Association of *Aryl Hydrocarbon Receptor* gene polymorphism with the risk of Lung Cancer in North Indian population

A

Thesis submitted

In partial fulfillment of the requirement of the degree

Of

Master of Science

In

Biotechnology

By

Apurva Bhardwaj

Roll no. 301201002

Under the supervision of

Dr. Siddharth Sharma

Assistant Professor

Department of Biotechnology

Thapar University

July, 2014

DEPARTMENT OF BIOTECHNOLOGY

THAPAR UNIVERSITY PATIALA,

PUNJAB

CERTIFICATE

This is to certify that the dissertation entitled "**Association of *Aryl Hydrocarbon Receptor* gene polymorphism with the risk of Lung Cancer in North Indian population**" being submitted by **Ms Apurva Bhardwaj** in partial fulfilment for the requirement of degree of **Masters of Sciences** in **Biotechnology** in the **Department of Biotechnology, Thapar University, Patiala** is a bonafide work carried out under the esteemed supervision and conception of **Dr Siddharth Sharma**, Assistant Professor , Department of Biotechnology, Thapar University, Patiala.

15/7/2014

Dr Siddharth Sharma
Assistant Professor
Department of Biotechnology
Thapar University , Patiala

(countersigned by)

Dr Dinesh Goyal
Professor & Head
Department of Biotechnology
Thapar University, Patiala

Dr S K Mohapatra
Dean of Academic Affairs
Thapar University, Patiala

DECLARATION

I, the undersigned, hereby declare that the research work presented in the M. Sc dissertation entitled ***"Association of Aryl Hydrocarbon Receptor gene polymorphism with the risk of Lung Cancer in North Indian population."*** has been carried out by me under the supervision and guidance of ***Dr. Siddharth Sharma***, Department of Biotechnology, Thapar University, Patiala. Further, I declare that no part of this dissertation has been submitted for a degree or any other qualification of any other university or examining body in India/elsewhere.

Apurva Bhardwaj

(Reg. No. 301201002)

ACKNOWLEDGEMENT

Thoughts push us to idealise, which then leads to Observation, and finally a validation to what one observes and proves gives birth to Science.. and this path can never be traversed alone by a novice no matter easy or tough. Though I have yet not trotted it the hard way, with all my gratuitous expressions of gratitude in the most raw but yet polished manner, my sincerest regards and thanks to my Advisor, ***Dr Siddharth Sharma***, one with exemplary passion towards research, with his immaculate and sharp deadlines, he exposed me to the truth of meeting ends and excelling. I faltered and I had him to my back to rectify, reconcile and reward.

I place on records my sincere thanks to ***Dr Dinesh Goyal*** , Professor and Head, Department of Biotechnology, Thapar University, for he made sure that the infrastructure suerely should suffice with our requirements

For a mansion, the pillars need to be the strongest and she personified to be one in my life, deepest of my regards and sincerest of thanks to ***Charu Bahl ,*** who truely is a substance of intelligence and genius, a help that is selfless and honest topped with a dose of extended release of encouragement and never to lose. Thanks to ***Amrita Singh***, for helping me bear the grunt of failures and falls. A warm and enriched thanks to all my friends who stood by me at the time of high and low tides. It shall cost me too much of space quoting all of them so under the broad head of ***Friends*** I wanna thank you all.

They made me what i am today and groomed me with this aura of positivity and safety around me. Ofcourse without Her incessant prayers and love and His pat on the back which gave me the strength to move on without expecting for rewards, this manuscript would not been possible .Though they cant be thanked for the long list but thank you ***Maa and Paa*** for being there always.

Almighty, the above mentioned list of acknowledgements will fall short provided you would not have paved a way for them to assist and support me. Patience, perseverance, consistency and dedication is what this dissertations journey instilled in me and I owe it all to all the above mentioned.

ABBREVIATIONS

ADCC	Adenocarcinoma
SQCC	Squamous cell carcinoma
SCLC	Small Cell Carcinoma
AhR	Aryl Hydrocarbon Receptor
C.I.	Confidence interval
CYP	Cytochrome P450
LC	Lung Cancer
PCB	Pentachlorobiphenyls
PAH	Polycyclic aromatic hydrocarbons
SNP	Single Nucleotide Polymorphism
HAH	Halogenated Aromatic Hydrocarbons
TCDD	2,3,7,8-tetrachlorodibenzo-*p*-dioxin
BaP	Benzo-*a*-pyrene
PCDF	Polychlorinated dibenzofurans
IAA	Indole-3-acetic acid
ICZ	Indole [3,2-b-] Carbazole
I3C	Indole-3-carbinol
DRE	Dioxin Response Elements
UGT	Uridine diphosphate glycosyltransferase
GST	Glutathione *S*-transferase
XAP 2	Hepatitis B Virus Associated Protein 2
HSP	Heat Shock Protein

INDEX

LIST OF TABLES

LIST OF FIGURE

ABSTRACT

Identification of host differences in the susceptibility to carcinogenesis, especially to cigarette smoking is essential in predicting if one is at a higher risk of acquiring the disease. AhR, an orphan receptor, is a key regulator for the transcription of metabolizers like *CYP* and the one playing as a housekeeper by maintaining homeostasis, is one of the most widely researched receptor.

The present study was focused to assess the vulnerability of an individual towards Lung cancer having a polymorphism in the exon 10 (rs2066853) and two polymorphism in the intronic region (rs10252882) and (rs 2282885) of AhR, which is a well documented and researched gene. A case control study was performed on the North Indian population. Odds ratio and Class intervals were used to assess the potency of the association. p values were also calculated to find out the statistical significance of the analysis.

Cumulative cigarette smoking was interrogated to be a lead player along with other parameters responsible for the susceptibility towards Lung Cancer. The combinatorial effect of genotype and histological forms with combined effect of polymorphism was also studied and significant associations were found to exist.

The present study provided us with wider platform to study the *rs10252882*, *rs2066853* and *rs228288* polymorphism which are not well studied yet and Statistically significant association was seen amongst the genetic variant of *Ahr* and were found to have a great risk towards Lung Cancer. Tobacco smoking was affirmed as congener along with other environmental contaminants and showed high ORs and therefore a strong association with lung cancer risk. Cumulative effects of polymorphism amongst the variants were also shown to be highly associated with the risk towards Lung cancer. *AhR rs10252882 a*nd

rs2282885 were found to be maximally associated with the risk for Lung cancer.

APPENDIX 1

1. **0.5M EDTA**: Dissolved 9.306g of disodium salt of EDTA in 20ml of deionsed water, and then adjusted the pH to 8.0 by 1M sodium hydroxide. Sterilized the solution by autoclaving.
2. **10% SDS**: Dissolved 1g of SDS in 10ml of deionsed water.
3. **100mM Tris-Cl (pH 8.0**): Dissolved 0.32g of Tris-Cl in 10ml of deionsed 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 ºC overnight.
6. **5% DMSO**: Mixed 50ml of 100% DMSO in 50ml of deionised sterile water. Sterilize the solution by autoclaving and stored at -20 ºC.
7. **5M Sodium chloride (NaCl)**: Dissolved 5.85g of sodium chloride in 20ml of deionised water. Sterilize the solution by autoclaving.
8. **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.
9. **6X loading dye**: Mixed 0.050g of bromophenol blue, 0.050g of xylene cyanol and 8g of sucrose. Mixed it in a small amount of TE buffer , then made up the volume upto 20ml by adding more TE buffer.
10. **Ethidium Bromide (10mg/ml)**: Dissolved 1g of ethidium bromide in 100ml of water. Mixed the solution properly.
11. **Magnesium chloride (MgCl2) (100mM)**: Dissolved 0.41g of MgCl2 in 20ml of deionised water and sterilized by autoclaving.
12. **Sucrose (1M)**: Dissolved 3.41g of sucrose in 10ml of deionised water and sterilized by autoclaving.
13. **TE buffer (pH 8.0)**: Added 1ml of 100mM Tris-Cl (pH 8.0) and 200µl of 0.5M EDTA solution to 8.8ml of deionised water. Sterilized the solution by autoclaving.
14. **Triton X- 100 (10%)**: 100µl of TritonX-100 mixed with 900µl of deionised water and mixed properly.

Chapter 1

Introduction

Introduction

Epidemiological studies have contributed incessantly to the growing awareness of the importance of the genetic and acquired susceptibility factors which are governed to modulate the risk associated with the encounter with the environmental contaminants. Knowledge, prevalence, distribution and understanding of these environmental contaminants in the population wherein an individual's vulnerability, which is predisposed in the form of his genetic makeup, differs and accounts for the risk of acquiring endangerments like Cancer. A devotion towards understanding the mechanism of toxicity in order to assess the risk and mitigate the harm deemed to occur, is taken care of under the rubric of Molecular Epidemiology while taking into consideration the sources of systematic errors and random errors (confounding and bias).

Cancer is a multifactorial disease that results from the complicated interactions between the environment and genetic factors. It has been seen that 80-90% of Lung Cancer incidence can be attributed to cigarette smoking but only about 10% acquire Lung Cancer. This clearly demonstrates the role of susceptibility parameters in the host. The WHO estimates that in 2030 the number of deaths attributable to the consumption of tobacco will be 100 million accompanied by an increased incidence of Lung Cancer. Differential susceptibilities amongst individuals are accounted by the complexity of exposure to carcinogen, number of multiple alleles present for an enzyme to be encoded and most importantly the predisposed genes for the xenobiotic metabolism in the individual. Specific alterations in the genome or an abnormality in the pathways for detoxification and metabolism of environmental contaminants, glues together to form one of the important etiological factors for Lung Cancer.

Aryl Hydrocarbon Receptor (AHR) serves as an archetype for the toxicity mediated response studies. *AhR*, a ubiquitously expressed ligand-activated transcription factor, upon its encounter with the foreign ligands activates the transcriptional machinery of genes encoding for biotransformation enzymes like CYP1A1 and hence mediate the metabolism of PAH s and nitrosamines which account for the maximally found carcinogen in the cigarette smoke. Dioxin Receptor ,the other name to *AhR* , shows a set of vivid and diverse consequences on

its exposure to carcinogens which may be due to structural difference in *AhR* gene or in the *AhR* repressor or in the *AhR* nuclear translocator (ARNT).

Cigarette smoke contains thousands of chemicals of which fifty are known carcinogens, whereas others are present in the form of procarcinogens, which undergo enzymatic conversions to form potent reactive carcinogens. Cigarette smoking induces genotoxicity which further traverses the path towards carcinogenesis and eventually causes Lung Cancer. A contradiction is supported by the presence of various endogenous ligands in the human body causing toxicity within the system. Therefore, synergistically genotoxicity and non genotoxicity play an important deciding factor in the advent of acquiring Lung Cancer.

Polymorphism in specific genes can modulate the formation of DNA carcinogen adduct and favour the generation of mutations which lead to Lung Cancer especially in Smokers. The Polymorphic variants of *AhR* play a significant role and are held responsible for disposing the individuals with greater chances of acquiring Lung Cancer. Intronic polymorphism over the time is being authenticated as biomarkers for Cancer and *AhR* stands at par to support the statement.

Therefore, to trace the role of these polymorphic variants in AhR in the susceptibility towards Lung cancer in North Indian population. We evaluated the association of one exonic SNP ***AhR Arg^{554}Lys rs2066853*** and two intronic region SNP ***rs10252882*** and ***rs2282885*** with the risk of development of Lung cancer in the North Indian population.

Chapter 2
Review of Literature

REVIEW OF LITERATURE

2.1 AhR : Introduction and structure

Aryl Hydrocarbon Receptor (*AhR*) is a cytosolic, ligand inducible, transcription factor regulating a myriad of genes. It plays a central role in the induction of xenobiotic - metabolizing enzymes and hence in xenobiotic detoxification by the process of biotransformation. The receptor is encoded by the gene *ahr* which resides on the short arm of chromosome 7(7q). The *AhR* gene is 50 kb in size and encompasses 12 exons and 10 introns. The *AhR* gene encodes for the enzyme Aryl Hydrocarbon hydroxylase. The receptor is seen to mediate a variety of toxic and carcinogenic effects of a wide variety of environmental contaminants along with other xenobiotics (Barouki *et al*., 2007). Activation of the *AhR* by these ligand leads to inappropriate modulation of gene expression or the disruption of other cellular signalling pathways or enzyme-induced generation of genotoxic intermediates.

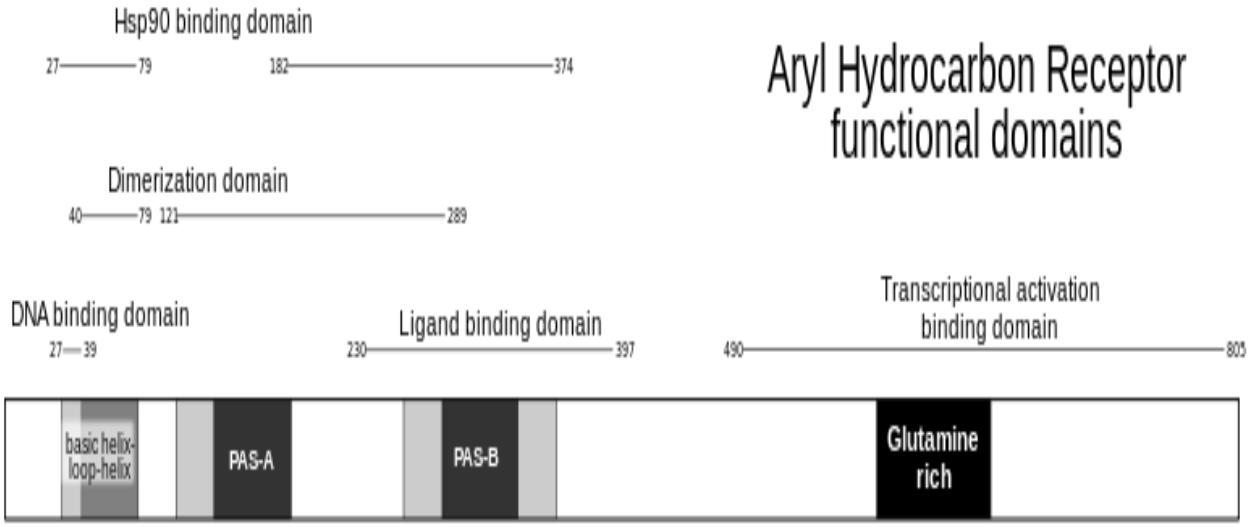

Figure 2.1: Aryl Hydrocarbon Receptor and its functional domains

(Schmidt.et al, 1996)

2.2 AhR and its Proteins

AhR belongs to the family of basic Helix Loop Helix (b HLH) Period –ARNT-Single Minded Protein (Per-ARNT-Sim) family. The three proteins (Per-ARNT-Sim) correspond to the proteins respectively found in Drosophila and Humans. The carboxy terminal PAS domain , houses the transcription activation /repressor Domain. Pas extends to a length of 260-310 amino acids usually and is categorically divided into two conserved regions, Pas (A) and Pas (B), separated by poorly conserved sequence of amino acids. The Pas (A) and (B) contain a copy of 44 amino acids repeat, uniquely characterized by Phenylalanine, followed by a string of variable amino acids terminating in Histidine and Aspartic acid. The Pas domain is responsible for binding of proteins to molecules. (Carver *et.al*., 1998). The Pas domain is also seen to be distributed amongst proteins essentially playing a role in the maintaining of circadian rhythm. The protein DNA binding is mediated by b HLH region of the b HLH –Pas Proteins. The b HLH function as sequence specific transcriptional regulators. These regulators harbour subdomains which play the role in DNA binding along the basic region and the Helix Loop Helix region assists in the dimerization (Meyer *et al.*,1998). One component of this heterodimer of b HLH –Pas is inducible and is expressed in specific tissues whereas the other component is expressed ubiquitously. Studies have clearly showed that nuclear localization and nuclear export signals required for the process of translocation resides in b HLH domain. This family of proteins play an essential role in sensing the environmental factors such as oxygen and light (Hooper *et al*., 2011)

2.3 ***AhR : Unliganded Cytosolic Complex***

The *AhR* complex is shelled by various proteins belonging to different classes and deemed to perform different roles. The complex forms a hetero tetrameric complex comprising of XAP 2 associated with p23 and two molecules of Heat Shock Protein 90(HSP 90) in the cytosol. (Petrulis *et al.,* 2002)

The most important amongst all is the XAP 2 (Hepatitis B Virus X Associated Protein 2). It is an immunophilin like protein, possessing a tetracopeptide repeat domain. It plays an important role in establishing the interaction between HSP 90 and AHR.(Carver *et al*.,1998) It acts as a leisch of the horse, providing protein stability along with the interaction of the complex with p23. XAP acts as an enhancer and repressor both for the *AhR* transcriptional activity.(Meyer *et al.,* 1998) (Ramadoss *et al.*,2005)

Chaperone protein HSP 90 plays a major role in the ligand binding fidelity of a number of steroid receptors. It binds to *AhR* and helps in maintaining the efficient conformation and helps to maintain the complex in its unliganded constitution in the cytosol.

p23 is a co chaperone to HSP 90,in the presence of ATP is stably bound to HSP 90 monomer and thereby interacts with *AhR* . Essentiality of p23 in the *AhR* complex is yet to be studied and validated. (Perdew *et al*., 1998)

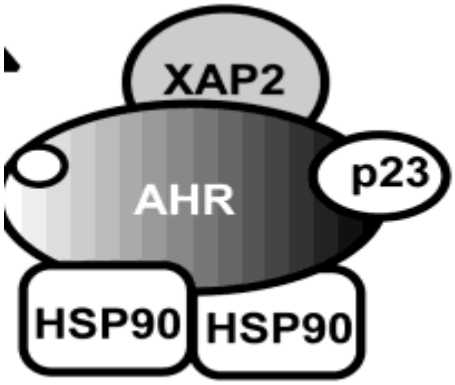

Figure 2.2: Ahr: Unliganded Cytosolic complex
AhR, ligand activated receptor resides in the cytosol in the form of a complex comprising of XAP2, two molecules of Heat Shock Proteins 90 along with co chaperone p23.

2.4 AhR –ARNT Heterodimerization and Translocation

The receptor lies in a dormant state in the cytosol complexed with a heterogenous group of proteins which assist the receptor in maintaining the integrity of its unliganded form and also keep the receptor masked until it is recognized by the attacking foreign ligand. The participating molecules for the unliganded complex play an essential role in preventing the nuclear translocation and ARNT dimerization with *AhR*. (Meyer *et al*., 1999)

Upon its encounter with the ligand, which usually is hydrophobic in nature and enters the cell by diffusion, an association is established leading to conformational changes. These changes lead to an increase in the affinity between the binding molecules and a dip in the rate of dissociation from the ligand. The *AhR*–ligand complex now translocates to the nucleus. It is here that it complexes with Aryl Hydrocarbon Receptor Nuclear Translocator (ARNT) found in the nucleus to form a *AhR-ARNT* heterodimer, which further recognizes a 6 nucleotide long consensus sequence (TNGCGTG) called DRE (Dioxin Responsive Elements) on the DNA in the regulatory domains of the gene coding for phase I and phase II drug

metabolizing enzymes and activates the transcription of these genes (Hankinson *et al., 1*995). DREs are defined as specific sequence present on an element which acts in response to a stimulus or a condition .*ARNT* recognizes the second half of the sequence (GTG) and *AhR* recognizes the (TNGC) half site of the sequence. After binding to DREs at the promotor/enhancer regions of *AhR* responsive genes, DNA bound *ARNT –AhR* heterodimer recruits various proteins for the basic transcriptional machinery. This paves the way to the activation of a number of genes involved in the Phase I and II metabolism ,e.g. *CYP1A1, GST,CYP1B1,*etc (Kazlauskas *et al*., 1999)

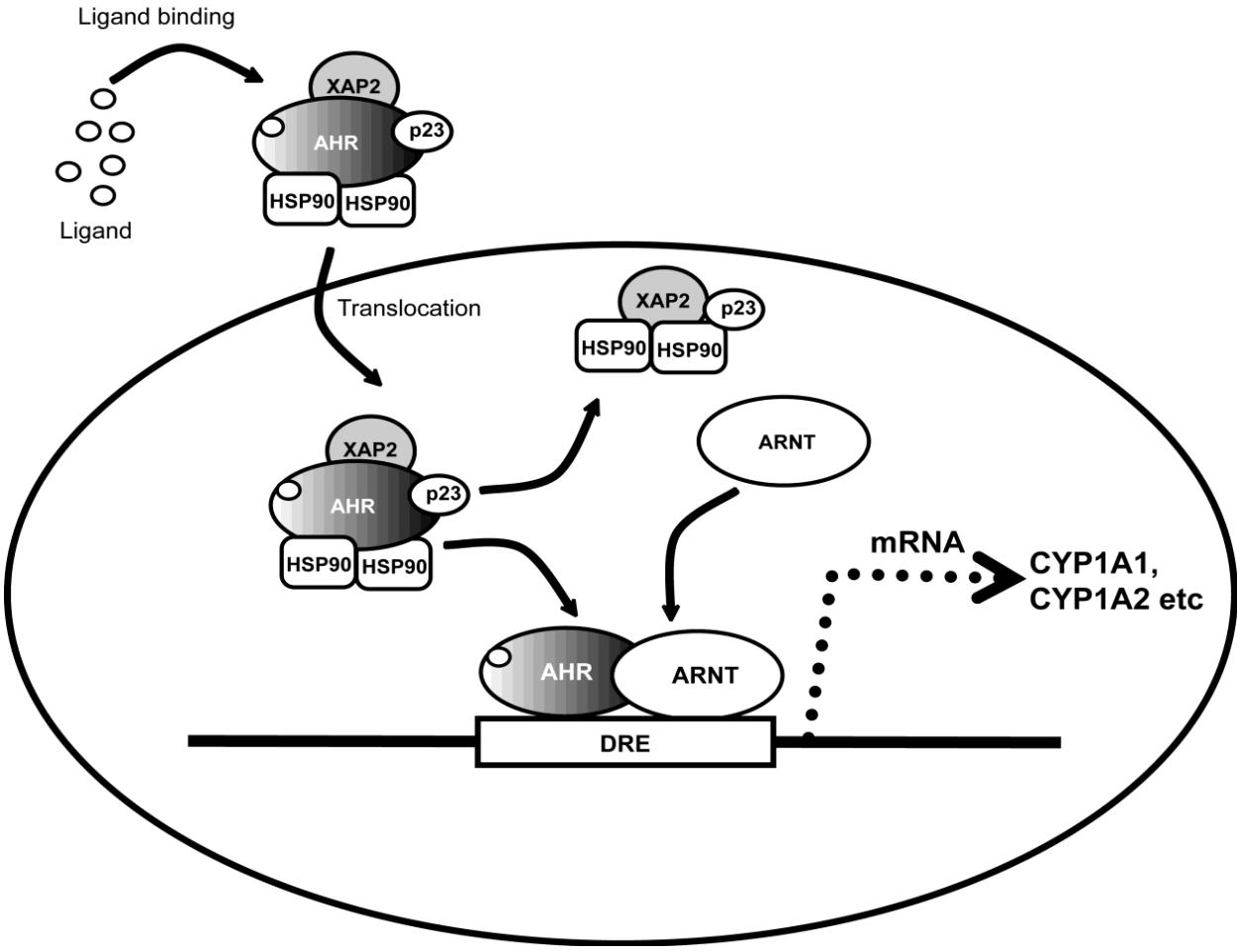

Figure 2.3: AhR Ligand Activation: Ligand activation of the AhR promotes receptor nuclear translocation and eventual heterodimerization with the ARNT. The ARNT/AhR heterodimer then binds to dioxin responsive elements (DREs) leading to AhR –target gene activation (Miller *et al.*, 2011)

2.5 Physiological Significance of AhR

The *AhR* has been seen to be expressed vividly in various tissues and organs in the human body. Studies have reflected that *AhR* has a promising role in maintaining the cells in harmony with the environment to which they are being exposed. Research continues to authenticate and support the indispensible roles of *AhR* in the various organ systems in the human body.

AhR is responsible for modulating diverse molecular pathways in concert with the induction of Phase I and II genes for detoxification.

Studies show that *AhR* function in direct and indirect modulation of transcription process by combining with other transcription factors, coactivators, or repressors and by altering signalling transductions cascades (Tan *et al*., 2002, 2004)

Li *et al.,* (2008) provide fascinating new insights into the physiological role by showing it to be essential for normal intestinal immune function. *AhR* gets activated by compounds like Indole 3 carbazol (I3C) present in the cruciferous vegetables (like cauliflower, cabbage, broccoli). When Indole 3 carbazol comes in contact with the gastric acid it gets converted to high affinity *AhR* ligands like (Indole-3,2-Carbazol) ICZ. It is now that *AhR* becomes activated in the cytosol and complexes with ARNT to form a heterodimer and thereafter translocates to the nucleus, eventually leading to carcinogenesis. (Hooper *et al.,* 2008)

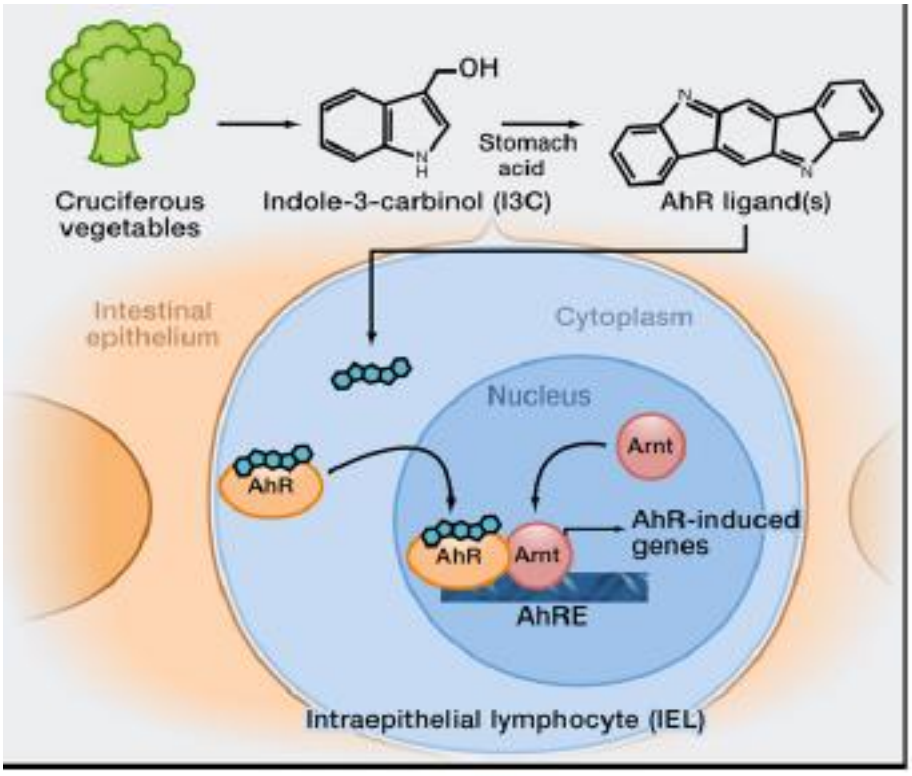

Fig 2.4: Dietary compounds regulate intestinal immunity through the Aryl Hydrocarbon Receptor (Hooper *et al*., 2008)

2.5.1 <u>Immune Response and AhR</u>: Immune Response is sharply governed by the intake of Xenobiotics compounds like PAH s and HAH s , causing thymic atropy and impairment of the immune system in mice.(Kimbrough et al 2003) It is also seen to have an enhanced incidence of lymphoma and leukemia formation in humans.(Silkworth *et al*.,1995)

TCDD has very prominent deleterious effects on the Maturation of bone marrow B lymphocytes in vivo (Thurmond *et al*., 2000) This is seen to occur due to the reduction of B lymphocytes precursor cells. TCDD has also been seen to wreck the expression of B lymphocyte-induced maturation protein which further inhibits and puts a halt to the differentiation of B lymphocytes and hence humoral response system.(Schneider *et al*., 2009) TCDD, a potent human carcinogen, is seen to perturb the T –cell IL -5 production and TH2 mediated immune response. *AhR* activation by TCDD also plays with the ability of the hematopoietic stem cells to respond to signals present within the microenvironment. It is seen that TCDD also affects the adaptive immune response of the body. (Fujimaki *et al*., 2002) Leukocyte activation promotes the endogenous expression of *AhR* and the translocation from the cytosol.

2.5.2 <u>Reproductive system and AhR</u> : Workers exposed to trichlorophenoxyacetic acid reported sexual dysfunction and reduced libido (Moses *et al.*,1998) . It was also seen that humans when exposed to TCDD displayed significant skewed offspring sex ratios. Studies also suggested that the exposure to dioxins disrupted the sexual differentiation and neuroendocrine functions in humans. Upon exposure to potent *AhR* ligands no definitive and marked reproductive consequence was observed. Dioxins have been found to obstruct sexual differentiation and neuroendocrine system in the animal models. (Mocarelli *et al*., 2000). Perturbed spermatogenesis was reported in males working in coke emission industries leading to male infertility (Izawa *et al.*, 2007)

2.5.3 <u>AhR and Organ development</u>: *AhR* holds an important place in the development of organ systems like liver, kidney, heart and skin. In a *AhR* knockout mice, irregular organ development was observed. It also extends a hand in the cardiovacular formation and angiogenesis.(Thackaberry *et al*.,2002). TCDD activated *AhR* led to irregular renal development and defective cardiovascular development. It is vividly expressed in the liver therefore plays a very important role in the liver development (Bunger *et al*., 2008)

The constituent proteins of *AhR* are found to play an important role in diverse physiological processes such as Circadian Rhythm, Organ Development, Neurogenesis, Metabolism and stress response to hypoxia.

The *AhR* is found to be a player in various other pathways involving the steroid receptors such as the oestrogen receptor and androgen receptor. In the former it downregulates ER mediated effects following various mechanisms whereas in case of the latter it is seen to inhibit the androgen mediated cell proliferation of LNCaP cells through protein phosphorylation modulations (Morrow *et al*., 2004)

2.6 *<u>AhR and its Ligands</u>*

AhR has often been designated as an orphan receptor for it becomes bioactivated only upon its encounter with its corresponding Ligands. *AhR* has a promiscuous binding site and therefore is able to bind a diverse array of ligands, classified as, Endobiotic ligands , Xenobiotic ligands and Atypical ligands

2.6.1 <u>Xenobiotic Ligands</u> encompasses a wide range of carcinogenic substances like PAHs , planar /aromatic hydrocarbons and halogenated hydrocarbons .(Denison *et.al*., 2003)

HAH (Halogenated Aromatic Hydrocarbon**)** form the most potent class of ligands and are capable of getting activated at a low concentration. They accumulate in the environment due to their chemical stability and hydrophobicity. This class includes PCBs (Pentachlorobiphenyls) and PCDFs.

TCDD (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) is another environmental contaminant which is formed during the process of pesticide and herbicide synthesis , burning of wood, burning of oil , etc. It has been classed as a Human Carcinogen by the IARC lately.

PAH s account for the maximum proportion in the cigarette smoke. It is also found in char broiled meat and vehicle exhaust .They have deleterious effect on lungs, skin and cause inflammatory disorders like Asthma. E.g. BaP (Skupinska *et al.,* 2004) (Gaulaunic *et al*., 2008)

BaP (Benzo –a-pyrene) is found to be bioactivated by *AhR* , by the enzymic conversion of BaP to a genotoxic intermediate, which causes the DNA adduct formation , promoting tumorigenesis.(Shimizu *et al.*,2000)

PCBs are synthetic organochlorine compounds with high thermal and chemical stability which helps they to persist in the environment. Surprisingly the main source of exposure in humans is through dietary consumption of fish. Studies show that PCB s cause huge amount of neurological and endocrine disruptions (Boucher *et al*.,2009)

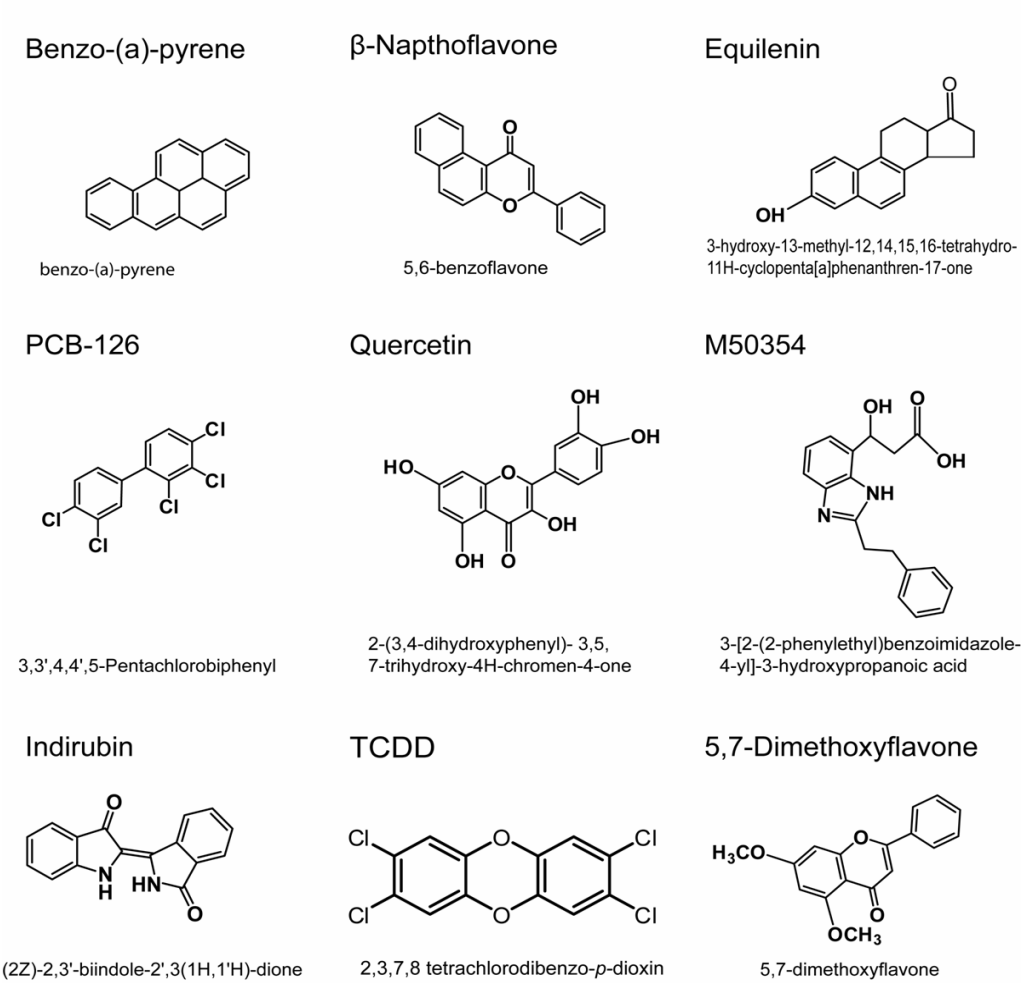

Fig 2.5: 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), 3,3`,4,4`,5-Pentachlorobiphenyl (PCB-126) and benzo-[*a*]-pyrene are environmental contaminants that cause negative health effects. (2Z)-2-3`-biindole-2`,3(1H,1`H)-dione indirubin is a plant tryptophan metabolite. 5,7-dimethoxyflavone and 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-4H-chromen-4-one (quercetin), 5,6-benzoflavone (β-NF) are synthetic and naturally occurring flavonoid compounds. 3-[2-(2-phenylethyl) benzoimidazole-4yl]-3-hydroxypropanoic acid (M50354) was a putative anti-cancer drug candidate. (Flaveny *et al.*, 2008*)*

2.6.2 <u>Endobiotic AhR Ligands</u>

The Endobiotic ligands for *AhR* are compounds which are synthesized endogenously and have been seen to activate the transcriptional machinery of the responsive genes encoding for the enzymes participating in the Detoxification process.

A splendid example is that of tetrapyroles like bilirubin and biliverdin , where the prior is a byproduct of hemoglobin in our body. It has a relatively low affinity for the *AhR* and under normal physiological conditions is not present in concentrations required to illicit *AhR* activation. (Bjeldanes *et.al.*, 1991). Studies show that they have a high affinity for *AhR* in yeast system but the same fails for mammalian system. In mammals the affinity for TCDD is found to be the highest.(Perdew *et al*., 2005)

Other examples are: Indole containing compounds e.g. IC3, IAA, Arachidonic Acid metabolites, Tryptophan photoproducts and metabolites.

Plant tryptophan metabolites like Indigo and Indirubin act as potent *AhR* activators. (Denison *et al.*, 2003)

Crucifers like broccoli, cauliflower, etc have been seen to play a significant role in conferring normal intestinal immune function. *AhR* is activated by compounds like I3C which gets converted to high affinity *AhR* ligands (ICZ) upon exposure to stomach acid (Hooper, 2011)

2.6.3 Atypical AhR Ligands

This category comprises of some unusual and atypical compounds which also act like ligands to *AhR* but donot go and bind to the receptor at its binding site or pocketted active site but yet manages to induce the activation of various genes and consequently encode for the enzymes responsible for various activities in the body. An example to support this is that of the CYP genes. Benezimidazole derivatives, Myriscitin and omeprazole are also the ones amongst the list. (Quattrochi *et al*., 1993)

Omeprazole is found to be a proton pump inhibitor. It has been used to cure peptic ulcers and heartburns. Studies reflect that omeprazole induced *AhR* activity was detected in human hepatoma cells along with human hepatocytes (Daujat *et al*., 1997). Literature also shows that there is a common critical ligand binding site for both TCDD and Omeprazole. Invivo the proton pump inhibitor is inhibited by tyrosine kinases, hence can be said that omeprazole induced AhR activity can be witnessed in cases where there is a faltered transduction pathway. (Murray *et al*., 2008)

2.7AhR As A Regulator

2.7.1 *Phase I Enzymes:*

In the first phase of Biotransformation, the polarity of the xenobiotic is increased by the addition reactions which help in the addition of water molecules to the target compounds. In this set of reactions the target molecule becomes water soluble and now surpasses to the second phase of the detoxification reactions.

In humans, cytochrome P450 is a family of metabolising enzymes that are predominantly found in the endoplasmic reticulum and lesser in the mitochondria. These enzymes metabolize a wide variety of substrates ranging from toxicants to steroids. A total number of fifty seven P450 s have been discovered in humans and are significantly being employed for the metabolism of xenobiotics. The most important amongst all the cytochrome isoforms are the CY3A4, CYP 1A1, CYP1A2, etc. expressed exclusively in extrahepatic tissues. Studies reveal that CYP1A1 is the only cytochrome which is directly dependent on *AhR* for its regulation. CYP1A1 is primarily an extrahepatic tissue which is potent enough to metabolize xenobiotics and bioactivate many carcinogens upon encounter.

Of all the CYP enzymes, CYP1A2 (hepatic enzyme) is the only one which is directly regulated by the AhR (Ding X *et al.,* 2004). Literature puts it across records that CYP1A2 is capable of metabolising arylamines and heterocyclins along with showcasing an enhanced gene expression on exposure to cigarette smoke. The isoform is capable of reducing caffeine to caffeine intermediates and are now being used as marker for assessing the activity of CYP1A2(Nakajima *et al*., 1999).

2.7.2 ***Phase II Enzymes*:**

In the second phase of biotransformation reactions, a set of chemical reactions like conjugation, oxidation and reduction take place and detoxification is carried out. Examples for this category are supported by UGT1A1, UGT1A6, GST and NaDPH-quinone – oxidoreductase and Aldehyde dehydrogenases.

GSTs act as the most potent and effective metabolizer for humans. GSTs are predominantly involved in the metabolism of xenobiotics, and the few drugs that are metabolised by GSTs are cytotoxic drugs used in chemotherapy (Daly *et al*., 2003)

Uridine diphosphate glucuronosyl transferase (UGT1A1) expression was first shown to be activated by TCDD and PCBs. The enhancer for UGT1A1 is solely responsible for the

inducing *AhR* therefore, independent of a ligand, as it houses a DRE in its enhancer region. In humans an increase in bilirubin levels lead to elevated binding with *AhR* hence an increase in the levels of UGT1A1.(Yueh *et al.*, 2003)

Aldehdye dehydrogenase (ALDH) is commonly known for the conversion of aldehydes and lipid and carbohydrate metabolism to carboxylic acids. Studies show that both induced and constitutive expression of ALDH is mediated by DREs. (Boesch *et al*., 1999)

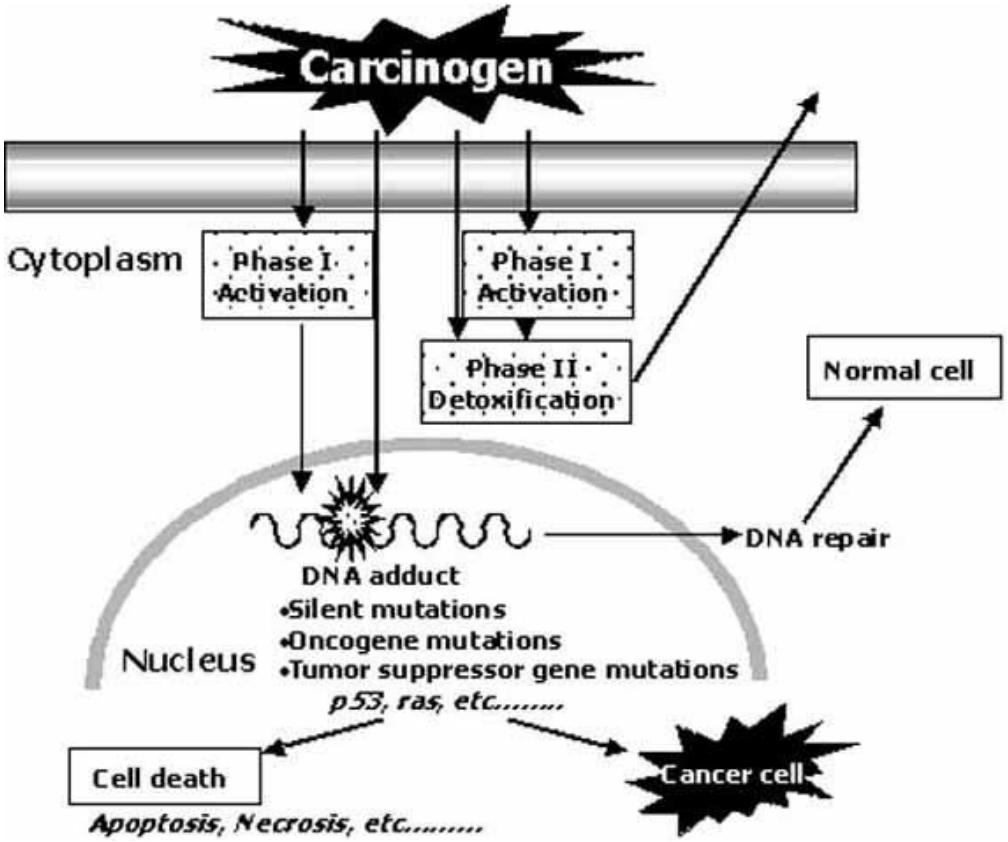

Figure 2.6: Metabolic pathways of carcinogen (Kawamoto *et al*., 2004)

2.8 AhR Polymorphism and Cancer

Genetic polymorphism are common variants in the genetic code. It is typically elucidated as a sample of interest comprising of a polymorphism seen in atleast 1% of the population. In this approach we assume that the low or medium penetrant genes account for the majority of the lung cancer susceptibility rather than those which are highly penetrating. A single nucleotide polymorphism (SNP) is an alteration in a single nucleotide present in the DNA sequence.

The molecular basis for enzyme polymorphism is backed by the occurrence of either of the three cases, ie.

- Either we will see a nucleotide variation in the coding region of a gene therefore altering the enzyme activity due to the amino acid substitution.
- There might be a deletion in the coding region of the gene leading to an inactive enzyme or even lack of protein synthesis.
- Variation in the polyadenylation site causing post transcriptional repression hence affecting the quantity of the transcript formed. (Boffeta 2000)

Different approaches and models are recruited in order to study the polymorphism in a particular gene. Amongst them are the Candidate gene approach, where the gene and the nucleotide polymorphism both are simultaneously selected and put to statistical and analytical studies. We too have followed the same approach for our study in the North India population.

Despite the high variabilitity in xenobiotic binding affinity displayed by *AhR*, high degree of amino acids conservation was observed. This fact is said to be a consequence to the dynamic and pivotal role of *AhR* in human physiology and homeostasis. (Kawajiri *et al*.,1995)

Most polymorphism cluster around exon 10 which encodes for the Transactivation domain of the receptor. Literature reflects that no polymorphism was observed at codon 678 of the human *AhR*, whereas protein function was adversely affected by polymorphism at codons 554 and 570. Also a 103 fold difference was observed in the CYP1A1 nad CYP2A2 inducibility in response to the TCDD, linked to polymorphism at codon 570.

Polymorphisms prominently seen in the transactivation domains are R554K,P517S and V570I.Amongst all the six variants of Human *AhR*, the R554K , is most widely and commonly investigated polymorphism, located at exon 10, which is a region responsible for the transactivity of the other genes (Harper *et al*.,2002)

$Arg^{554}Lys$ polymorphism is associated with an increased risk of breast cancer in women. (Long *et al.,* 2006)(Ng *et al.,* 2010) (Zhang *et al*., 2011). This SNP is not found to be associated closely with the risk of bladder cancer (Zhang *et al*.,2002) or Lung Cancer (Cauchi *et al*., 2001) (*Chen et al.,*2009*)*

In humans, polymorphisms have also been widely seen at $Lys^{554}Leu$, $Val^{570}Ile$ and $Pro^{571}Ser$, but these are found to be unrelated to the phenotype of TCDD toxicity (Jones *et al*., 2001). Wong *et al* reported an important association when the polymorphisms were studied in

combinations i.e. Lys^{554}Leu, and Val^{570}Ile mutations cumulatively were seen to abrogate CYP1A1 induction reflecting that a single variation in an individual can show an association when studied along with another polymorphism in the same subject. Studies showed that in combination, the individual showed decreased susceptibility of the individual towards the carcinogenic effects of xenobiotics.

Polymorphism at codon 517 was found to be prevalent in African specific population (Harper *et al*.,2001) Studies suggest absence of any association between the *Ahr* polymorphism with receptor ligand binding, dimerization and translocation event in *AhR* signalling, nuclear localization ,export and protein stability along with the recruitment of coactivators and other set of proteins and factors.

Studies also declare a close relationship between *AhR* and mammary gland tumorigenesis. *AhR* gene polymorphisms have been linked to an increased risk of acquiring Lung, Breast and pancreatic cancer (Kim JH, 2007)

In order to assist the science fraternity in exploring the facts, extending our boundaries of knowledge and wisdom and unravel the associative powers possessed by these polymorphism towards the risk of Lung cancer, we chose to study the following three SNPs in *AhR* gene.

AhR rs10252882

This polymorphism is witnessed in the intronic region of the gene where the T nucleotide is substituted by C. This polymorphism is located in the near gene regions towards the 5' end of the intron region. Much literature doesn't stand to support the risk of acquiring cancer with the prevalence of this polymorphism in the population.

AhR rs2282552

This polymorphism is also located in the intronic region of the gene where T nucleotide undergoes a substitution with a C nucleotide.

AhR Arg ^{554} Lys rs2066853

This falls under the head of one of the most widely studied polymorphisms in the AhR gene. Located on exon 10, single substitution occurs by replacing G with A, though it has been observed to a conserved region. The sequence responsible for the coding of Transactivation domains reside in exon 10 and hence plays an important role in increasing or delineating the transactivity of the genes.

AhR along with its heterodimeric partner ARNT ,has been associated with a host's response by which it safeguards itself from the adverse consequence of Carcinogenesis. An alteration in the nucleotide sequence disrupts the transcriptional machinery by dysregulating the gene

encoding for the enzymes to be recruited. Overexpression or less than the desired expression of the gene leads to an stupendous increase in the conversion of procarcinogens into carcinogens. *AhR* outbreeds other receptors by displaying a wide number of binding sites for the ligands ,which in this case are the xenobiotic compounds and contaminants present in the environment, dimerizing with the ligand undergoes translocation towards the nucleus and finally proceeds towards tumorigenesis. The process of carcinogenesis leads to this highly mortal and morbid endangerment called Cancer. The narration of this multistage disease clearly establishes the fact that Smoking cannot be authenticated as biomarker but an etiological factor for Lung cancer and also forms one of the important parameters to be interrogated upon, in being potent and responsible for causing Lung cancer to be assessed

.

Chapter3

Aims and objective

- To evaluate the vulnerability of Lung cancer with the gene polymorphism (rs10252882, rs2066853 and rs2282885) in *AhR* gene
- To find the susceptibility towards Lung cancer on the basis of histological subtypes in association with genetic polymorphism(rs10252882, rs2066853 and rs2282885) in *AhR* gene

- To investigate the role of smoking in acquiring Lung cancer in association with the polymorphism (rs10252882,rs2066853 and rs2282885) in *AhR* gene
- To assess the cumulative effect of different genotypic combinations of the three *AhR* SNPs (rs10252882, rs2066853 and rs2282885) in association with risk towards Lung cancer
- To estimate the cumulative effect of different histological subtypes for the combinations of the three *AhR* SNPs (rs10252882, rs2066853 and rs2282885) in association with risk towards Lung cancer

CHAPTER 4

MATERIALS AND PROTOCOLS

4.1 Study Subjects And Sample Collection

A total of 2O1 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 had been reviewed and approved by the Institute Ethics Committee of PGIMER. Informed written consent was obtained from all patients or their representatives. In brief, eligible cases included those patients who were newly diagnosed with primary Lung Cancer. All the recruited patients were histopathologically diagnosed as having NSCLC and SCLC. There were no age, gender, smoking, histological, or TNM stage restrictions, but patients with a prior history of Cancer were excluded from this study. In parallel 200 unrelated individuals with no evidence of lung or other cancer visiting the hospital for health check-ups were enrolled as control group. Each control was pair-matched by sex, age (±10 years) and smoking parameters .These characteristics allowed us to obtain control population without any possible risk bias for Lung Cancer. A detailed questionnaire was completed for each case and control by a trained interviewer. The questionnaire included information on demographic and smoking characteristics. 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]. While medical information of cases, including Histology, TNM classification, Clinical staging, Primary tumour size, involvement of lymph node and metastasis were obtained from medical records of the Hospital. Approximately 3-5ml of venous blood was collected from each Candidate recruited in the study population.

4.2 *Isolation Of DNA From Peripheral Blood*

Isolation of DNA from peripheral blood: Genomic DNA was isolated using standard Protein K digestion, phenol/chloroform extraction and ethanol precipitation method from whole blood samples of both cases and controls (Bartlett and White's method).

REQUIREMENTS:

- Washing buffer
- Lysis buffer

- Phenol:Chloroform:Isoamylalcohol (25:24:1)
- Chloroform:Isoamylalcohol (24:1)
- Isopropanol
- TE buffer

PROCEDURE:

Preparation of Buffers

Washing buffer, Lysis buffer and TE buffer were prepared as shown in tables below.

Table 4.1: Preparation of Washing Buffer

STOCK CONCENTRATION	WORKING CONCENTRATION
1M Sucrose	320 mM Sucrose
100% Triton X-100	1% Triton X-100
100m M Magnesium Chloride	5mM Magnesium Chloride
100m M Tris-HCl (p H=8.0)	10mM Tris-HCl pH (8.0)

Table 4.2: Preparation of Lysis Buffer

STOCK CONCENTRATION	WORKING CONCENTRATION

1 M Tris HCl Buffer(p H=8) 10% SDS 0.5M EDTA 5 M NaCl (p H=8.0) 10mg/ml Proteinase K	400 Mm Tris HCl Buffer (pH=8) 1% SDS 60 mM EDTA 150mM NaCl 10 µg/ml proteinase K

4.3 Procedure of DNA Isolation

- Took 5ml of blood and 5ml of Washing Buffer was added and mixed thoroughly. Centrifuged it at 3500 rpm for 5 minutes.
- Discarded the supernatant and added 5ml of Washing buffer (1.6ml 1M Sucrose,0.5 ml Triton X-100, 0.25ml MgCl2, 0.5 ml 100mM Tris HCl and 0.26ml of water) to the pellet and re suspended the pellet in the Buffer and centrifuged again (repeat this step thrice).
- Dissolved the pellet in 5ml of Lysis buffer (1 M Tris HCl 2ml,10% SDS 0.5ml ,0.5 M EDTA 0.6ml, 5M NaCl 0.15ml,10mg/ml Proteinase-K 0.05ml and water 1.7ml) and incubated at 44 °C overnight.
- Added an equal volume of Phenol: Chloroform: Isoamyl alcohol (PCI) 25:24:1(2.5ml Phenol, 2.4 ml Chloroform and 0.1ml isoamyl alcohol) and mixed the contents slowly.
- Centrifuged at 8000 rpm for 10 minutes at 4°C. Took the upper aqueous layer and again add PCI mix and centrifuged.
- Took the aqueous layer and added equal volume of Chloroform: Isoamyl alcohol (24:1).
- Centrifuged it at 6500 rpm for 5 minutes and took the upper layer.

- To the aqueous layer added equal volume of chilled Isopropanol or 2.5 times volume of absolute Ethanol and mixed it gently.
- Freeze it at -20°C for 1-2 hours.
- Centrifuged it at 12,000 rpm for 10 min at 4°C.The supernatant was discarded and the pellet of DNA was washed with chilled 70% Ethanol twice at 10,000 rpm for 5 minutes.
- Decant ethanol and air dry the pellet.
- Dissolved the pellet in 50µl-150µl Tris-EDTA buffer depending on the size of DNA pellet (Bartlett & White, 2003).

4.4 ***DNA Quantification***

The Thermo Scientific Nanodrop Spectrophotometer holds 1µl of sample without the need of traditional containment devices such as cuvettes and capillaries. Using fibre optic technology and surface tension, the sample is held in place between two optical surfaces that define the path length in vertical orientation. Removal of fixed containment devices from the system allows the path length to change in real time for a sample. This essentially eliminates the need to perform dilutions and hence less cumbersome.

Procedure

- Pipetted 1µl of Deionised water onto the lower optical surface of Nanodrop (Thermo Scientific) to clean it
- Opened the Nanodrop software and select Nucleic acid Module
- Performed a blank measurement by loading 1µl of TE and selecting “blank” from the screen
- Measured the Nucleic acid sample by loading 1µl of DNA sample and selecting “measure”
- Concentration and purity of DNA samples were calculated automatically

DNA concentration otherwise can be calculated as:

DNA concentration (µg/ml) = O.D at 260nm X 50 X Dilution factor

where 50µg/ml of DNA is equal to 1 O.D

Purity of DNA = O.D at 260nm/O.D at 280nm

REMARK : A ratio of ~1.8 indicates purity of DNA; a ratio of ~2.0 is generally accepted as pure for RNA. If the ratio is appreciably lower in either case, it may indicate the presence of protein, phenol or other contaminants.

4.4.*1 Resolution Of DNA Fragments On Agarose Gels*

Requirements

- Electrophoresis buffer (TAE or TBE)
- Ethidium bromide solution
- Electrophoresis-grade Agarose
- 6X loading dye
- DNA molecular weight markers
- Horizontal gel electrophoresis apparatus
- Gel casting platform
- Gel combs (slot formers)
- DC power supply

Procedure

Preparing 5X TBE (1000ml)

Tris base - 54 g

Boric Acid - 27.5g

EDTA (0.5M) - 20ml

Make up final volume with water.

Preparing 6X Loading Dye (20ml)

0.25% Bromophenol blue - 0.05gm

0.25% Xylene Cyanol - 0.05gm

40% Sucrose - 8gm

Make up final volume with TE buffer.

4.4.2 ***Preparation of the Agarose Gel for Electrophoresis***

- Prepared an adequate volume of electrophoresis buffer.
- Added the desired amount of Electrophoresis-grade Agarose to a volume of Electrophoresis buffer sufficient for preparing the gel. For example, for genomic DNA 0.7% gel (0.7g agarose in 100ml 0.5X TBE) was prepared while for the PCR products 1.7% gel (1.7g agarose in 100ml 0.5X TBE buffer) was prepared.
- Melted agarose was cooled to 55°C in a water bath before pouring onto the gel platform to prevent warping of the gel apparatus.
- Before pouring Ethidium bromide solution was added to the melted well mixed agarose gel to a final concentration of 0.3μg/ml to facilitate visualization of DNA when seen under UV Transilluminator.
- Poured the melted Agarose onto the gel casting apparatus between 0.5 and 1 cm thick and with the gel comb inserted prior to pouring, making sure that no bubbles are trapped underneath the combs and all bubbles on the surface of the agarose were removed before the setting of the gel.

Loading and running the gel

- After the gel got solidified, the gel comb was withdrawn with proper care without disrupting the sample wells.
- Placed the gel casting platform containing the set gel in the Electrophoresis tank. Added sufficient Electrophoresis buffer to cover the gel until the tops of the wells are submerged. Made sure no air pockets were trapped within the wells.
- DNA samples were prepared by mixing 5µl DNA with 2µl of 6X loading dye and 2ul water in case of genomic DNA or by mixing 5µl DNA with 2µl of 6X loading dye in case of PCR product.
- Samples were typically loaded into the wells with micropipette. Care was taken to prevent mixing of the samples between wells.
- Appropriate DNA molecular weight marker were also loaded in case of PCR and digestion products
- Connected the electrodes to a power pack and allowed the Electrophoresis apparatus to run at 60 V until the marker dyes migrated the desired distance.
- Turned off the electric power, disconnected the leads, and discarded the electrophoresis buffer from the reservoirs
- DNA was visualized by placing the gel on a UV transilluminator and then photographed using Gel Documentation.

4.5 POLYMERASE CHAIN REACTION (PCR):AMPLIFICATION OF AhR

The polymerase chain reaction (PCR) is an enzymatic process that allows for the detection of specific genes within an environmental DNA sample. PCR utilizes short, user defined DNA sequences called oligonucleotide primers, the sequences of which are complementary to target regions of genes known to encode for specific microbial functions (e.g. contaminant degradation). Polymerase Chain Reaction (PCR) is a very sensitive assay in which a single DNA molecule can be amplified and single-copy genes can be extracted out of complex mixtures of genomic sequences In brief, the DNA sample is denatured to produce single stranded DNA, called template DNA, to which the oligonucleotide primers can bind. The enzyme DNA polymerase then adds nucleotide bases to the end of each primer, using the template DNA as a guide to extend the primer thereby producing new double stranded DNA. This process is repeated for a number of cycles to enrich the DNA sample for the desired genes targeted by the oligonucleotide primers (Laboratory of Environmental Pathogen Research). PCR can also be utilized for rapid screening and/or sequencing of inserts directly from aliquots of individual phage plaques or bacterial colonies.

Requirements

- 10X PCR buffer
- BSA
- Forward Primer
- Reverse Primer
- dNTPs
- Taq DNA polymerase
- Water
- DNA sample

PCR Primers are short fragments of single stranded DNA (15-30 nucleotides in length) that are complementary to DNA sequences that flank the target region of interest. The purpose of PCR primers is to provide a "free" 3'-OH group to which the DNA polymerase can add dNTPs. Primer sequences need to be chosen uniquely. Specific region of DNA should be chosen and primer should be designed ubiquitously, avoiding the possibility of mis hybridization to a similar sequence nearby. Primers should not easily anneal with other primers in the mixture (either other copies of same or the reverse direction primer); this phenomenon can lead to the production of 'primer dimer' products contaminating the mixture. Primers should also not anneal strongly to themselves, as internal hairpins and loops could hinder the annealing with the template DNA. Given below are the primers used.

Table 4.3: Primers used in PCR amplification of AhR SNPs

GENE	ALTERATION	PRIMER USED	PRODUCT SIZE
AhR rs 10252822	T → C	Forwad primer : 5-TGA TGC TTG GTA TGG GGT CTG AGT G-3 **25** Reverse primer: 5-CCT CCG TGG GCT GAA GAA TAT GTG T-3	412 bp

AhR rs 2282885	T → C	Forwad primer: 5-AAC TGC ACT TGA CTT GGA TTA CGC T-3 **25** Reverse primer: 5-AAG ATA GCA TTC AGA CTG GCA TTG G-3 **25**	377 bp
AhR rs 2066853	G → A	Forwad primer: Reverse primer:	152 bp

(Ping bin *et al.*,2011)

Table 4.4 :Reaction mixture for PCR carried out for *AhR* rs 10252822			
Reagent	**Stock Concentration**	**Final Reaction Concentration**	**Quantity Used**
Additive 1 BSA	1000 µg/ml	100 µg/ml	33 µl
PCR Buffer (Mg Conc.)	10 X 25mM	1 X 1.5mM total	33 µl
AhR rs 10252882 Primer (Forward)	10µM	0.5	16.50 µl
AhR rs 10252882 Primer (Reverse)	10µM	0.5	16.50 µl
Magnesium Chloride	50mM	2 mM	2.42 µl
Taq Polymerase	2.0U/µl	1.5Mm	3.52 µl
dNTP	10mM each	0.2mM each	6.60 µl
PCR Grade Water			130.48 µl

DNA Template	100ng/µl	400ng	2 µl

Table 4.5 : Cycling profile of PCR for *AhR rs 10252882*		
Steps	Temperature	Time
Initial Denaturation	95°C	5min
Denaturation	94°C	30 seconds
Annealing	61°C	45 seconds
Polymerization	72°C	30 seconds
Final Extension	72°C	5 min

The above reaction was carried out for 30 cycles

Table 4.6: Reaction mixture for PCR carried out for *AhR rs 2282885 and AhR Arg554 Lys rs 2066853 (*20 reactions)			
Reagent	**Stock Concentration**	**Final Reaction Concentration**	**Quantity Used**
Additive 1 BSA	1000 µg/ml	100 µg/ml	44 µl
PCR Buffer (Mg Conc.)	10 X 25mM	1 X 1.5mM total	44 µl
AhR rs 2282885 Primer (Forward)	10µM	0.5	22 µl
AhR rs 2282885 Primer (Reverse)	10µM	0.5	22 µl
Taq Polymerase	2.0U/µl	1.5mM	5.87 µl

dNTP	10mM each	0.2mM each	8.80 µl
PCR Grade Water			201.81 µl
DNA Template	100ng/µl	400ng	2 µl

Table 4.7 : Cycling profile of PCR for *AhR rs 2282885 and AhR Arg554 Lys rs 2066853*

Steps	Temperature	Time
Initial Denaturation	95°C	5min
Denaturation	94°C	30 seconds
Annealing	65°C (***AhR rs 2282885)***	45 seconds
	54°C (***AhR Arg554 Lys rs 2066853)***	30 seconds
Polymerization	72°C	30 seconds
Final Extension	72°C	5 min

- The reaction is carried out for 29 cycles for amplification of ***AhR rs 2282885*** and the reaction is carried out for 35 cycles for amplification of ***AhR Arg554 Lys rs 2066853***

4.5 Restriction Digestion of AhR SNPs

This enzymatic technique can be used to cleave the DNA amplicons at specific sites ensuring that all the all DNA fragments that contain a particular sequence have the same size; furthermore, each fragment that contains the desired sequence has the sequence located at exactly the same position within the fragment. The enzymes used were:

BsmA1*:** The source microorganism from which it has been isolated is *Bacillus stearothermophilus* A664. This enzyme was used for the digestion of amplicons obtained from the amplification of ***AhR rs 10252882. The enzyme was used along with NEB4 Buffer (10X) which helps in maintaining the homeostatic conditions with respect to ions within the reaction mixture.

5'...GTCTC(N)1^...3'

3'...CAGAG(N)5^...5'

BspH1: The source microorganism from which it has been isolated is from *Bacillus* species H (NEB 394). This enzyme was used for the digestion of amplicons obtained from the amplification of ***AhR rs 2282885.*** It recognizes the T→C site in the amplified stretch of DNA sequence.

5' . . . T^{v}C A T G A . . . 3'

3' . . . A G T A C$_{\wedge}$T . . . 5'

Dde1 : the source microorganism from which it has been isolated is *Desulfovibrio desulfuricans.* This enzyme cleave on a GA nucleotide change.

4.6 Statistical Analysis

Differences in the distributions of demographic characteristics between the cases and controls were evaluated using the Chi-square tests (χ^2 test) for the categorical data and student *t* test for continuous variables. The Hardy–Weinberg equilibrium 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 both in cases and controls to calculate the genotype frequencies of *AhR* gene polymorphism using χ^2 test. Pearson's χ^2 test was used to determine whether there was any significant difference in allele and genotype frequencies between cases and controls. To assess the risk for lung cancer and polymorphisms adjusted Odds Ratio (ORs) along with 95% Confidence Intervals (CI) were calculated using logistic regression analysis with adjustment for possible confounders (age and pack-years of smoking as continuous variables; and gender as a nominal variable). The homozygous wild genotype TT of the AhR rs10252882 and rs2282552 was used as the reference and for polymorphism in AhR rs2066853 GG was taken as the homozygous wild genotype in calculating the ORs and 95%CI. 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 version 9.3.6.0 (Medcalc Software, Ostend, Belgium)

Results

5.1Genotyping

DNA was isolated from the peripheral blood and samples were run on 0.7% gel. The total DNA from peripheral blood was preserved as a stock of the candidate s DNA and a part of it was then diluted to a concentration of 100ng/µl. The diluted DNA was then used as a template in different volumes in the PCR as deduced by Biometra for the amplification reactions.

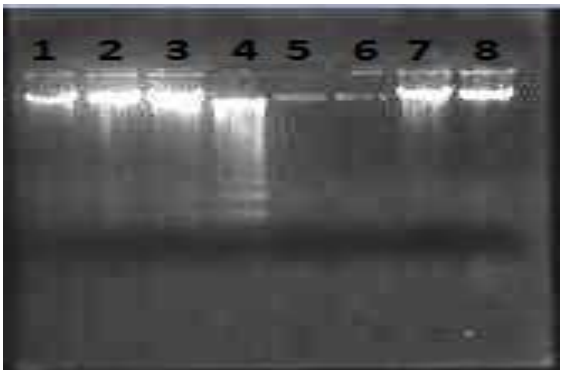

Figure 5.1: Representative example of the Agarose gel (0.7%) showing the presence of genomic DNA in the samples isolated using the White and Bartlet's method

Primers designed corresponding to the regions of the gene to be amplified were used for the amplification of the desired region of the gene. The amplicons obtained after the amplification process were run on agarose gel (1.7%). The agarose gel is stained with ethidium bromide which stacks between the basepairs of the DNA and is responsible for the illuminating when seen in the UV light. This acts as a test for whether the DNA has undergone amplification process or not.

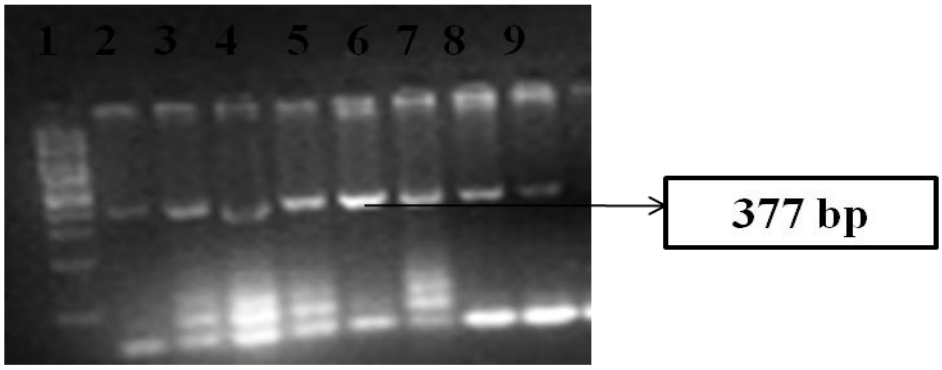

Figure 5.2: Representative example of the PCR product of AhR rs2282885 with amplicon size 377bp Lane 1: 100 bp Ladder (G Biosciences); Lane 2,3,4,5,6,7,8,9: PCR products of amplification reaction

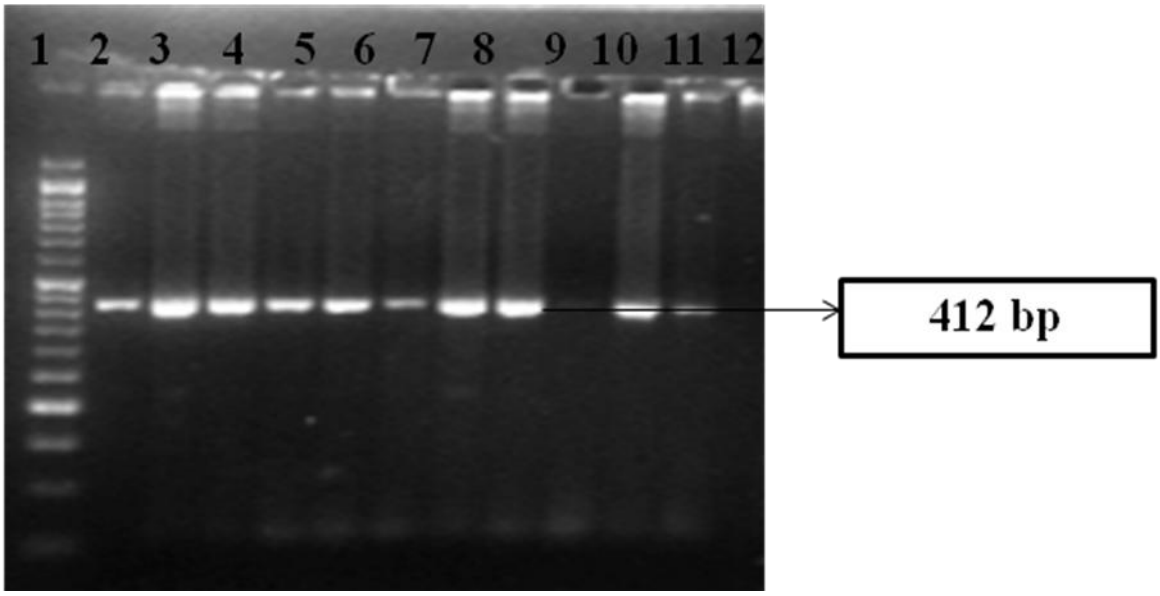

Figure 5.3: Representative example of PCR product of AhR rs 10252882 with amplicon size 412bp

Lane 1:100 bp Ladder(G Biosciences)

Lane 2,3,4,5,6,7,8,9: PCR products

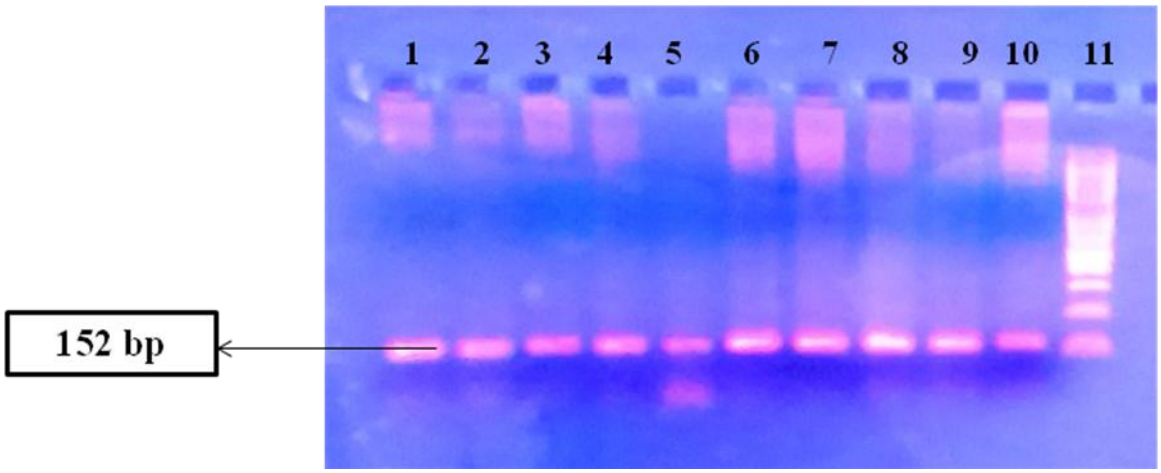

Figure 5.4: Representative example showing the PCR product of Ahr Arg 554Lys rs2066853 (amplicon size:152bp)

Lane 1-10:PCR Products of AhR Arg 554 Lys rs2066853

Lane 11: 50 bp Ladder (G Biosciences)

The PCR products were then digested with their corresponding enzymes capable of excising the PCR product at the unique cleavage site which remains specific to a single nucleotide sequence and enzyme. The digestion reaction is again checked by running the sample in (2.5%) agarose gel for ***AhR rs10252882*** and ***AhR rs2282885***. the digestion product was

visualized by running it one (3%) agarose gel in case of ***AhR rs2066853*** due to the closely spaced digestion products hence obtaining clear and distinct gel results.

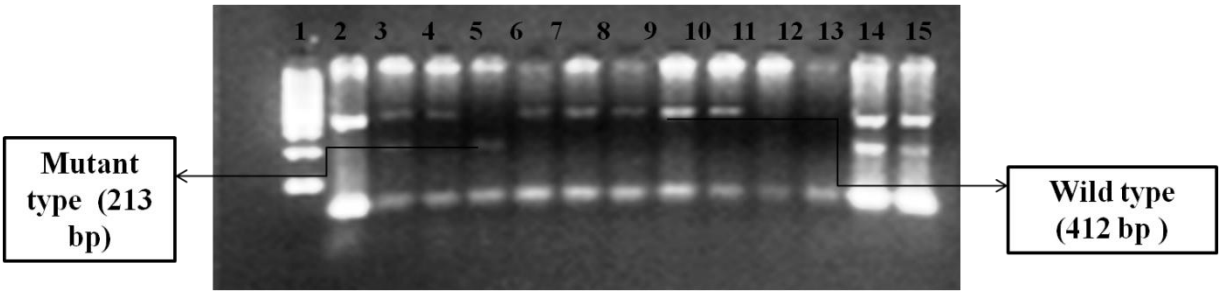

Figure 5.5: Representative example for the Restriction digestion of PCR product of Ahr rs10252822 showing the digested products (wild:412bp ;mutant:213 bp;heterozygote:412bp,213bp,
Lane 1:100 bp Ladder (G Biosciences);
Lane 2: uncut ;Lane 5: Mutant ;Lane 3,4,6,7,8,9,10,11,12,13: Wild type
Lane 14,15: Heterozygote type

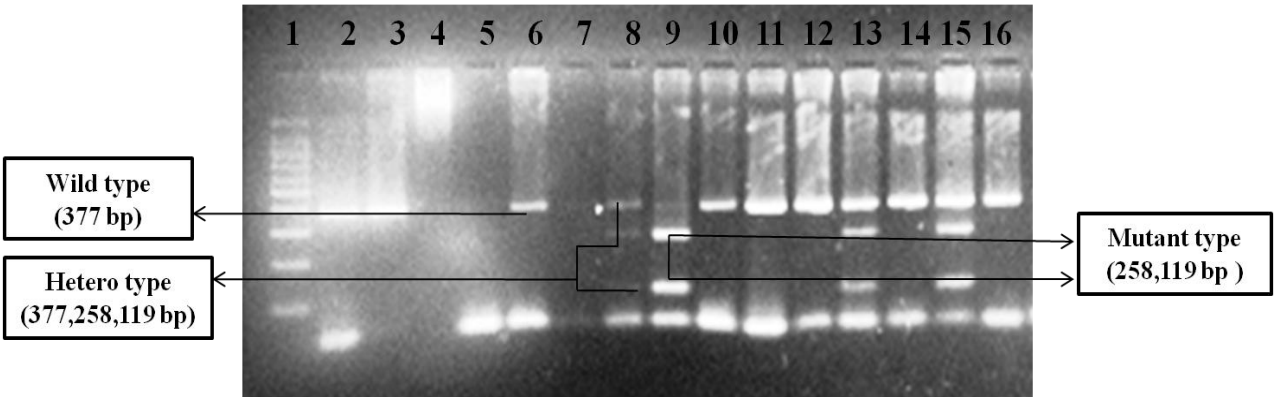

Figure 5.6: Representative example showing the restriction digestion of PCR product of AhR rs 2282885 showing the digested products (wild type: 377bp ; mutant type: 258bp,119bp; hetero type:377bp,258bp,119bp)
Lane 1:100 bp Ladder (G.Biosciences); Lane 2,3,6,10,11,12,14,16: Wild type
Lane 8,13,15:heterozygote type; Lane 9:Mutant type

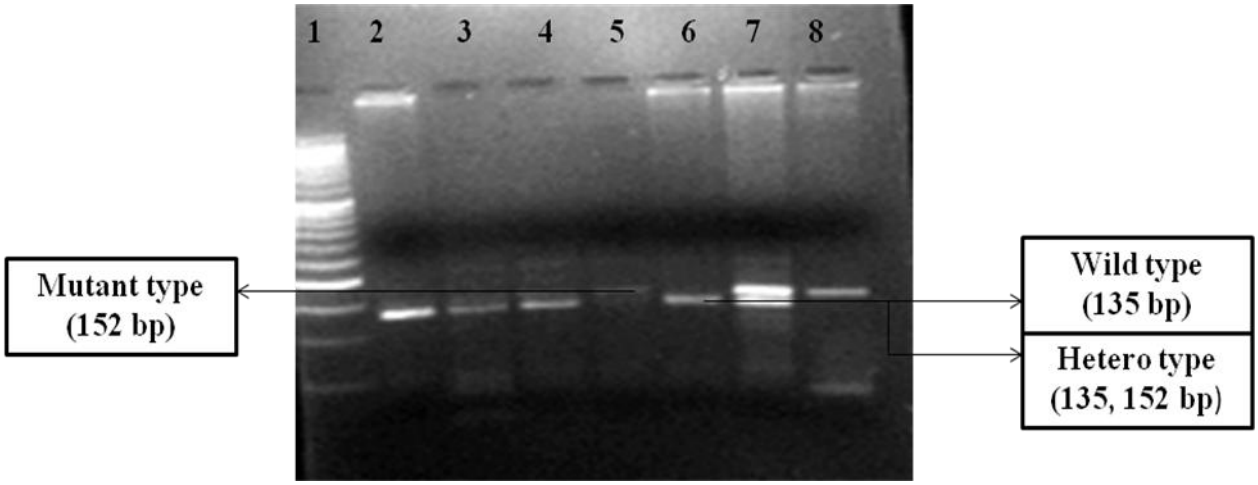

Figure 5.7: Representative example of restriction digestion of PCR products of Ahr rs2066853 showing the digested products (wild type:135bp, hetero type:135bp,152bp; mutant type:152bp)
Lane 1:Ladder (G.biosciences); Lane 2:Uncut
Lane 3,4,6:Wild type; Lane 5:Mutant type; Lane 7: Heterozygote type

5.2 Epidemiology for AhR rs10252882 polymorphism

TABLE 5.1: Distribution of demographic characteristics for AhR rs10252882 polymorphism amongst cases and controls			
Variable	**Cases, n (%) N = 201**	**Controls, n (%) N = 195**	***p* – value**
Age (years) Mean ± SD Range	57.7±10.04 31-86	52.09±9.82 32-80	P<0.0001
Gender Male Female	176 (87.56) 25 (12.43)	191 (97.94) 4 (2.05)	0.0002
Smoking Status Smokers Non – Smokers	163 (81.09) 38 (18.9)	147 (75.381) 47 (24.10)	0.24
Pack Years Mean ± SD	27.94±33.4	20.04±21.11	0.0056
Histological Types SQCC ADCC SCLC	84 (42) 61(30) 56 (29)		
TNM Staging I II III IV Unclassified	2 (1) 11 (54.7) 96 (47.76) 79 (39.30) 13 (6.46)		
Tumor Size T1 T2 T3 T4 Unknown	11 (5.4) 22 (10.9) 34 (16.9) 84 (41.8) 43 (21.44)		

Lymph Node Involvement N0 N1 N2 N3 N4 Unknown	 37 (18.4) 18 (8.9) 84 (41.8) 55(27.36) 4 (2) 3(1.5)		
Metastasis M0 M1 Unknown	 108 (53.73) 90 (44.77)		

Demographic characteristics of the study groups including age, gender, smoking status, pack years, histological subtypes, TNM staging and other clinical parameters are shown in the above table. The case control study pertains to 201 lung cases and 195 controls. The mean age of the cases was 57.7±10.04(31-86) whereas the mean age of all the control subjects were 52.09±9.82 (32-80). The study comprised of 176(87.56%) males and 25(12.5%) females in the rubric of cases and 191(97.5%) males and 4(2.5%) females in controls. Significant difference in the distribution of males and females was seen. In the present study, 81.09% cases were smokers and 19 % were non smokers, whereas in the control group 75.38 % were smokers and 24.10% were non smokers. Upon statistical testing it was inferred from the values that smoking came out to be an important parameter for the assessment of risk towards acquiring Lung Cancer. In addition to this it was also seen that the pack years for the smokers were significantly higher in cases as compared to controls (27.94±33.4 vs 20.04±21.11: p <0.0056). Variables like age, gender, smoking exposure were further adjusted for any residual confounding effect using multivariate logistic regression analyses. Of all the 201 Lung Cancer cases 84(42%) were squamous cell carcinoma, 56(30%) were Adenocarcinoma ,56(29.0%) were small cell carcinoma. TNM stage data was also evaluated for the patients (Stage I: 2,II: 11, III:96, IV: 79). Amongst them 13(6.5) cases remained unclassified out of the total 201 cases.

5.3 Epidemiology for AhR rs2282885 polymorphism

TABLE 5.2: Distribution of demographic characteristics for AhR rs2282885 polymorphism amongst cases and controls			
Variable	**Cases, n (%) (200)**	**Controls, n (%) (201)**	***p* – value**

Age (years) Mean ± SD Range	57.5±10.18 31-86	52.7±10.3 30-80	< 0.001
Gender Male Female	173(86.5) 27 (13.5)	198 (98.5) 3(1.5)	<0.0001
Smoking Status Smokers Non – Smokers	160(80) 40(20)	156(77.61) 47(23.38)	0.5170
Pack Years Mean ± SD	33.35±31.25	25.6±20.86	0.0001
Histological Types SQCC ADCC SCLC	84 (42) 56 (28) 53 (26.5)		
TNM Staging I II III IV Unclassified	1 (0.5) 12 (6) 89 (44.5) 85 (42.5) 17 (8.5)		
Tumor Size T1 T2 T3 T4 Unknown	8 (4) 19 (9.5) 40 (20) 80 (40) 54 (27)		
Lymph Node Involvement N0 N1 N2 N3 N4 Unknown	34 (17) 16 (8) 63 (31.5) 31 (15.5) 4 (2) 50 (25)		
Metastasis M0 M1 Unknown	76 (38) 68 (34) 50(25)		

Demographic characteristics of the study groups including age, gender, smoking status, pack years, histological subtypes, TNM staging and other clinical parameters are shown in the above table. The case control study pertains to 201 lung cases and 200 controls. The mean age of the cases was 57.5±10.18 (31-86), whereas the mean age of all the control subjects were 52.7±10.3 (30-80). The study comprised of 173(86.5%) males and 27(13.5%) females in the rubric of cases and 198(98.5%) males and 3(1.5%) females in controls. Significant difference in the distribution of males and females was seen. In the present study, 80% cases were smokers and 20 % were non smokers, whereas in the control group 77.61 % were

smokers and 23.38% were non smokers. Upon statistical testing it was inferred from the values that smoking came out to be an important parameter for the assessment of risk towards acquiring Lung Cancer. In addition to this it was also seen that the pack years for the smokers were significantly higher in cases as compared to controls (33.35± 31.25 vs 25.6± 20.86: p <0.0001). Variables like age, gender, smoking exposure were further adjusted for any residual confounding effect using multivariate logistic regression analyses. Of all the 200 Lung Cancer cases 84(42%) were squamous cell carcinoma, 56(28%) were Adenocarcinoma, 53(26.5%) were small cell carcinoma and 7 (3.5%) undifferentiated carcinoma. TNM stage data was also evaluated for the patients (Stage I: 1, II: 12, III:89, IV: 85). Amongst them 17(8.5) cases remained unclassified out of the total 200 cases.

5.4 Epidemiology for AhR rs2066853 polymorphism

TABLE 5.3: Distribution of demographic characteristics for AhR rs2066853 polymorphism amongst cases and controls

Variable	**Cases, n (%) N = 196**	**Controls, n (%) N = 191**	***p* – value**
Age (years) Mean ± SD Range	57.9±10.62 29-85	54.6±10.22 30-83	0.0018
Gender Male Female	173 (88.26) 23 (11.73)	179(93.71) 12 (6.28)	0.0906
Smoking Status Smokers Non – Smokers	162 (82.65) 34 (17.34)	151 (79.05) 40 (21)	0.441
Pack Years Mean ± SD	26.62±30.61	20.57±20.74	0.023

Histological Types SQCC ADCC SCLC	 88(45) 51(26.02) 50 (25.51)		
TNM Staging I II III IV Unclassified	 2 (1.02) 12 (6.12) 115 (58.7) 108 (55.10) 42 (2.42)		
Tumor Size T1 T2 T3 T4 Unknown	 8 (4.08) 25 (12.75) 39 (20) 76 (39) 50 (26)		
Lymph Node Involvement N0 N1 N2 N3 N4 Unknown	 25 (12.75) 14 (7.1) 63(32.1) 39 (20) 3 (1.5) 45 (23)		
Metastasis M0 M1 Unknown	 79(40.3) 72 (37) 45 (23)		

Demographic characteristics of the study groups including age, gender, smoking status, pack years, histological subtypes, TNM staging and other clinical parameters are shown in the above table. The case control study pertains to 196 lung cancer cases and 191 controls. The mean age of the cases was 57.9±10.62 (29-85) whereas the mean age of all the control subjects were 54.6±10.22 (30-83). The study comprised of 173(88.26%) males and 23(11.7%) females in the rubric of cases and 179(93.7%) males and 12(6.3%)females in controls. Significant difference in the distribution of males and females was seen. In the present study, 82.65% cases were smokers and 17.34% were non smokers, whereas in the control group 79.05 % were smokers and 21% were non smokers. Upon statistical testing it was inferred from the values that smoking came out to be an important parameter for the assessment of risk towards acquiring Lung Cancer. In addition to this it was also seen that the pack years for the smokers were significantly higher in cases as compared to controls (26.62±30.61 vs 20.57±20.74: $p < 0.400$). Variables like age, gender, smoking exposure were further adjusted for any residual confounding effect using multivariate logistic regression analyses. Of all the 196 Lung Cancer cases 88(45%) were squamous cell carcinoma, 51(26.02%) were Adenocarcinoma, 50(25.50%) were small cell carcinoma. TNM stage data was also evaluated for the patients (Stage I: 2, II: 12, III: 115, IV: 108). Amongst them 42(21.42) cases remained

unclassified out of the total 196 cases.

5.5 Genotypic and allelic frequency distribution of AhR rs10250822 and its association with risk of acquiring Lung Cancer

Table 5.4 :Genotypic frequency distribution of AhR rs10250822 among patients and controls and its susceptibility towards Lung Cancer				
Gene	**Cases, N (%) N = 201**	**Controls, N (%) N = 195**	**Adjusted OR (95% CI)**	***p* value**
AhR rs10250822				
TT	81 (40.29)	120 (61.53)	1.00 (Reference)	
CT	102 (49.75)	56(28.717)	2.5 (1.6-4.02)	**0.0001**
CC	18 (8.995)	18 (9.23)	1.35(0.62-2.91)	0.44
CT +CC vs TT	120(59.701)	74 (37.94)	2.20(1.43-3.38)	**0.0003**
T	264(65.67)	296(75.89)	1.00(Reference)	
C	138(34.32)	258(66.15)	0.599(0.46-0.78)	**0.0002**
Minor allele frequency	0.34	0.24		

The genotypes of cases and controls recruited for the population study for ***Ahr rs 10250822*** were obtained by PCR-RFLP. Of the total cases (201) studied , (40.3%)of the individuals were found to have homozygous wild type (TT) genotype, (50.70%) had heterozygous genotype (CT) and (9.00%) individuals had homozygous mutant genotype. On the other hand out of all the controls (195) studied, (62.3%) of the individuals were found to have homozygous wild type (TT) genotype, (29.2%) of them had heterozygous genotype (CT) and (9.23%) individuals had homozygous mutant genotype (CC). Due to the low prevalence of the homozygous mutant type genotype, we combined the heterozygous and mutant genotypes under a single genotype. This combined genotype was then compared to the homozygous wild genotype which is taken as the reference genotype.

With reference to TT genotype of the *AhR* gene polymorphism, it was observed that the CT genotype shows an increased risk of Lung Cancer (OR =2.5; 95% CI =1.6-4.02;p=0.0001). On considering only the mutant genotype, a dip in the risk towards Lung Cancer was observed but was found to be statistically insignificant (OR=1.35;95% CI =0.62-2.91; p= 0.44). Due to the low prevalence of homozygous variant genotype, we combined both the heterozygous (CT) and the homozygous variant genotype (CC) as a single genotype and

compared it with reference genotype (TT). An exorbitant increase in the statistical values was observed validating the association of Lung Cancer with the genotype taken in combination (OR=2.20; 95%CI =1.43-3.38; p=0.0003)

The allelic and genotypic frequencies of the AhR rs10250822 C/T polymorphism are also summarised in the above table. The genotypic frequencies of AhR rs10250822 C/T polymorphism in the control group (n= 195: $\chi 2$=7.93: df=2: p=0.0002) were in Hardy Weinberg equilibrium, suggesting no sampling bias and absence of population stratification. Distribution of *AhR* rs10252882 genotype was significantly distinct between cases and controls respectively (TT=81,CT=102,CC=18) : (TT=120,CT=56,CC=18) respectively. The Minor allele frequency for cases and controls were 0.34 and 0.24 respectively.

5.6 Genotypic Distribution and association of AhR gene rs10250822 among the patients with different histological forms of Lung Cancer

Table 5.5 :Genotypic frequency distribution of AhR rs10250822 among patients and controls on the basis of histological subtypes of Lung Cancer

***AhR rs 10250822* (SQCC)**	**Cases, N (%) N = 84(42)**	**Controls, N (%) N = 195**	**Adjusted OR (95% CI)**	***p* value**
TT	36 (43)	120 (61.53)	1.00 (Reference)	
CT	44 (52)	56(28.717)	2.41(1.17-3.47)	**0.002**
CC	4 (5)	18 (9.23)	0.76 (0.22-2.60)	0.66
CT +CC vs TT	48 (57)	74 (37.94)	2.01(1.17-3.47)	**0.01**

AhR rs 10250822 **(ADCC)**	**Cases, N (%) N = 61(30)**	**Controls, N (%) N =195**	**Adjusted OR (95% CI)**	***p*** **value**
TT	24 (39)	120 (61.53)	1.00 (Reference)	
CT	29 (48)	56(28.717)	2.88 (1.31-4.88)	**0.003**
CC	8 (13)	18 (9.23)	2.00 (0.07-5.9)	0.21
CT+CC vs TT	37 (61)	74 (37.94)	2.53 (1.31-4.88)	**0.005**
AhR rs 10250822 **(SCLC)**	**Cases, N (%) N = 56(28)**	**Controls, N (%) N =195**	**Adjusted OR (95% CI)**	***p*** **value**
TT	21 (37)	120 (61.53)	1.00 (Reference)	
CT	29 (52)	56(28.717)	2.40 (1.21-4.76)	**0.012**
CC	6 (11)	18 (9.23)	1.84 (0.61-5.55)	0.277
CT+CC vs TT	35 (62.5)	74 (37.94)	2.17 (1.13-4.17)	**0.019**

Among the cases studied in the population (42%) of the cases were of those who were diagnosed with SQCC, (28%) had SCLC and (30%) were diagnosed with ADCC. On further stratification of the population on the basis of the genotypes it was found that 43%, 39% , and 28% of the individuals of SQCC ,ADCC and SCLC respectively had homozygous wild genotype whereas (52%), (48%) and (52%) had heterozygote genotype . Mutant homozygous genotype was shown by (5%), (13%) and (11%) of the population cases respectively.

The table above shows the statistical values obtained for the association of C/T polymorphism in the AhR gene with that of acquiring Lung Cancer. On stratifying the data on histologically different forms of Lung Cancer it was seen that individuals with heterozygote genotype having Adenocarcinoma portrays a highly significant statistical value along with an approximate threefold increase in the risk towards acquiring Lung Cancer (OR=2.88; 95%CI =1.31-4.88; p= 0.003).Individuals with heterozygote genotype having SQCC showcased a two fold increase in the risk towards Lung Cancer (OR= 2.41;95%CI=1.71-3.47;p=0.002).The SCLC subgroup amongst the population was also competent in showcasing significant statistical values in the heterozygote genotypical subgroups (OR=2.40;95%CI = 1.2-4.76 ;p= 0.012) and stands at par with the former . The joint combination of genotypes (mutants and heterozygotes) also showed significant values in case of all the three subgroups of histological stratification. ADCC outstands the rest by

displaying a more than two fold increase in the risk (OR=2.53; 95%CI=1.31-4.88;p= 0.005) followed by SCLC (OR=2.17;95%CI=1.13-4.17;p= 0.019) and SQCC ranking the lowest

Table 5.6 : Genotypic frequency distribution of AhR rs10250822 among patients and controls on the basis of Smoking characteristics (Pack years) and its association with the risk towards Lung Cancer

(OR=2.01;95%CI =1.17-3.47;p=0.01)amongst all.

AhR rs 10250822 (Smokers)	***Cases, N (%) N = 163***	***Controls, N (%) N = 157***	***Adjusted OR (95% CI)***	***p value***
TT	63(38.6)	91(58)	1.00 (Reference)	
CT	86(52.8)	55(35)	2.22 (1.36-3.64	**0.0014**
CC	14(8.6)	11(7)	1.87(0.78-4.48)	0.15
CT +CC vs TT	100(61.3)	66(42)	2.41(1.50-3.88)	**0.0003**
AhR rs10250822 (Heavy smokers)	***Cases, N (%) N =108 (66)***	**Controls, N (%) N =82 (52.2)**	**Adjusted OR (95% CI)**	***p* value**
TT	42 (39)	49 (60)	1.00 (Reference)	
CT	56 (52)	26 (32)	3.59(1.08-3.94)	**0.02**
CC	10 (9)	7 (8)	1.76(0.59-5.14)	0.31
CT+CC vs TT	66 (61)	33 (40)	1.99(1.087-3.67)	**0.025**
AhR rs10250822 (Light smokers)	***Cases, N (%) N =54 (33.1)***	***Controls, N (%) N =75 (48)***	***Adjusted OR (95% CI)***	***p value***
TT	20 (37)	45 (60)	1.00 (Reference)	
CT	32 (60)	25 (33.33)	2.06(1.61-8.02)	**0.0018**
CC	2 (3.7)	5 (6.66)	1.89(0.38-9.25)	0.42
CT+CC vs TT	34 (63)	30(40)	3.10(1.42-6.75)	**0.0043**
AhR rs10250822 (Non Smokers)	***Cases, N (%) N =37***	***Controls, N (%) N =49***	***Adjusted OR (95% CI)***	***p value***
TT	18 (49)	29(59.18)	1.00 (Reference)	
CT	15 (41)	12(24.48)	1.97(0.65-5.92)	0.22
CC	4 (4.3)	7(14.28)	0.30(0.04-2.32)	0.25
CT+CC vs TT	19(51.35)	20(41)	0.93(0.33-2.57)	0.89

5.7 Distribution of genotypes of AhR rs10250822 among smokers and non smokers and evaluation of risk towards acquiring Lung Cancer

The population recruited for the epidemiological study were categorized as Smokers and Non smokers to validate the association between cigarette smoking and the risk of developing Lung Cancer. (81.09%) of the cases studied were smokers and (19%) were non smokers. On

the contrary (80.51%) of the controls were smokers whereas (20%) were non smokers. Smokers were further categorized as Heavy smokers and Light smokers on the basis of the pack years *i.e.* the no of cigarettes smoked in a day by an individual multiplied with the total number of years smoked. Heavy smokers were those with pack yrs > and equal to 20 and the light smokers were those for whom the values of pack yrs fell below 20. In cases the heavy smokers accounted to (66%) and light smokers accounted to (33.1%), whereas in controls, (52.22%) were heavy and (48%) were light smokers.

Cigarette smoking showcases an additive effect on the polymorphism of *AhR* in association with Lung Cancer .In line with the epidemiological studies, the above table displays statistically significant values for the Smokers in contrast to that of Non Smokers .On categorizing the smokers as Heavy and Light a profound three -fold increase was witnessed in case of Heavy smokers with heterozygotes genotype (OR=3.59; 95%CI=1.61-8.02; p=0.0018). Individuals with heterozygote genotype amongst the Light Smokers showed two fold increase (OR=2.06; 95%CI=1.08-3.94,p=0.02) whereas the Non Smokers failed to show statistically significant values. The combinatorial genotypes were considered due to low incidence of the homozygous variants, hence grouping the homozygous variants (CC) and heterozygotes (CT) together, amongst which the Light Smokers showed highest statistically significant values with threefold increase in the risk towards acquiring Lung Cancer (0R= 3.10; 95%CI=1.42-6.75;p=0.0043) followed by the Heavy smokers (OR=1.99;95%CI=1.08-3.67; p=0.025)

5.8 Genotypic and allelic frequency distribution of AhR rs2282885 and its association with risk of acquiring Lung Cancer

Table 5.7 :Genotypic frequency distribution of AhR rs2282885 among patients and controls and its susceptibility towards Lung Cancer				
***AhR* rs 2282885**	**Cases, n (%) N = 200**	**Controls, n (%) N =201**	**Adjusted OR (95% CI)**	**p value**
TT	141(70.5)	149(74.129)	1.00 (Reference)	
CT	48(24)	42(20.89)	1.29(0.78-2.14)	**0.30**
CC	11(5.5)	10(4.975)	1.312(0.67-2.54)	**0.422**
CT+CC vs TT	59(29.5)	52(26.8)	1.197(0.75-1.90)	**0.441**
T	330(82.5)	340(85)	1.00(Reference)	
C	70(17.5)	62(15)	1.22(0.58-2.59)	**0.5896**
Minor allele frequency	0.18	0.15		

The genotypes of cases and controls recruited for the population study for *Ahr rs2282885* were obtained by PCR-RFLP. Of the total cases (200) studied , (70.5%)of the individuals were found to have homozygous wild type (TT) genotype, (24%)had heterozygous genotype (CT) and (5.5%) individuals had homozygous mutant genotype. On the other hand, out of all the controls (201) studied, (74.13%) of the individuals were found to have homozygous wild type (TT) genotype,(21.0%)of them had heterozygous genotype (CT) and (5.8%) individuals had homozygous mutant genotype(CC). Due to the low prevalence of the homozygous mutant type genotyp , we combined the heterozygous and mutant genotypes under a single genotype .In cases, the number was found to be (29.5%) whereas in controls they were found to be (26.8%) This combined genotype was then compared to the homozygous wild genotype which is taken as the reference genotype .

In accordance with the reference genotype (homozygous wild TT), it was seen that though none of the other genotypes (homozygous mutants and heterozygous genotype) were found to be statistically significant but on the contrary the homozygous mutants were seen to have a marginally higher association with the risk of Lung Cancer (OR=1.312; 95%CI=0.67-2.54; p=0.422) followed by those individuals with heterozygote genotype (OR=1.29; 95%CI=0.78-2.14; p=0.30). The combined effect of the genotypes also showcased association with the risk of developing Lung Cancer.

The allelic and genotypic frequencies of the ***AhR rs2282885*** C/T polymorphism are also summarised in the above table. The genotypic frequencies of ***AhR rs2282885*** C/T polymorphism in the control group (n= 201:$\chi 2$=7.96: df=2 : p=0.586) were in Hardy Weinberg equilibrium, suggesting no sampling bias and absence of population stratification. Distribution of ***AhR rs2282885*** genotype was significantly distinct between cases and

controls respectively (TT=141, CT=48, CC=11): (TT=149,CT=42,CC=10) respectively. The Minor allele frequency for cases and controls were 0.18 and 0.15 respectively.

5.9 Genotypic Distribution and association of AhR gene rs2282885 among the patients with different histological types of Lung cancer

Table 5.8:Genotypic frequency distribution of AhR rs2282885 among patients and controls subgrouped on the basis of histological forms of Lung Cancer

***Ahr* rs 2282885 (SQCC)**	**Cases, n (%) N = 84 (42)**	**Controls, n (%) N = 201**	**Adjusted OR (95% CI)**	***p value***
TT	55 (65)	149(74.129)	1.00 (Reference)	
CT	20 (24)	42(20.89)	1.36 (0.71-2.60)	0.3
CC	9 (11)	10(4.975)	2.80 (1.03-7.63)	**0.04**
CT +CC vs TT	29 (34)	52(25.87)	1.64 (0.92-2.93)	0.09
***Ahr* rs 2282885 (ADCC)**	**Cases, n (%) N =56 (28)**	**Controls, n (%) N = 201**	**Adjusted OR (95% CI)**	***p value***
TT	41 (73)	149(74.129)	1.00 (Reference)	
CT	14 (25)	42(20.89)	1.19 (0.56-2.52)	0.63
CC	1 (2)	10(4.975)	0.15 (0.01-2.25)	0.16
CT+CC vs TT	15 (27)	52(25.87)	0.96 (0.46-2.01)	0.93
***Ahr* rs 2282885 (SCLC)**	**Cases, n (%) N = 53 (27)**	**Controls, n (%) N = 201**	**Adjusted OR (95% CI)**	***p value***
TT	39 (73)	149(74.129)	1.00 (Reference)	
CT	13 (25)	42(20.89)	1.37 (0.63-2.9)	**0.04**
CC	1 (2)	10(4.975)	0.59 (0.06-5.14)	0.63
CT+CC vs TT	14 (26)	52(25.87)	1.24 (0.59-2.58)	0.55

Among the cases studied in the population (42%) of the cases were of those who were diagnosed with SQCC, (27%) had SCLC and (28%) were diagnosed with ADCC. On further stratification of the population on the basis of the genotypes it was found that 65%, 73%, and 73% of the individuals of SQCC, ADCC and SCLC respectively had homozygous wild genotype. Whereas (24%), (25%) and (25%) had heterozygote genotype. The mutant

homozygous genotype was shown by (11%), (2%) and (2%) of the population cases respectively. Due to low incidence of occurrence of the mutant genotype, we combined the homozygous mutant genotype with the heterozygous genotype under one single genotype. The count was found to be 34%, 27% and 26% respectively for SQCC, ADCC and SCLC, the different histological types of lung cancer.

Statistical analysis showed an increased association of the homozygous mutants (CC) for ***AhR rs 2282885*** with the risk of lung cancer which further upon sub grouping the population according to the histologically different forms of Lung cancer, fell in accordance with the former genotype based results. Homozygous mutants (CC) for SQCC gave statistically significant values and confers an almost three fold increase in the advent of developing the disease (OR=2.80; 95%CI=1.03-7.63, p =0.04). On the contrary individuals with a heterozygote genotype (CT) falling in the subgroup of SCLC, showed a slight dip in the values (OR=1.37; 95%CI=0.63-2.90; p =0.04). Patients having Adenocarcinoma were found to showcase insignificant values and hence can be said not to be associated with the risk of Lung Cancer. The joint genotype encompassing the homozygous mutants and the heterozygotes did not give statistically significant values and hence proved a non association with the risk of Lung Cancer.

5.10 Distribution of genotypes of AhR rs2282885 among smokers and non smokers and evaluation of risk towards acquiring lung cancer

Table 5.9:Genotypic frequency distribution of AhR rs2282885 among patients and controls on the basis of Smoking characteristics (Pack years) and its susceptibility towards Lung Cancer

***Ahr* rs 2282885 (Smokers)**	**Cases, n (%) N =160 (80)**	**Controls, n (%) N =154 (76.61)**	**Adjusted OR (95% CI)**	***p* value**
TT	110 (68)	119 (77)	1.00 (Reference)	
CT	41 (26)	29 (19)	1.47 (0.813-2.52)	0.21
CC	9 (6)	6 (4)	1.67(0.56-4.97)	0.35
CT +CC vs TT	50 (31)	35 (22)	1.47(0.87-2.48)	0.14

***Ahr* rs 2282885 (Heavy smokers)**	**Cases, n (%) N =108 (67.5)**	**Controls, n (%) N =80 (40)**	**Adjusted OR (95% CI)**	***p* value**
TT	32 (30)	62 (77)	1.00 (Reference)	
CT	69 (64)	15 (19)	2.20(1.04-4.64)	**0.03**
CC	7 (6)	3 (4)	2.21(0.52-9.31)	0.27
CT+CC vs TT	76 (70)	18 (22)	2.16(1.08-4.31)	**0.02**
***Ahr* rs 2282885 (Light smokers)**	**Cases, n (%) N =52 (33)**	**Controls, n (%) N =74 (48)**	**Adjusted OR (95% CI)**	***p* value**
TT	40 (57)	57 (77)	1.00 (Reference)	
CT	9 (13)	14 (18)	1.63(0.69-3.83)	0.262
CC	3 (5)	3 (4)	1.69(0.30-9.51)	0.54
CT+CC vs TT	12 (28)	17 (22)	1.05(0.44-2.46)	0.90
***Ahr* rs 2282885 (Non Smokers)**	**Cases, n (%) N =40 (20)**	**Controls, n (%) N =47 (23.38)**	**Adjusted OR (95% CI)**	***p value***
TT	31 (78)	30 (64)	1.00 (Reference)	
CT	7 (17)	13 (28)	1.97(0.65-5.92)	0.225
CC	2 (5)	4 (8)	0.30(0.04-2.32)	0.25
CT+CC vs TT	9 (22)	17 (36)	0.93(0.33-2.57)	0.89

The population recruited for the epidemiological study were categorized as Smokers and Non smokers to validate the association between cigarette smoking and the risk of developing Lung cancer. (80%) of the cases studied were smokers and (20%) were non smokers. On the contrary (76.61%) of the controls were smokers whereas (23.38%) were non smokers. Smokers were further categorized as Heavy smokers and Light smokers on the basis of the pack years i.e. the no of cigarettes smoked in a day by an individual multiplied with the total number of years smoked for. Heavy smokers were those with pack yrs > and equal to 20 and the Light smokers were those for whom the values of pack yrs fell below 20. In cases the heavy smokers accounted to (67.5%) and light smokers accounted to (33%), whereas in controls, (40%) were heavy and (48%) were light smokers.

In the epidemiological studies, Smokers in comparison to the Non Smokers showed marginally high propensity of acquiring the risk towards Lung Cancer, especially for those carrying homozygous mutant genotype (0R=1.67,95%CI=0.56-4.97;0.35) though not statistically significant. On grouping the population on the basis of Smoking and Pack years, significant results were seen for the subgroup of Heavy Smokers, depicting a two fold increase in the risk of developing Lung Cancer in the individuals carrying the heterozygote genotype (OR=2.20;95%CI=1.04-4.64;p =0.03). Candidates with mutant genotype under the same subgroup of Heavy Smokers were found to be statistically insignificant but contrarily a two fold increase in the chances of developing Lung Cancer (OR=2.21; 95%CI=0.52-9.31;p =0.27).The homozygous mutant and heterozygote genotypes in joint combination for Heavy Smokers showed an exorbitant increase in the risk of Lung Cancer development (OR=2.16;95%CI=1.08-4.31;p =0.02). Individuals sub grouped in Non smokers escaped from projecting statistically significant values.

5.11Genotypic and allelic frequency distribution of AhR rs 2066853 and its association with risk of acquiring lung cancer

Table 5.10 :Genotypic frequency distribution of AhR rs2066853 among patients and controls and its susceptibility towards Lung Cancer

AhR Arg^{554}Lys rs 2066853	**Cases, n (%) N = 196**	**Controls, n (%) N =191**	**Adjusted OR (95% CI)**	***p* value**
GG	156(79.6)	130 (68)	1.00 (Reference)	
GA	32(16.3)	42(22)	0.60 (0.35-1.03)	**0.06**
AA	8(4.1)	19(10)	0.34 (0.14-0.81)	**0.01**
GA+AA vs GG	40(20.4)	61(32)	0.52(0.32-0.83)	**0.007**
G	344(88)	302(77.04)	1.00(Reference)	
A	48(12)	80(21)	0.52(0.35-0.77)	**0.0013**
Minor Allele Frequency	0.12	0.21		

The genotypes of cases and controls recruited for the population study for ***Ahr rs2066853*** were obtained by PCR-RFLP. Of the total cases (196) studied , (79.6%)of the individuals were found to have homozygous wild type (GG) genotype, (16.3%)had heterozygous genotype (GA) and (4.1%) individuals had homozygous mutant genotype. On the other hand out of all the controls (191) studied, (68%) of the individuals were found to have homozygous wild type (GG) genotype (22%)of them had heterozygous genotype (GA) and (10%) individuals had homozygous mutant genotype(AA). Due to the low prevalence of the homozygous mutant type genotype, we combined the heterozygous and mutant genotypes under a single genotype. This combined genotype was then compared to the homozygous wild genotype which is taken as the reference genotype. The count for this combined genotype was found to be (20.4%) and (32%) respectively for cases and controls in the study population.

In the table above, the individuals with the homozygous mutant genotype (AA) are found to show statistically significant (0R=0.34; 95%CI=0.14-0.81;p=0.01) whereas those with heterozygote genotype fail to stand at par with the former by a slight margin though it showed an increment in the values which validate the chances of developing lung cancer (OR=0.60;95%CI=0.35-1.03;0.06). The joint effect of homozygote mutants and heterozygotes showed statistically significant (OR=0.52; 95%CI=0.32-0.83; p=0.007)

The allelic and genotypic frequencies of the ***AhR rs2066853*** polymorphism are also summarised in the above table. The genotypic frequencies of ***AhR rs2066853*** polymorphism in the control group (n= 191:$\chi 2$=21.55: df=2 : p=0.0002) were in Hardy Weinberg equilibrium, suggesting no sampling bias and absence of population stratification. Distribution of ***AhR rs2066853*** genotype was significantly distinct between cases and controls respectively (GG=156, GA=32,AA=8) : (GG=68,GA=22,AA=10) respectively. The Minor allele frequency for cases and controls were 0.12 and 0.21

5.12 Genotypic Distribution and association of AhR rs2066853 among the patients with different histological types of Lung cancer

Table 5.11 :Genotypic frequency distribution of AhR rs2066853 among patients and controls subgrouping on the basis of histological forms of Lung Cancer				
***Arg ^{554}Lys rs 2066853* (SQCC)**	**Cases, n (%) N = 88 (44.89)**	**Controls, n (%) N = 191**	**Adjusted OR (95% CI)**	***p* value**
GG	71 (81)	130 (68)	1.00 (Reference)	
GA	13 (15)	42(22)	0.56 (0.27-1.13)	0.11
AA	4 (4)	19(10)	0.31 (0.10-0.98)	**0.04**
GA+AA vs GG	17 (19)	61(32)	0.47 (0.25-0.88)	**0.01**
***Arg^{554} Lys rs 2066853* (ADCC)**	**Cases, n (%) N = 51 (26.02)**	**Controls, n (%) N =191**	**Adjusted OR (95% CI)**	***p* value**
GG	40 (78)	130 (68)	1.00 (Reference)	
GA	10 (20)	42(22)	0.63 (0.27-1.45)	0.28
AA	1 (2)	19(10)	0.18 (0.02-1.47)	0.11
GA+AA vs GG	11 (21)	61(32)	0.51 (0.23-1.10)	0.08
***Arg ^{554}Lys rs 2066853* (SCLC)**	**Cases, n (%) N =50 (25.51)**	**Controls, n (%) N =191**	**Adjusted OR (95% CI)**	***p* value**
GG	40 (80)	130 (68)	1.00 (Reference)	
GA	8 (16)	42(22)	0.54 (0.22-1.33)	0.18
AA	2 (4)	19(10)	0.33 (0.07-1.50)	0.15
GA+AA vs GG	10 (20)	61(32)	0.47 (0.21-1.04)	0.06

Among the cases studied in the population (45.1%) of the cases were of those who were diagnosed with SQCC, (25.51%) had SCLC and (26.02%) were diagnosed with ADCC. On further stratification of the population on the basis of the genotypes it was found that 81%, 78% and 80 % of the individuals of SQCC, ADCC and SCLC respectively had homozygous

wild genotype. Whereas (15%), (20%) and (16%) had heterozygote genotype. The mutant homozygous genotype was shown by (4%), (2%) and (4%) of the population cases respectively. Due to low incidence of occurrence of the mutant genotype, we combined the homozygous mutant genotype with the heterozygous genotype under one single genotype. The count was found to be 19%, 21% and 20% respectively for SQCC, ADCC and SCLC, the different histological types of lung cancer.

Amongst all the histological forms of Lung cancer, in this population recruited for study , we saw that individuals belonging to the SQCC subgroup with a homozygous mutant genotype showed statistically significant values with a highly prominent protective effect .(OR=0.31,95%CI=0.10-0.98;p= 0.04)in contrast to patients having Adenocarcinoma and SCLC who failed to show significant values. The combined effect of homozygous mutants and heterozygotes in the subgroup of SQCC showed a statistically significant values (OR=0.47;95%CI=0.25-0.88;0.01)though it too shows a protective effect towards the risk of acquiring Lung Cancer.

5.13Distribution of genotypes of Ahr rs2066853 among smokers and non smokers and evaluation of risk toward acquiring Lung Cancer

Table 5.12 :Genotypic frequency distribution of AhR rs2066853 among patients and controls on the basis of Smoking characteristics (Pack years) and its susceptibility towards Lung Cancer

Arg554 Lys rs 2066853 **(Smokers)**	**Cases, n (%) N = 162 (84.18)**	**Controls, n (%) N = 151 (79.05)**	**Adjusted OR (95% CI)**	***p* value**
GG	131 (81)	100 (66)	1.00 (Reference)	
GA	25 (15)	32 (21)	0.63 (0.34-1.157)	0.137
AA	6 (4)	19 (13)	0.211(0.076-0.57)	**0.0025**
GA+AA vs GG	31 (19)	51 (34)	0.46(0.27-0.79)	**0.005**
Arg554 Lys rs 2066853 **(Heavy smokers)**	**Cases, n (%) N =106 (54.08)**	**Controls, n (%) N =84 (55.6)**	**Adjusted OR (95% CI)**	***p* value**
GG	91 (86)	60 (71)	1.00 (Reference)	
GA	11 (10)	15 (18)	0.503(0.20-1.20)	**0.123**
AA	4 (4)	9 (11)	0.21(0.05-0.80)	0.022
GA+AA vs GG	15 (14)	24 (28)	0.38(0.17-0.81)	**0.02**
Arg554 Lys rs 2066853 **(Light smokers)**	**Cases, n (%) N =69 (43)**	**Controls, n (%) N =79 (52.31)**	**Adjusted OR (95% CI)**	***p* value**
GG	51 (74)	50 (63)	1.00 (Reference)	
GA	16 (23)	19 (24)	0.84(0.38-1.85)	0.67
AA	2 (3)	10 (13)	0.21(0.04-1.04)	0.54
GA+AA vs GG	18 (26)	29 (37)	0.62(0.30-1.28)	0.19
Arg554 Lys rs 2066853 **(Non Smokers)**	**Cases, n (%) N =34 (17.34)**	**Controls, n (%) N =40 (21)**	**Adjusted OR (95% CI)**	***p* value**
GG	25 (73)	30 (75)	1.00 (Reference)	
GA	7 (21)	10 (25)	0.66 (0.20-2.15)	0.49
AA	2 (6)	-	-	-
GA+AA vs GG	9 (26)	10 (25)	0.93(0.31-283)	0.90

The population recruited for the epidemiological study were categorized as Smokers and Non smokers to validate the association between cigarette smoking and the risk of developing Lung cancer. (84.18) of the cases studied were smokers and (17.34) were non smokers. On the contrary (79.05) of the controls were smokers whereas (21) were non smokers. Smokers were further categorized as Heavy smokers and Light smokers on the basis of the pack years i.e. the no of cigarettes smoked in a day by an individual multiplied with the total number of

years smoked. Heavy smokers were those with Pack Years > and equal to 20 and the Light smokers were those for whom the values of pack yrs fell below 20. In cases the heavy smokers accounted to (54.08%) and light smokers accounted to (43%), whereas in controls, (55.6%) were heavy and (52.31%) were light smokers.

A vivid contrast between the Smokers and Non Smokers is depicted by this set of population ,where excluding the Smokers with heterozygote genotype, the ones with homozygous mutant genotype shows a highly significant value though it showcases a very low risk of being associated with Lung Cancer. In parallel to this the Heavy smokers with a homozygous mutant genotype (AA) also show a highly statistically significant value along with an incredible protective effect towards the propensity of acquiring the disease (OR=0.21, 95%CI=0.05-0.80; p= 0.022).The light smokers fail to show any significance pertaining to Lung Cancer. Due to low prevalence of homozygous mutant genotype, a joint effect of heterozygote genotype along with the homozygous mutant genotype is evaluated. The joint genotype (AA+GA) shows highly significant values in case of smokers (OR=0.46; 95%CI=0.27-0.79; p=0.005) and heavy smokers (OR=0.38; 95%CI=0.17-0.81; p=0.02) in contrast to Light smokers and Non smokers. Smokers with ***AhR Arg^{554}Lys*** polymorphism showcase a protective effect towards the risk of acquiring Lung Cancer.

5.14 Cumulative effect of the polymorphism in AhR rs10250822(C/T) and Ahr rs 2282885(C/T) and its association with the risk of acquiring Lung Cancer

Table 5.13: Distribution showing the combinatorial effect of AhR rs10250822 and Ahr rs 2282885 polymorphism on the vulnerability towards Lung Cancer

AhR rs 10250822* and *Ahr rs 2282885	**Cases, n (%) N = 69**	**Controls, n (%) N = 91**	**Adjusted OR (95% CI)**	***p*value**
T/T + T/T	43 (62)	81 (89)	1.00 (Reference)	
C/T + C/T	26 (38)	10 (11)	5.87 (2.46-13.98)	**0.0001**

Studies show that polymorphism in *AhR* in simultaneous combinations with each other show an enhanced effect towards the incidence of acquiring Lung Cancer.The two polymorphism effectively enhance the incidence levels. This is validated by studying the combinatorial effects of the two polymorphism in the same gene by skimming common genotypes ,which in this case is heterozygotes(CT) and compairing it with the reference wild genotype (TT).

From the above table we can see that of the total 401 cases from the respective polymorphism study only 69(17.20) cases and similarly 91(22.69) controls out of the 401 controls were found to be those which had a common heterozygote genotype (CT). This was compared with that of the reference genotype.

The table showcases a stupendously high significance between the two polymorphism in ***AhR rs10250822 and rs2282885***. This shows that an individual with a heterozygote genotype (CT) having polymorphism for *Ahr* gene has developed a fivefold risk towards acquiring the endangerment (OR=5.87;95%CI=2.46-13.98;p= 0.0001)

Procastination, when is not deliberate, leads to an observation and eventually a science evolves, this is well proved by the above observation. No prior work has been done to study the combinatorial effects of the polymorphism in the same gene in case of *AhR*.

5.15 Cumulative effect of combined polymorphism in Ahr rs 2282885(C/T) and AhR rs2066853(G/A) and its association with the propensity towards Lung Cancer

Table5.14: Distribution showing the combinatorial effect of AhR rs2282885 and Ahr rs 2066853 polymorphism on the vulnerability towards Lung Cancer

AhR rs 2282885 and AhR Arg 554 Lys rs 2066853	**Cases, n (%) N = 86**	**Controls, n (%) N =83**	**Adjusted OR (95% CI)**	***p* value**
T/T + G/G	83 (96)	76 (92)	1.00 (Reference)	
C/T + G/A	3 (4)	7 (8)	0.47 (0.11-2.03)	0.31

Combined analysis of the two SNP s to evaluate the potentially modifying effect of combined genotypes towards the propensity of acquiring Lung cancer was done. An insignificant association was found between the two SNP s towards the risk of developing Lung Cancer as is reflected by the values (OR=0.47;95%CI=0.11-2.03;p = 0.31). Due to absence of eminent homozygous mutants for the two SNP s we could not evaluate the cumulative effect of the combined polymorphism in the homozygous mutant genotype .

5.16 Cumulative effect of combined polymorphism in Ahr rs 10250822(C/T) and AhR rs 2066853(G/A) and its association with the propensity towards Lung Cancer

Table 5.15: Distribution showing the combinatorial effect of AhR rs10250822 and Ahr rs 2066853 polymorphism on the vulnerability towards Lung Cancer				
AhR rs 10250822(C/T)_and AhRArg554Lysrs2066853(G/A)	**Cases, n (%) N = 59**	**Controls, n (%) N =67**	**Adjusted OR (95% CI) [b]**	***p* value**
T/T + G/G	47 (80)	61 (91)	1.00 (Reference)	
C/T + G/A	12 (20)	6 (9)	2.99 (0.99-8.9)	**0.05**

The combined presence of the two polymorphisms showed a complete threefold increase in the vulnerability of an individual towards developing the Lung Cancer. The combinatorial effect also showcased a fairly significant statistical value on the applying of statistical parameters and testing (OR=2.99; 95%CI=0.99-8.9; p= 0.05). Due to the low prevalence of the homozygous mutants in the study population, the combinations were restrained to the heterozygote genotype.

5.17 Cumulative effect of joint polymorphism AhR rs 10252882 and AhR rs 2282885 on Lung Cancer vulnerability, subgrouped on the basis of histology

Table 5.16: *Distribution showing the combinatorial effect of AhR rs10252882 and Ahr rs2282885 polymorphism on the subgroups of histological forms of Lung cancer*				
***AhR rs 10252882 and AhR rs 2282885* SQCC)**	**Cases, n (%) N =24**	**Controls, n (%) N = 91**	**Adjusted OR (95% CI)**	***p* value**
T/T + T/T	16 (67)	81 (89)	1.00 (Reference)	
C/T + C/T	8 (33)	10 (11)	3.98 (1.28-12.41)	**0.01**
***AhR rs 10252882 and AhR rs 2282885* (ADCC)**	**Cases, n (%) N =25**	**Controls, n (%) N = 91**	**Adjusted OR (95% CI)**	***p* value**
T/T + T/T	15 (60)	81 (89)	1.00 (Reference)	
C/T + C/T	10 (40)	10 (11)	5.57 (1.82-17.08)	**0.0026**
***AhR rs 10252882 and AhR rs 2282885* (SCLC)**	**Cases, n (%) N = 18**	**Controls, n (%) N =91**	**Adjusted OR (95% CI)**	***p* value**
T/T + T/T	12 (67)	81 (89)	1.00 (Reference)	
C/T + C/T	6 (33)	10 (11)	3.78 (1.06-13.43)	**0.039**

Simultaneous presence of ***AhR rs 10252882***(C/T) and ***AhR rs 2282885*** (C/T) polymorphisms showed exorbitantly elevated statistically significant values. Amongst all the histological forms of Lung cancer, individuals diagnosed with ADCC showed fivefold elevated vulnerability towards Lung cancer (OR=5.57, 95%CI=1.82-17.08; p=0.0026). In case of SQCC, the susceptibility dipped by one fold to being fourfold (0R=3.98; 95%CI=1.28-12.41;p=0.011).In parallel a slight dip was witnessed in the vulnerability towards Lung cancer by individuals diagnosed with SCLC (OR=3.78;95%CI=1.06-13.43;p =0.039).The combinatorial effect for these SNPs is immensely important and must be taken note of with respect to the susceptibility of an individual with either of the three histological forms showing a polymorphism for these SNPs towards acquiring Lung cancer.

5.18 Combinatorial effect of polymorphism AhR rs10252882 and AhR rs 2066853 on Lung Cancer vulnerability, subgrouped on the basis of Histology

Table 5.17: ***Distribution showing the combinatorial effect of AhR rs10252882 and Ahr rs2066853 polymorphism on the subgroups of histological forms of Lung cancer***

AhR rs10252882 & AhR rs 2066853 (SQCC)	**Cases, n (%) N = 28**	**Controls, n (%) N = 67**	**Adjusted OR (95% CI) [b]**	**p value**
T/T + G/G	22 (79)	61 (91)	1.00 (Reference)	
C/T + G/A	6 (21)	6 (9)	3.11 (0.85-11.37)	0.085
AhR rs10252882 & AhR Arg 554Lys rs 2066853 (ADCC)	**Cases, n (%) N =15**	**Controls, n (%) N =67**	**Adjusted OR (95% CI) [b]**	**Pvalue**
T/T + G/G	13 (87)	61 (91)	1.00 (Reference)	
C/T + G/A	2 (13)	6 (9)	2.46 (0.416-14.62)	0.32
AhR rs10252882 & AhR Arg 554Lys rs 2066853 (SCLC)	**Cases, n (%) N =16**	**Controls, n (%) N =67**	**Adjusted OR (95% CI) [b]**	**p value**
T/T + G/G	12 (75)	61 (91)	1.00 (Reference)	
C/T + G/A	4 (25)	6 (9)	3.31 (0.68-16.05)	0.13

The combined presence of ***AhR rs10252882 C/T*** and ***AhR rs2066853 G/A*** polymorphisms showed increased risk for SCLC as well as SQCC. Three fold elevated risk was observed for SCLC, marginally higher than that for SQCC when the combined heterozygote and mutant genotype were considered in ***AhR rs10252882 C/T*** and ***AhR rs2066853 G/A*** combination though statistically the values were insignificant for these combinations. In case of ADCC patients with a heterozygote and mutant combined genotype, two fold risk was witnessed towards Lung Cancer.

Discussions

Our findings regarding the polymorphism in the variant of ***AhR rs 10250822*** and ***AhR rs 2282885*** are very much in sync with those seen in a study done in coke exposed workers. In the coke exposed workers increased levels of hydroxypyrene was accounted in the urine which also is found to be a metabolite of PAHs. A detoxified byproduct of PAH, indirectly is involved in the onset or inducing the CYP1A1 activity and hence the process of metabolism. This study provides us with an insight towards the probable chances for a coke exposed worker to carry higher chances for involving the detoxifiying enzymes than those who are not undergoing such occupational hazards. (Antilla *et al.,* 2001) The ***Ahr*** mediated signalling is seen to participate in the regulation of PAH metabolic activation and contribute to susceptibility of PAH exposure (Bin et al 2008).

In conclusion, the alteration in PAH metabolic pathway may interact with environmental exposure and contribute towards tumorigenesis. Polymorphism in ***rs10252882*** donot showcase an association towards the risk of acquiring infertility in males as was stated in Iranian population (Berwick *et al*., 2004). This SNP continues to be unexplored vividly by researchers and we thus donot have enough instances to validate our work with.

Study in the Chinese population suggests that an increase in the pack years of smokers simultaneously increased the OR and validated the hypothesis that suggests that as the number of pack years increase the susceptibility of an individual towards acquiring lung cancer increases (OR=3.36; 95%CI=1.07-10.55; p=0.035). It showcased a similar trend in the sub grouping of the population in smokers and non smokers followed by light and heavy smokers on the basis of pack years. Our study in the north Indian population falls very much in line with the former which also hold good for ***AhR rs2066853*** and proves a significant association of cumulative cigarette smoking on the susceptibility towards lung cancer. As per the study in Chinese population, individuals with SQCC showed the maximum response towards the risk of lung cancer by displaying high ORs. In our study it was seen that the individuals diagnosed with SQCC showed statistically significant values, which confirms the findings reported by the study in Chinese population as well.

Another study suggests a significant effect of ***AhR rs2282885*** and ***rs2066853*** polymorphism on the CYP1A2 inducibility, which confirmed the involvement of the AhR mediated pathway (Nebert et al 2004). It was seen that if the individual was exposed to more smoke or was a

heavy smoker its inbuilt possessed capacity to detoxify the inhaled carcinogen would increase leading to an enhanced CYP1A2 activity. (Eap *et al*., 2013)

AhR rs 2066853 being nonsynonymous is thought to play a vital role in the area of proteins crucial to Enzyme activity (Smart et al 2000).A complete contradiction comes from similar epidemiological studies done in Korean, Japanese and French populations, where no association was seen between the polymorphism at ***AhR Arg ^{554} Lys rs 2066853*** with that of lung cancer and French populations , where no association was seen between the polymorphism at ***AhR Arg 554 Lys rs 2066853*** with that of Lung cancer

The possible explanation of the various finding regarding the ***Ahr rs 2282885*** polymorphism is that it happens to be located in the intronic region of the *AhR* gene, wherein the gene expression is dysregulated leading to the decrease or increase in the gene transcription levels and it has been seen to influence the proper splicing of RNA leading to alternatively spliced RNA variants (Hirose *et al.,* 2008). For e.g. an intronic region in the *AhR* gene which alters RNA splicing at either 38 or 43 amino acid near the end of the carboxy terminus results in the deletion from the Transactivation Domain of the Receptor, therefore these intronic mutations are accountable for differences in sensitivity to the xenobiotic induced toxicity.

In similar study populations, SNPS ***rs 2158041*** and ***rs 7811989*** both residing in the intronic regions were reported to be associated with higher risk of Lung cancer. Another study reports the association of *AhR* rs 2282885 with the inducibility of CYP1A2 gene, which validates the involvement of *AhR* mediated pathway and also a higher risk towards Lung Cancer.

In an Iranian population study, it was seen that *AhR* rs2282885 SNP with a homozygous mutant genotype showed a threefold increase to acquire infertility in males. Literature backs the fact by holding the release of PAHs from the diesel exhaust responsible for the decreased sperm production due to perturbed spermatogenesis and testicular functions (Izawa *et al.,* 2007)

As of now no vivid studies on ***rs 2282885*** has been reported or seen in association with the risk towards acquiring Lung cancer however this SNP shows a strong association with Idiopathic Male factor infertility which is a direct repercussion of differential sensitivity towards TCDD induced carcinogenesis.

Conclusions

In the three *AhR* variants (*AhR* rs10252882, rs2066853, rs2282885) statistically significant values (High ORs validate the association of these SNPs with the risk of Lung cancer.

On subgrouping the population on the basis of histologically different forms of Lung cancer *AhR* rs10252882 showed stupendous results. Patients from the ADCC subgroup showed maximum risk towards Lung cancer. Trends changed for rs2282885 where patients with SQCC showed the highest association with the endangerment. rs2066853 projected a protective effect towards the risk of Lung cancer though it showed statistically significant values.

The results validate the interrogation about Cigarette smoking association with Lung cancer. Heavy smokers showcased a higher risk towards the advent of Lung cancer in case of rs10252882, whereas 2282885 showed increased risk only towards lung cancer only when combinations were considered. rs2066853 failed to show a risk towards the disease and showed a protective effect.

Combinatorial study of rs10252882 and 2282885 showed exorbitantly high OR values suggesting a fivefold increase in the risk towards acquiring lung cancer in case of simultaneous polymorphism in the *AhR* gene. In case of rs10252882 and rs2066853 also a threefold risk towards lung cancer was seen.

Histological subgrouping in combination studies showed the highest risk for patients with ADCC (fivefold) followed by SQCC(fourfold) and SCLC (threefold) in case of rs10252882 and rs2282885. On the contrast in the combination of rs 10252882 and rs2066853, maximum risk was for those falling in the subgroup of SQCC(threefold) followed by SCLC which was marginally less than the former and ones with ADCC showed a twofold elevated risk towards lung cancer.

Bibliography

- Kawajiri K, Watanabe J, Eguchi H, Nakachi K, Kiyohara C, Hayashi S.:*Polymorphisms of human Ah receptor gene are not involved in lung canc*er. Pharmacogenetics 1995; 5:151–158.

- Cauchi S, Stucker I, Solas C, Laurent-Puig P, Cenee S, Hemon D, et al. *Polymorphisms of human aryl hydrocarbon receptor (AhR) gene in a French population: relationship with CYP1A1 inducibility and lung cancer*. Carcinogenesis 2001; 22:1819–1824.

- Kim JH, Kim H, Lee KY, Kang JW, Lee KH, Park SY, et al. *Aryl hydrocarbon receptor gene polymorphisms affect lung cancer risk*. Lung Cancer 2007; 56:9–15.

- Harper PA, Wong JY, Lam MS, Okey AB. *Polymorphisms in the human AH receptor.* Journal of Chemico Biological Interact 2002; 141:161–187.

- Bjeldanes LF, Kim JY ,Grose Kr; *Aromatic hydrocarbon responsiveness receptor agonists generated from I3C invitro and invivo*(1991)

- Denison, M.S. and Nagy S.R., *Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals.* Annual Review of Pharmacology and Toxicology, 2003. **43**: p. 309-34

- Perdew G, Craig, Ramadoss*: role of aryl hydrocarbon receptor in drug metabolism:2005*

- Petrulis Jr, Perdew *GH :the role of chaperone proteins in the aryl hydrocarbon receptor core complex*. Journal of Chemico Biological Interact. 2002

- Morrow D, Qin C,Smith R :*Aryl hydrocarbon receptor mediated inhibition of LNCaP prostrate cancer cell growth and hormone induced Transactivation*. Journal of Steroid Biochemistry and Molecular Biology 2004

- Dzeletovic N, MCguire J, Daujat M *et al* :*Regulation of dioxin receptor function by omeprazole.* Journal of Biological Chemitry. (1997)272:12705-12713

- Ding X, Kaminsky LS: *Human extrahepatic cytochromes P450: function in xenobiotic metabolism and tissue-selective chemical toxicity in the respiratory and gastrointestinal tracts.* Annual Review of Pharmacology and Toxicology. (2003) 43: 149-173

- Sachse C, Brockmoller J, Bauer S, Roots I: *Functional significance of a C->A polymorphism in intron 1 of the cytochrome P450 CYP1A2 gene tested with caffeine.* British Journal of Clinical Pharmacology (1999)

- Nakajima M, Yokoi T, Mizutani M, Kinoshita M, Funayama M, Kamataki T: *Genetic polymorphism in the 52-flanking region of human CYP1A2 gene: effect on the CYP1A2 inducibility in humans*. Journal of Biochemistry (Tokyo) (1999) 125:803-808

- Schlezinger JJ, Liu D, Farago M, Seldin DC, Belguise K, Sonenshein GE, Sherr DH: *A role for the aryl hydrocarbon receptor in mammary gland tumorigenesis.* Journal of Biological Chemistry (2006), 387:1175-1187

- Kim JH, Kim H, Lee KY, Kang JW, Park SY, Yoon HI, Jheon SH, Sung SW, Hong YC: *Aryl hydrocarbon receptor gene polymorphisms affect lung cancer risk*. Lung Cancer 2007, 56:9-15

- Lin P, Chang H, Tsai WT, Wu MH, Liao YS, Chen JT, Su JM: *Overexpression of aryl hydrocarbon receptor in human lung carcinomas.* Toxicology Pathology 2003, 31:22-30

- Rowlands CJ, Staskal DF, Gollapudi B, Budinsky R. *The human AHR:identification of single nucleotide polymorphisms from six ethnic populations.* Pharmacogenetics Genomics. 2010;20(5):283-290

- Inouye K., Pan X., Imai N., Ito T., Takei T., Tohyama C., and Nohara K., *T cell-derived IL-5 production is a sensitive target of 2,3,7,8-tetrachlorodibenzo-p-dioxin* (TCDD). Chemosphere, 2005. **60**(7): p. 907-13.

- Ito T, Inouye K, Fujimaki H, Tohyama C, and Nohara K., *Mechanism of TCDD-induced suppression of antibody production: effect on T cell-derived 182 cytokine production in the primary immune reaction of mice*. Journal of Toxicological Sciences, 2002. **70**(1): p. 46-54.

- Amazit L, Pasini L, Szafran A.T., Berno V, Wu R.C, Mielke M., Jones E.D., Mancini M.G., Hinojos C.A., O'Malley B.W., and Mancini M.A., *Regulation of SRC-3 intercompartmental dynamics by estrogen receptor and phosphorylation*. Journal of Molecular Cell Biology, 2007. **27**(19): p. 6913-32.

- Nambu J.R., Lewis J.O., Wharton K.A Jr., and Crews S.T., *The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development.* Cell, 1991. **67**(6): p. 1157-67.

- Hankinson O., *The role of the aryl hydrocarbon receptor nuclear translocator protein in aryl hydrocarbon receptor action*. Trends in Endocrinology and Metabolism, 1994. **5**(6): p. 240-4.

- Skupinska K., Misiewicz I, and Kasprzycka-Guttman T.,*Polycyclic aromatic hydrocarbons: physicochemical properties, environmental appearance and impact on living organisms.* Acta Poloniae Pharmaceutica, 2004. **61**(3): p. 233-40.

- Goulaouic S., Foucaud L., Bennasroune A., Laval-Gilly P., and Falla J., *Effect of polycyclic aromatic hydrocarbons and carbon black particles on pro-inflammatory cytokine secretion: impact of PAH coating onto particles.* Journal of Immunotoxicology, 2008. **5**(3): p. 337-45.

- Shimizu Y., Nakatsuru Y., Ichinose M., Takahashi Y., Kume H., Mimura J., Fuji-Kuryama Y., and T. I., *Benzo(a)pyrene carcinogenicity is lost in mice lacking the aryl hydrocarbon receptor.* Proccedings of National Academy of Sciences. USA, 2000. **97**: p. 779-782.

- Boucher O., Muckle G., and Bastien C.H., *Prenatal exposure to polychlorinated biphenyls: a neuropsychologic analysis.* Environment and Health Perspect, 2009. **117**(1): p. 7-16.

- Crawford R.B., Holsapple M.P., and Kaminski N.E., *Leukocyte activation induces aryl hydrocarbon receptor up-regulation, DNA binding, and increased Cyp1a1 expression in the absence of exogenous ligand.* Molecular Pharmacology, 1997. **52**(6): p. 921-7.

- Quattrochi L.C. and Tukey R.H., *Nuclear uptake of the Ah (dioxin) receptor in response to omeprazole: transcriptional activation of the human CYP1A1 gene.* Molecular Pharmacology, 1993. **43**(4): p. 504-8.

- Murray I.A. and Perdew G.H., *Omeprazole stimulates the induction of human insulin-like growth factor binding protein-1 through aryl hydrocarbon receptor activation.* Journal of Pharmacology and Experimental Therapeutics, 2008. **324**(3): p. 1102-10

- Carver L.A., LaPres J.J., Jain S., Dunham E.E., and Bradfield C.A., *Characterization of the Ah receptor-associated protein, ARA9.* Journal of Biological Chemistry, 1998. **273**(50): p. 33580-7.

- Carver L.A. and Bradfield C.A., *Ligand-dependent interaction of the aryl hydrocarbon receptor with a novel immunophilin homolog in vivo.* Journal of Biological Chemistry, 1997. **272**(17): p. 11452-6.

- Meyer B.K., Pray-Grant M.G., Vanden Heuvel J.P., and Perdew G.H., *Hepatitis B virus X-associated protein 2 is a subunit of the unliganded aryl hydrocarbon receptor core complex and exhibits transcriptional enhancer activity.* Molecular Cell Biology, 1998. **18**(2): p. 978-88.

- Petrulis J.R., Kusnadi A., Ramadoss P., Hollingshead B., and Perdew G.H., *The hsp90 Co-chaperone XAP2 alters importin beta recognition of the bipartite nuclear localization signal of the Ah receptor and represses transcriptional activity.* Journal of Biological Chemistry, 2003. **278**(4): p. 2677-85.

- Silkworth J.B., Lipinskas T., and Stoner C.R., *Immunosuppressive potential of several polycyclic aromatic hydrocarbons (PAHs) found at a Superfund site: new model used to evaluate additive interactions between benzo[a]pyrene and TCDD.* Toxicology, 1995. **105**(2-3): p. 375-86.

- Ramadoss, P. and Perdew G.H., *The transactivation domain of the Ah receptor is a key determinant of cellular localization and ligand-independent nucleocytoplasmic shuttling properties.* Biochemistry, 2005. **44**(33): p. 11148-59.

- Perdew G.H. and Bradfield C.A., *Mapping the 90 kDa heat shock protein binding region of the Ah receptor.* International Journal of Biochemistry and Molecular Biology , 1996. **39**(3): p. 589-93.

- Meyer B K, Perdew G H, Biochemistry 38 (1999) 8907–8917.

- Kazlauskas, L. Poellinger, I. Pongratz, Journal of Biological Chemistry. 274 (1999)13519–13524.

- Kimbrough, R.D. and Krouskas C.A., *Human exposure to polychlorinated biphenyls and health effects: a critical synopsis.* Toxicology Review, 2003. **22**(4): p. 217-33

- Thurmond, T.S., Staples J.E., Silverstone A.E., and Gasiewicz T.A., *The aryl hydrocarbon receptor has a role in the in vivo maturation of murine bone marrow B lymphocytes and their response to 2,3,7,8-tetrachlorodibenzo-p-dioxin.* Toxicological Applications of Pharmacology, 2000. **165**(3): p. 227-36

- Schneider, D., Manzan M.A., Yoo B.S., Crawford R.B., and Kaminski N., *Involvement of Blimp-1 and AP-1 dysregulation in the 2,3,7,8-Tetrachlorodibenzo-p-dioxin-mediated suppression of the IgM response by B cells.* Toxicological Sciences, 2009. **108**(2): p. 377-88

- Fujimaki H., Nohara K., Kobayashi T., Suzuki K., Eguchi-Kasai K., Tsukumo S., Kijima M., and Tohyama C., *Effect of a single oral dose of 2,3,7,8-tetrachlorodibenzo-*

p-dioxin on immune function in male NC/Nga mice. Toxicological Sciences, 2002. **66**(1): p. 117-24.

- Mocarelli P., Gerthoux P.M., Ferrari E., Patterson J., Donald G, Kieszak S.M., Brambilla P., Vincoli N., Signorini S., and Tramacere P., *Paternal concentrations of dioxin and sex ratio of offspring.* The Lancet, 2000. **355**(9218): p. 1858-1863.

- Moses M., Lilis R., Crow K.D., Thornton J., Fischbein A., Anderson H.A., and Selikoff I.J., *Health status of workers with past exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin in the manufacture of 2,4,5-trichlorophenoxyacetic acid: comparison of findings with and without chloracne.* American Journal of Internal Medicine, 1984. **5**(3): p. 161-82.

- Thackaberry E.A., Gabaldon D.M., Walker M.K., and Smith S.M., *Aryl hydrocarbon receptor null mice develop cardiac hypertrophy and increased hypoxia-inducible factor-1alpha in the absence of cardiac hypoxia.* Journal of Cardiovascular Toxicology, 2002. **2**(4): p. 263-74.

- Bunge, M.K., Glover E., Moran S.M., Walisser J.A., Lahvis G.P., Hsu E.L., and Bradfield C.A., *Abnormal liver development and resistance to 2,3,7,8-tetrachlorodibenzo-p-dioxin toxicity in mice carrying a mutation in the DNA-binding domain of the aryl hydrocarbon receptor*. Toxicological Sciences, 2008. **106**(1): p. 83-92.

- Boffeta P, Molecular epidemiology. Journal of Internal Medicine 2000.248:447-454

- Jones L.C. and Whitlock Jr J.P., *Dioxin-inducible transactivation in a chromosomal setting. Analysis of the acidic domain of the Ah receptor.* Journal of Biological Chemistry, 2001. **276**(27): p. 25037-42

- Kawajiri K., Watanabe J., Eguchi H., Nakachi K., Kiyohara C., and Hayashi S., *Polymorphisms of human Ah receptor gene are not involved in lung cancer.* Pharmacogenetics, 1995. **5**(3): p. 151-8