > A and *sFRP4* Arg³⁴⁰Lys G > A. Being characterized as a tumor suppressor gene, sFRP4 has been investigated for its altered expression levels in various tumors such as endometrial, prostate, ovarian, cervical, bladder, colorectal, mesothelium, oesophageal etc. (Pohl et al., 2014). *sFRP4* Pro³²⁰Thr has been studied in colorectal cancer where a significant association was observed between the variant genotype and increased risk of colorectal cancer where it was inferred that this genetic variant leads to dysregulation of Wnt signalling pathway which results in the altered expression of the downstream gene targets that are involved in cell proliferation and apoptosis (Frank et al., 2010). Our results also exhibited an increased risk of lung cancer in association with the variant genotype of *sFRP4* Pro³²⁰Thr. This gives us an insight that this non-synonymous SNP which is being classified as “probably damaging” by the Polyphen software might have an influence on the functional properties of the sFRP4 protein. This SNP has also been evaluated to determine its association with renal carcinoma where there was no significant conclusion about the role of this variant was observed (Hirata et al., 2009). There have been very few studies focussing on the modulation of susceptibility towards lung cancer due to sFRP4 polymorphisms. However, the two polymorphic variants have been studied for their role in bone mineralization (Boudin et al., 2012) and *sFRP4* Pro³²⁰Thr has been evaluated for its role in craniotubular hyperostosis (Boudin et al., 2013). *sFRP4* Arg³⁴⁰Lys G > A is least studied polymorphic site, there are not many reports available to compare the results obtained from the present study. The findings from the study done on Turkish population state that rs1802074 GG and GA genotype confers a decreased risk of lung cancer when compared with AA genotype (Yilmaz et al., 2015). In the present study, this polymorphic site GA + AA genotype showed a marginal role in modifying the susceptibility towards lung cancer, however the stratified analysis did reveal certain high risk subgroups.

Due to variation in the pathogenesis of different histological subtypes of lung cancer, they exhibit slightly different clinical manifestations, also they are mediated through distinct alterations in expression of various genes and proteins. Therefore, it can be concluded that an individual's predisposition towards a certain type of lung cancer might be influenced by different genetic variants (Sodhi et al., 2015). In the present study, *sFRP3* Arg²⁰⁰Trp did show a major risk associated with ADCC, however a similar trend was observed in case of *sFRP3* Arg³²⁴Gly but with a much lesser risk associated with its variant genotype. *sFRP4* Pro³²⁰Thr variant genotype did show a higher risk towards ADCC as compared to SQCC and SCLC. Similarly, *sFRP4* Arg³⁴⁰Lys variant genotype also displayed a higher risk towards ADCC. Recently, experiments done in murine models have confirmed that Wnt signalling pathway plays a major role in NSCLC pathogenesis (Stewart, 2014). Evidences are also available which clearly demonstrate that aberrant activation of Wnt signalling pathway might lead in lung adenocarcinoma progression (Coscio et al., 2014). Further, due to less sample size in the stratified categories, the findings from the present study do need further validation. In order to figure the role of smoking which is a major risk factor for lung cancer, we analysed the prevalence of the variant genotypes in smokers and non-smokers, as it is important to detect the gene-environment interactions in order to demarcate the subgroups who are at higher risk of acquiring lung cancer due to environmental factors (Li et al., 2010). In the present study, non-smokers were found to be at a higher risk as compared to smokers if they had a variant genotype for a polymorphic variant. In case of all the three genetic variants namely *sFRP3* Arg²⁰⁰Trp, *sFRP4* Pro³²⁰Thr and *sFRP4* Arg³⁴⁰Lys, non-smoker subjects were found to be at a higher risk of developing lung cancer. The molecular mechanism for development of

adenocarcinoma which is the only subtype of lung cancer that is weakly related to smoking and the process of carcinogenesis in non-smokers both are highly unclear (Li et al., 2010). The findings from the present study can be an aid in predicting the molecular changes occurring during the process of tumorigenesis. Combinatorial genotypic analysis was carried out to find out the higher risk subgroups with a particular genotypic combination. Of all the different combinations assessed, the major significant risk was observed in the two genotypic combinations namely *sFRP3* Arg²⁰⁰Trp & *sFRP4* Pro³²⁰Thr and *sFRP3* Arg²⁰⁰Trp & *sFRP4* Arg³⁴⁰Lys.

5. Conclusion

The above findings will be helpful in estimating the individuals risk towards lung cancer in accordance to different genotypic combinations for multiple genes. The present study is probably the first attempt in to find the possible implication of these germline polymorphic variations in secreted frizzled related protein, a Wnt antagonist in lung cancer patients. Wnt pathway is highly intervening signalling cascade having diversified roles and effects on the normal cell development and proliferation. This study uncovers another aspect of this pathway by evaluating the role these genetic variants in modulating the predisposition towards lung cancer. Certain interesting findings have been observed, as *sFRP4* Pro³²⁰Thr variant have been found to play a functional role in altering lung cancer risk. Also, the adenocarcinoma patients were found to show a certain trend for variant genotype for all polymorphic sites. This study is a preliminary one and further investigations and studies with much larger sample size are required to confer that the sFRP can act as a biomarker and predictive tool for detection and prognosis of lung cancer.

Acknowledgements

We would like to express our gratitude to all the subjects who participated in this current study and the work was supported by grant from the Indian Council of Medical Research, New Delhi, India.

Funding

This work was supported by grant from the Indian Council of Medical Research, New Delhi, India (Grant No. 5/13/126/2011/NCD-III).

Conflict of interest

The authors have no conflict of interest to declare.

References

- Akiyama, T., 2000. Wnt/ β -catenin signaling. Cytokine Growth Factor Rev. 11, 273–282.
- Alanazi, M.S., Parine, N.R., Shaik, J.P., Alabdulkarim, H., Ajaj, S.A., Khan, Z., 2013. Association of single nucleotide polymorphisms in Wnt signaling pathway genes with breast cancer in Saudi patients. PLoS One 8, e59555.
- Amos, C.I., Wu, X., Spitz, M.R., 1999. Is there a genetic basis for lung cancer susceptibility? Recent Results Cancer Res. 151, 3–12.
- Berndt, S.I., Huang, W.Y., Yeager, M., Weissfeld, J.L., Chanock, S.J., Hayes, R.B., 2009. Genetic variants in frizzled-related protein (FRZB) and the risk of colorectal neoplasia. Cancer Causes Control 20, 487–490.
- Boudin, E., Piters, E., Nielsen, T.L., et al., 2012. Single nucleotide polymorphisms in sFRP4 are associated with bone and body composition related parameters in Danish but not in Belgian men. Mol. Genet. Metab. 106, 366–374.
- Boudin, E., Piters, E., Fijalkowski, I., et al., 2013. Mutations in sFRP1 or sFRP4 are not a common cause of craniotubular hyperostosis. Bone 52, 292–295.
- Coscio, A., Chang, D.W., Roth, J., et al., 2014. Genetic variants of the Wnt signaling pathway as predictors of recurrence and survival in early-stage non-small cell lung cancer patients. Carcinogenesis 35, 1284–1291.
- Frank, B., Hoffmeister, M., Klopp, N., Illig, T., Chang-Claude, J., Brenner, H., 2010. Single nucleotide polymorphisms in Wnt signaling and cell death pathway genes and susceptibility to colorectal cancer. Carcinogenesis 31, 1381–1386.
- Fukui, T., Kondo, M., Ito, G., Maeda, O., Sato, N., Yoshioka, H., Yokoi, K., Ueda, Y., Shimokata, K., Sekido, Y., 2005. Transcriptional silencing of secreted frizzled related

- protein 1 (SFRP1) by promoter hypermethylation in non-small-cell lung cancer. *Oncogene* 24, 6323–6327.
- Hirata, H., Hinoda, Y., Nakajima, K., et al., 2009. Wnt antagonist gene polymorphisms and renal cancer. *Cancer* 115, 4488–4503.
- Hirata, H., Hinoda, Y., Ueno, K., Majid, S., Saini, S., Dahiya, R., 2010. Role of secreted frizzled-related protein 3 in human renal cell carcinoma. *Cancer Res.* 70, 1896–18905.
- Jones, S.E., Jomary, C., 2002. Secreted frizzled-related proteins: searching for relationships and patterns. *BioEssays* 24, 811–820.
- Li, Y., Sheu, C.C., Ye, Y., et al., 2010. Genetic variants and risk of lung cancer in never smokers: a genome-wide association study. *Lancet Oncol.* 11, 321–330.
- Licchesi, J.D., Westra, W.H., Hooker, C.M., Machida, E.O., Baylin, S.B., Herman, J.G., 2008. Epigenetic alteration of Wnt pathway antagonists in progressive glandular neoplasia of the lung. *Carcinogenesis* 29, 895–904.
- Loughlin, J., Dowling, B., Chapman, K., et al., 2004. Functional variants within the secreted frizzled-related protein 3 gene are associated with hip osteoarthritis in females. *Proc. Natl. Acad. Sci. U. S. A.* 101, 9757–9762.
- Luo, J., Chen, J., Deng, Z.-L., et al., 2007. Wnt signaling and human diseases: what are the therapeutic implications? *Lab. Invest.* 87, 97–103.
- Mazieres, J., He, B., You, L., Xu, Z., Jablons, D.M., 2005. Wnt signaling in lung cancer. *Cancer Lett.* 222, 1–10.
- Mothersill, C., Seymour, C.B., O'Brien, A., 1991. Induction of c-myc oncoprotein and of cellular proliferation by radiation in normal human urothelial cultures. *Anticancer Res.* 11, 1609–1612.
- Páez, D., Gerger, A., Zhang, W., et al., 2014. Association of common gene variants in the WNT/ β -catenin pathway with colon cancer recurrence. *Pharm. J.* 14, 142–150.
- Pohl, S., Scott, R., Arfuso, F., Perumal, V., Dharmarajan, A., 2014. Secreted frizzled-related protein 4 and its implications in cancer and apoptosis. *Tumor Biol.* 36, 143–152.
- Shanmugam, K.S., Brenner, H., Hoffmeister, M., Chang-Claude, J., Burwinkel, B., 2007. The functional genetic variant Arg324Gly of frizzled-related protein is associated with colorectal cancer risk. *Carcinogenesis* 28, 1914–1917.
- Shi, Y., He, B., You, L., Jablons, D.M., 2007. Roles of secreted frizzled-related proteins in cancer. *Acta Pharmacol. Sin.* 28, 1499–1504.
- Shih, Y.L., Shyu, R.Y., Hsieh, C.B., et al., 2006. Promoter methylation of the secreted frizzled-related protein 1 gene SFRP1 is frequent in hepatocellular carcinoma. *Cancer* 107, 579–590.
- Sodhi, K.K., Bahl, C., Singh, N., Behera, D., Sharma, S., 2015. Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians. *Future Oncol.* 11, 2159–2173.
- Stewart, D.J., 2014. Wnt signaling pathway in non-small cell lung cancer. *J. Natl. Cancer Inst.* 106, 1–11.
- Surana, R., Sikka, S., Cai, W., et al., 2013. Secreted frizzled related proteins: implications in cancers. *Biochim. Biophys. Acta - Rev. Cancer.* 1845, 53–65.
- Taguchi, Y., Iwadata, M., Umeyama, H., 2016. SFRP1 is a possible candidate for epigenetic therapy in non-small cell lung cancer. *BMC Med. Genet.* 9, 28.
- Takada, T., Yagi, Y., Maekita, T., et al., 2004. Methylation-associated silencing of the Wnt antagonist SFRP1 gene in human ovarian cancers. *Cancer Sci.* 95, 741–744.
- Veeck, J., Dahl, E., 2012. Targeting the Wnt pathway in cancer: the emerging role of Dickkopf-3. *Biochim. Biophys. Acta - Rev. Cancer.* 1825, 18–28.
- Veeck, J., Niederacher, D., An, H., et al., 2006. Aberrant methylation of the Wnt antagonist SFRP1 in breast cancer is associated with unfavourable prognosis. *Oncogene* 25, 3479–3488.
- Wogan, G.N., Hecht, S.S., Felton, J.S., Conney, A.H., Loeb, L.A., 2004. Environmental and chemical carcinogenesis. 14, 473–486.
- Yang, Y., 2012. Wnt signaling in development and disease. *Cell. Biosci.* 2, 14.
- Yilmaz, M., Donmez, G., Kacan, T., Sari, I., Babacan, N.A., Sari, M., Kilickap, S., 2015. Significant association between polymorphisms of Wnt antagonist genes and lung cancer. *J. Investig. Med.* 63, 935–941.
- Yokota, J., Shiraishi, K., Kohno, T., 2010. Genetic basis for susceptibility to lung cancer: recent progress and future directions. *Adv. Cancer Res.* 109, 51–72.
- Zhang, Y.W., Miao, Y.F., Yi, J., Geng, J., Wang, R., Chen, L.B., 2009. Transcriptional inactivation of secreted frizzled-related protein 1 by promoter hypermethylation as a potential biomarker for non-small cell lung cancer. *Neoplasia* 57, 228–233.
- Zhang, Y., Wang, R., Song, H., et al., 2011. Methylation of multiple genes as a candidate biomarker in non-small cell lung cancer. *Cancer Lett.* 303, 21–28.
- Zi, X., Guo, Y., Simoneau, A.R., et al., 2005. Expression of Frzb/secreted frizzled-related protein 3, a secreted Wnt antagonist, in human androgen-independent prostate cancer PC-3 cells suppresses tumor growth and cellular invasiveness. *Cancer Res.* 65, 9762–9770.

ORIGINALITY REPORT

10%

SIMILARITY INDEX

7%

INTERNET SOURCES

6%

PUBLICATIONS

%

STUDENT PAPERS

PRIMARY SOURCES

1	Singh, Amrita, NavneetSingh, DigambarBehera, and Siddharth Sharma. "Association and multiple interaction analysis among five XRCC1 polymorphic variants in modulating lung cancer risk in North Indian population", DNA Repair, 2016. Publication	2%
2	Sodhi, Kushaldeep Kaur, Charu Bahl, Navneet Singh, Digamber Behera, and Siddharth Sharma. "Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians", Future Oncology, 2015. Publication	1%
3	academic.oup.com Internet Source	1%
4	www.onko-i.si Internet Source	1%
5	www.nice.org.uk Internet Source	1%

6	www.uchospitals.edu Internet Source	<1 %
7	www.iarc.fr Internet Source	<1 %
8	www.ecogen.com Internet Source	<1 %
9	www.ncbi.nlm.nih.gov Internet Source	<1 %
10	www.pubmedcentral.nih.gov Internet Source	<1 %
11	Sharma, Neha, Amrita Singh, Navneet Singh, Digamber Behera, and Siddharth Sharma. "Genetic polymorphisms in GSTM1, GSTT1 and GSTP1 genes and risk of lung cancer in a North Indian population", Cancer Epidemiology, 2015. Publication	<1 %
12	edoc.ub.uni-muenchen.de Internet Source	<1 %
13	www.coursehero.com Internet Source	<1 %
14	www.health.gov.au Internet Source	<1 %
15	atc.wustl.edu Internet Source	<1 %

16	preview-jhonline.biomedcentral.com Internet Source	<1 %
17	www.sinobiological.com Internet Source	<1 %
18	discovery.ucl.ac.uk Internet Source	<1 %
19	www.ashg.org Internet Source	<1 %
20	www.nationaldiagnostics.com Internet Source	<1 %
21	www.monzir-pal.net Internet Source	<1 %
22	www.benegene.com.cn Internet Source	<1 %
23	www.mdpi.com Internet Source	<1 %
24	scripties.umcg.eldoc.ub.rug.nl Internet Source	<1 %
25	www.zymoresearch.com Internet Source	<1 %
26	respiratory-research.com Internet Source	<1 %
27	Hauser, Russ Chen, Zuying Pothier, Lucil. "The	

relationship between human semen parameters and environmental exposure to polychlorinated biphen", Environmental Health Perspectives, Sept 2003 Issue

Publication

<1 %

28

www.science.gov

Internet Source

<1 %

29

www.h4c.net

Internet Source

<1 %

30

www.funakoshi.co.jp

Internet Source

<1 %

31

www.jica-ngo.ph

Internet Source

<1 %

32

Zolezzi Moraga, Juan Manuel <1979>(Hrelia, Patrizia). "Cellular and Molecular Biomarkers for Risk Assessment of Colorectal cancer", Alma Mater Studiorum - Università di Bologna, 2012.

Publication

<1 %

33

clinicaltrials.gov

Internet Source

<1 %

34

www.hrpub.org

Internet Source

<1 %

35

Chauhan, Anuradha(and Höfer, Thomas). "Cell cycle", Mathematisch-Naturwissenschaftliche

<1 %

36	www.jofamericanscience.org Internet Source	<1 %
37	Sequaris, Gilles Frédéric Brice. "Molecular mechanism of Indian Hedgehog signalling in human sebocytes", Kölner UniversitätsPublikationsServer, 2010. Publication	<1 %
38	www.doria.fi Internet Source	<1 %
39	minerva-access.unimelb.edu.au Internet Source	<1 %
40	hmg.oxfordjournals.org Internet Source	<1 %
41	scielo.isciii.es Internet Source	<1 %
42	brage.bibsys.no Internet Source	<1 %
43	mtcancer.org Internet Source	<1 %
44	www.faqs.org Internet Source	<1 %
45	Nanri, A. Mizoue, T. Matsushita, Y. Sasa.	<1 %

"Serum folate and homocysteine and depressive symptoms among Japanese men and women.(ORIGINAL ARTICLE", European Journal of Clinical Nutrition, March 2010 Issue

Publication

46

espace.library.uq.edu.au

Internet Source

<1 %

47

www.amoydx.com

Internet Source

<1 %

48

herkules oulu.fi

Internet Source

<1 %

49

www.cancerfacts.com

Internet Source

<1 %

50

www.articleshub.org

Internet Source

<1 %

51

www.jthoracdis.com

Internet Source

<1 %

52

ctep.info.nih.gov

Internet Source

<1 %

53

Brichta, Lars. "Molecular genetic investigations of histone deacetylase inhibitors as potential neurotherapeutics for autosomal recessive proximal spinal muscular atrophy (SMA)", Kölner UniversitätsPublikationsServer, 2011.

Publication

<1 %

54	www.sigmaaldrich.com Internet Source	<1 %
55	highlandcountypress.com Internet Source	<1 %
56	rosemonteis.us Internet Source	<1 %
57	www.fda.gov Internet Source	<1 %
58	220.247.224.39 Internet Source	<1 %
59	www.apjcpcontrol.org Internet Source	<1 %
60	pathkids.com Internet Source	<1 %
61	donnachac.hubpages.com Internet Source	<1 %
62	vinci.inesc.pt Internet Source	<1 %
63	intl.psychosomaticmedicine.org Internet Source	<1 %
64	www.tribuneindia.com Internet Source	<1 %
65	www.zymoresearch.de Internet Source	

<1 %

66

ijmeg.org

Internet Source

<1 %

67

www.medworm.com

Internet Source

<1 %

68

Silvestri, Gerard A., and James R. Jett. "Clinical Aspects of Lung Cancer", Murray and Nadel's Textbook of Respiratory Medicine, 2010.

Publication

<1 %

Exclude quotes On

Exclude matches < 10 words

Exclude bibliography On