Prognostic significance of *ABCB1* (MDR-1) gene variants (C1236T and G2677T) in lung cancer patients treated with platinum-based doublet chemotherapy

A Dissertation submitted in partial fulfilment of the requirement for

the award of degree of

Master of Science

in

Biotechnology

Submitted by

Kirti Rai

Roll no.: 301601015

Under the guidance of

Dr. Siddharth Sharma

Associate Professor

Department of Biotechnology

Thapar Institute of Engineering and Technology

(Deemed to be University)

Patiala, Punjab

July, 2018

DECLARATION

I, the undersigned, hereby declare that the work presented in the M.Sc. dissertation entitled "**Prognostic significance of *ABCB1* (MDR-1) gene variants (C1236T and G2677T) in lung cancer patients treated with platinum-based doublet chemotherapy**" has been carried out by me under the supervision and guidance of **Dr. Siddharth Sharma**, Department of Biotechnology, Thapar Institute of Engineering and Technology, 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.

Kirti Rai
(Reg. No. 301601015)

CERTIFICATE

This is to certify that the dissertation entitled "**Prognostic significance of *ABCB1* (MDR-1) gene variants (C1236T and G2677T) in lung cancer patients treated with platinum-based doublet chemotherapy**" submitted for the degree of Master of Science in the subject of Biotechnology, Thapar Institute of Engineering and Technology (TIET), Patiala is a bonafide work carried out by Ms. Kirti Rai under the supervision of Dr. Siddharth Sharma, Ph.D., Associate Professor, Thapar Institute of Engineering and Technology (TIET), Patiala and that no part of this work has been submitted for any other degree.

The assistance and help received during the course of investigation has been fully acknowledged.

Dr Siddharth Sharma
Associate Professor
Department of Biotechnology
Thapar Institute of Engineering Technology, Patiala

Acknowledgement

Firstly, I would like to express my sincere gratitude to **Dr. Siddharth Sharma**, Associate Professor, Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala for guiding me through the entire project. His constant motivation and brilliant supervision gave me an opportunity to learn a lot and simultaneously enjoy my work.

I express my humble gratitude to all the faculty members of Department of Biotechnology, Thapar Institute of Engineering and Technology, for their constant support, help and encouragement.

A special thanks to **Mr. Kinshuk Sharma** and **Ms. Harleen Walia**, Department of Biotechnology, Thapar Institute of Engineering and Technology, Patiala for their valuable guidance, expedient advice and affection during the work. I am also indebted to one and all, who have, directly or indirectly, lent their helping hand during the course.

Furthermore I am thankful to **Anu Gupta** and **Divya Ahluwalia**, Thapar Institute of Engineering and Technology, Patiala for their incomparable help.

I extend my sincere and heartfelt gratitude to all lab assistants and non-teaching staff. Finally, yet importantly, I would like to express my heartfelt thanks to my beloved parents for their blessings.

Above all I remember with gratitude the grace of Almighty for the successful completion of this work.

Kirti Rai

Roll no.: 301601015

Abbreviations

ABC	ATP-binding cassette
ABCB1	ATP-binding cassette subfamily B member 1
ABCC1	ATP-binding cassette subfamily C member 1
ABCG2	ATP-binding cassette subfamily G member 2
ADCC	Adenocarcinoma
AOR	Adjusted odds ratio
ASCO	American Society of Clinical Oncology
ATP	Adenosine triphosphate
BCRP	Breast cancer resistance protein
BSA	Bovine serum albumin
CR	Complete response
CTR1	Copper transporter 1
dNTPs	Deoxynucleotide triphosphate
ECF	Energy coupling factor
ECOG	Eastern Cooperative Oncology Group
EDTA	Ethylenediaminetetraacetic acid
EtBr	Ethidium Bromide
GSH	Glutathione
HCl	Hydrochloric acid
HR	Hazard ration
KPS	Karnofsky Performance Status
LCC	Large cell carcinoma
LC	Lung cancer
MDR1	Multidrug resistance 1
$MgCl_2$	Magnesium Chloride
MRP1	Multidrug resistance associated protein
MST	Median survival time
NaCl	Sodium Chloride
NBD	Nucleotide-binding domain
NSCLC	Non-small cell lung cancer
nsSNP	Non-synonymous single nucleotide polymorphism

OR	Odds ratio
OS	Overall survival
PAGE	Polyacrylamide gel electrophoresis
PCI	Phenol Chloroform Isoamyl alcohol
PCR	Polymerase chain reaction
PD	Progressive disease
PGIMER	Postgraduate Institute of Medical Education and Research
P-gp	Permeability-glycoprotein
Polyphen2	Polymorphism phenotyping v2
PR	Partial response
PROVEAN	Protein Variation Effect Analyzer
Pt	Platinum
RFLP	Restriction fragment lemgth polymorphism
SCLC	Small-cell lung cancer
SD	Stable disease
SDS	Sodium dodecyl sulphate
SIFT	Separating intolerant from tolerant
SNP	Single nucleotide polymorphism
SQCC	Squamous cell carcinoma
TAE	Tris acetate EDTA
Taq	*Thermus aquaticus*
TBE	Tris borate EDTA
TE	Tris-EDTA
TNM	Tumor, Node, Metastasis
TMD	Transmembrane binding domain
WHO	World health organization

Index

List of Tables

List of Figures

Abstract

ABCB1 gene encodes P-glycoprotein which is involved in the transport of substrates in and out of the cell. However, some of the substrates of P-glycoprotein include numerous structurally unrelated chemotherapeutic drugs which play a major role in drug resistance. Also, single nucleotide polymorphisms of *ABCB1* gene are known to affect the structure and function of P-glycoprotein thus affecting their efflux system. The most common SNPs of *ABCB1* gene that has been widely studied include rs1128503 (C1236T) and rs2032582 (G2677T). Various studies have evaluated the association of the above mentioned genetic polymorphisms of *ABCB1* gene with overall survival of cancer (colorectal cancer, breast cancer, osteosarcoma, acute myeloid leukemia) on the basis of age, gender, smoking status, performance status and chemotherapy response. However, not much study has been done with regard to *ABCB1* gene polymorphism and lung cancer. The aim of this study is to find the association between genetic polymorphisms rs1128503 and rs2032582*,* individually, with overall survival and treatment outcomes in lung cancer patients receiving platinum-based doublet chemotherapy in North Indian population.

Methods: A total of 192 patients for C1236T and 153 patients for G2677T polymorphisms were evaluated. Genotyping was performed using PCR-RFLP method.

Results: In the present work, we found no significant association of *ABCB1* gene polymorphisms rs1128503 and rs2032582 with overall survival of lung cancer patients. However, statistically significant association of rs2032582 with SQCC was observed in patients harbouring heterozygous genotype GT (HR = 0.13, 95% CI = 0.00063 to 27.19, **P = 0.01**), mutant genotype TT (HR = 0.04, 95% CI = 0.0000071 to 301.98, **P < 0.0001**) and combined genotype (HR = 0.08, 95% CI = 0.000087 to 74.56, **P = 0.001**). Further, association of *ABCB1* gene polymorphisms with chemotherapy response were examined. The results suggested no significant association of these polymorphisms with chemotherapy response.

Keywords: *ABCB1,* C1236T, G2677T, lung cancer, overall survival, chemotherapy response

Chapter 1

Introduction

Cancer is unusual cell growth that arises from heterogeneous set of mutations in somatic cells (Croce *et al*., 2008). Worldwide it is proposed that before 2030 cancer burden will possibly increase to 22.2 million new cancer patients with 13.2 million cancer-related deaths (Bray *et al*., 2012). Among different types of cancer described, malignancy in lung is considered as common type of cancer for both men and women (Zappa *et al.,* 2016). In India the second most frequent disease is cancer which is applicable for higher mortality rate with approximately 0.3 million deaths per year. The accepted form of different malignancy found in Indian population are that of blood, bladder, breast, stomach, lungs, rectum, oesophagus, liver, prostate, cervix, mouth etc. It is seen that incidence percentage of lung malignancy in Indian population are increasing with alarming rate and it is predicted as 1.45 million new cases per year, for which the 5-year survival rate is estimated as 17.8 % (Sharma, 2016; Zappa *et al*., 2016).

Predominantly, malignancy of lung is classified into small cell lung cancer (SCLC) and Non-small cell lung cancer (NSCLC). NSCLC is further subclassified into squamous cell carcinoma, large cell carcinoma and adenocarcinoma among which adenocarcinoma alone accounts for 40% of total lung cancer patients in both smokers and non-smokers (second-hand smoke). SCLC contributes for 15% of lung cancer patients whereas NSCLC being the most accepted form of lung cancer holds 80-85% of total lung cancer patients. According to literature survey, there are many risk factors responsible for making an individual susceptible to lung cancer some of which include tobacco smoking, ionization radiation, asbestos, ancestral history, lung diseases and pollution. (Zappa *et al*., 2016).

The widely accepted treatment for lung cancer includes surgery, chemotherapy and radiotherapy (Shewach *et al*., 2009). With the advances in regimen, doublet platinum-based chemotherapy is given to patient for better survival, but arrival of drug resistance in patients by influx and efflux of chemotherapeutic drugs, cure of lung cancer has become difficult thus affecting the survival of patients. The two (Gly^{412}Gly, rs1128503 and Ala^{893}Ser, rs2032582) major polymorphic

variants of *ABCB1* gene are liable for affecting the survival of patients. Therefore, to trace the function of above polymorphic variants in overall survival of north Indian population, we evaluated the significance of $Gly^{412}Gly$, rs1128503 and $Ala^{893}Ser$, rs2032582 polymorphic variants in the response rate of patients treated with doublet platinum-based chemotherapy.

Chapter 2

Review of Literature

2.1. Cancer

Cancer is a multifactorial disease caused as a result of combination of various molecular alterations in either proto-oncogenes or tumor suppressor genes. An understanding of these molecular abnormalities is essential in order to target better therapeutics and make an impact on cancer (Singh *et al*., 2014). It is known that the accumulation of multiple molecular alterations can lead to cancer some of which include:

1. Mutations in proto-oncogenes leading to autonomous growth signals,
2. Cells being unresponsive to anti-proliferative signals as a consequential effect of mutation in tumor suppressor genes
3. Evasion of cells from apoptotic effects either due to down-regulation of pro-apoptic molecules or up-regulation of anti-apoptotic molecules
4. Telomerases may get activated causing continuous replication
5. uninterrupted angiogenesis and
6. property of contact inhibition is lost leading to tissue invasion and propagation to distant sites (Hanahan *et al*., 2011)

According to GLOBOCAN report of 2012, worldwide there were 8.2 million cancer deaths worldwide, out of which the 45% cancer deaths occurred in less developed regions.
Cancer statistics in India according to GLOBOCAN report, 2012 is mentioned below:

Population in 2012:	1258.3m
People newly diagnosed with cancer (excluding NMSC) / yr:	1,014,900
Age-standardised rate, incidence per 100,000 people/yr:	94.0
Risk of getting cancer before age 75:	10.1%
People dying from cancer /yr:	682,800

The estimated incidence and mortality rates in men is highest for lung cancer, accounting for 16.8% and 23.6% respectively of total with 8.3% of 5 year prevalence whereas in case of women it ranks fourth followed by cancers of breast, colorectum and cervix uteri (GLOBOCAN report, 2012).

2.2. Overview of Lung cancer

The term lung cancer indicates the growth of abnormal cells lining the air passages of lung tissue, where the cells repeatedly divide to form a tumor. Lung cancer is the leading cause of cancer related deaths worldwide. According to Jama oncology report, 2015, there were an estimated 1.8 million incident cases and 1.6 million death cases of Tracheal, bronchus and lung cancer in the year 2013.

2.2.1 Risk factors associated with development of lung cancer

a) Smoking

Smoking is the cause approximately 80% of all lung cancer deaths. It is the major preventable factor responsible for lung cancer (Wangari *et al*., 2013). Cigarette contains a variety of experimentally proven carcinogenesis stimulating polyaromatic hydrocarbons such as benzo(a)pyrene, nickel, ethyl carbamate, cadmium, chromium, arsenic and hydrazine that potentially induce lung tumors. Second hand smoke has been observed as the cause of lung cancer in non-smokers (Zappa *et al.,* 2016*).*

b) Radon gas

Radon which is naturally found in soil and rocks, is a odorless, radioactive gas that is produced on decay of uranium-238 which ultimately turns into alpha and beta emitting particles. The link between radon and lung cancer dates back to 16^{th} century where mine workers were found to be at a greater risk of acquiring respiratory problems. Nowadays, it has been found that exposure to radon gas, both in houses as well as mining areas can damage the lungs which can lead to lung cancer (Singh *et al.,* 2014).

c) **Asbestos**

Asbestos are group of minerals that naturally occur as fibres and are also used in industries. It is labelled as the most potent occupational carcinogen. On inhalation of these particles, they are lodged in the lungs thus damaging it and increasing the risk for lung cancer. Measures must be taken by the government to reduce the use asbestos in manufacturing and other industrial purposes. Other potential carcinogens include mustard gas, chemicals produced from vehicles (diesel exhaust), vinyl chloride, coal products etc (Singh *et al.,* 2014).

d) **Lung diseases**

Patients with lung diseases such as tuberculosis are at greater risk of acquiring lung cancer in areas of the lung scarred from tuberculosis (Ali *et al.,* 2011).

e) **Genetic factors**

Family history of lung cancer also increases the chances for lung cancer. There are certain genes or chromosomes more prone to mutation due to environmental factors that may affect the signalling pathway and make an individual more susceptible to lung cancer (Zappa *et al.,* 2016*)*.

2.2.2 Symptoms of lung cancer

Symptoms of lung cancer are not often felt unless the disease has reached an advanced stage. The most common symptoms observed include: frequent chest pain, coughing, weight loss, haemoptysis, dyspnea, lung infection such as pneumonia, visible change in voice such as hoarseness, joint pain, muscle weakness, drooping eyelids, breast development in men, loss of appetite, chills, swelling of the face, swollen lymph nodes and shortness of breath.

2.2.3 Types of lung cancer

Lung cancer is not a single disease; instead it is classified into its subtypes based on histology as

a) Small cell lung cancer (SCLC)
b) Non-small cell lung cancer (NSCLC)

a) Small cell lung cancer (SCLC)

SCLC comprises of about 15% of all lung cancer cases and is the most aggressive, fast growing among lung cancer subtypes. This carcinoma is strongly associated with cigarette smoking. Its prognosis is generally poor due its aggressive nature and hence more than 95% of SCLC patients eventually die. This subtype can readily spread to other parts of the body. When observed under microscope the SCLC cells appear small and oval shaped. The risk of acquiring SCLC increases with the increase in number of cigarettes smoked per day and duration of smoking **(**Zappa *et al.,* 2016).

b) Non-small cell lung cancer (NSCLC)

NSCLC is the commonest type of lung cancer comprising of about 80-85% of lung cancer cases. Based on the type of cell and location of tumor, it is further classified into three types each of which will be discussed here briefly.

i. **Adenocarcinoma**

Adenocarcinoma comprises of 40% of NSCLC cases and is the most common type of lung cancer in both smokers and non-smokers irrespective of their gender and age. It arises from mucus secreting, type II alveolar cells, lining the airway. This type of tumor usually occurs in the periphery of lung which may be due to addition of filters in cigarettes which prevents larger

particles from entering the lungs. It spreads slowly as compared to other subtypes and is occasionally diagnosed at an early stage.

ii. **Squamous cell carcinoma**

Being strongly associated with cigarette smoking, unlike adenocarcinoma, this subtype comprises of about 25-30% of all lung cancer cases. It arises from the squamous cells lining the epithelial airway in the bronchial tubes.

iii. **Large cell carcinoma**

Large cell carcinoma accounts for about 5-10% of all lung cancer cases. It is an undifferentiated form of lung cancer. On microscopic examination, large, rounded cells are observed. This subtype is also associated with smoking and has a high tendency to spread tp other parts of the body **(**Zappa *et al.,* 2016).

2.2.4 Treatment of Lung cancer

The various treatment options available for lung cancer depend on the histological type of disease and clinical stage of the disease. For patients at initial stage of NSCLC (Stage I-II), surgery is the vital element, provided they are medically fit for surgical resection (Pallis *et al.,* 2012). Radiotherapy is the choice of treatment for those patients at early stage who refuse to go for surgery, are at an older age or showcase any sort physical malaise. However, the survival rates with radiotherapy are lower when compared to surgery (Rowell *et al.,* 2001). Adjuvant chemotherapy is performed either after surgical resection or radiotherapy (stage II or IIIA). However, meta-analysis study reported 5.4% increase in five year survival rate of patients irrespective of whether adjuvant chemotherapy was conducted after surgical resection alone or after combination of surgery and radiotherapy (Douillard *et al.,* 2006). In patients with locally advanced disease (at stage IIIA or IIIB) when tumors are unresectable, chemoradiotherapy is preferred in which the patient is exposed to concurrent chemotherapy and radiotherapy. Nevertheless, chemoradiotherapy is preferred for patients with good performance status so as to be able to sustain the toxicity level of treatment.

Chemotherapy is the keystone in treatment of metastatic cancer stage IV disease. According to a meta-analysis conducted by NMAC group in 2008, chemotherapy enhanced the overall survival of 9% at 12 months offering an absolute improvement. According to study published by Delbaldo *et al* (2004) and Lilenbaum *et al.,* (2005), combination chemotherapy (doublet) offered superior results in terms of overall survival benefit as compared to single agent treatment. Also, no difference in overall survival was observed in three-drug combinations. According to ASCO (American Society of Clinical

Oncology), cisplatin-based regimens offer exceptional result in terms of overall survival with a minimum of 1-year survival benefit in comparison to platinum-free regimens (Azzoli *et al.,* 2010). Classification of chemotherapeutic agents can be done on the basis of cell cycle phase they are active in as mentioned below (Page *et al,* 2004):

a) **S-phase dependent:** Antimetabolites such as doxorubicin, gemcitabine etc.
b) **M-phase dependent:** Taxanes such as Docetaxel, paclitaxel, Vinca alkaloids such as vincristine, vinblastine and podophyllotoxins such as etoposide
c) **G1-phase dependent:** Asparaginase, corticosteroids
d) **G2-phase dependent:** Irinotecan, topotecan *etc.*

2.3. Platinum-based chemotherapy

Platinum complexes are classified under the class of alkylating agents that impair the cell function by alkylating some most biologically important molecules such as DNA, RNA and proteins. They form covalent bonds with the phosphate, amino, carboxyl or sulfyhydryl groups of these molecules. Cisplatin [PtII $(NH_3)_2Cl_2$], which is a inorganic heavy metal complex and carboplatin, which has similar diamine platinum active moiety, are most commonly used platinum based agents for chemotherapy (Page *et al., 2005)*. However, carboplatin is less toxic when compared with cisplatin as it contains carboxylate group that modulates its water solubility and allows slower hydrolysis to the alkylating platinum complex (Kelland *et al.,* 2007, Page *et al*.,2005). These agents act by forming DNA adducts by producing interstrand and intrastrand DNA crosslinks thereby inhibiting replication of DNA, synthesis of RNA and proteins as shown in figure 2.1 below.

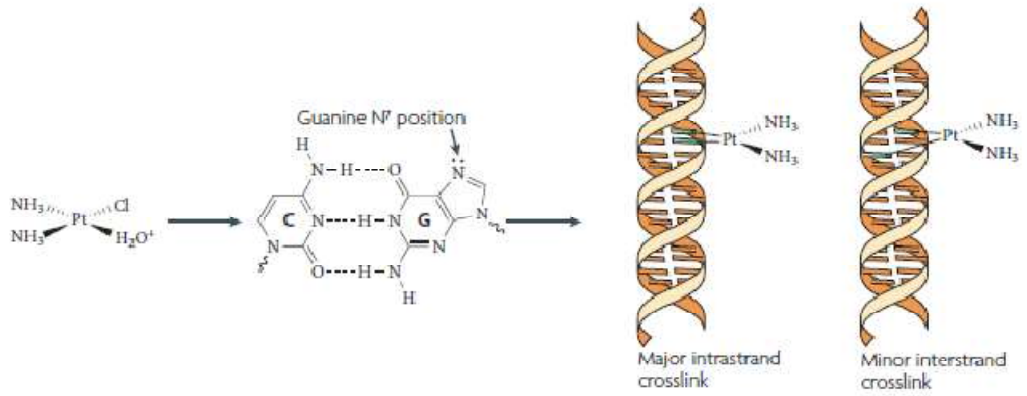


Figure 2.1 Formation of interstrand and intrastrand DNA cross-links with platinum agents (Source: Kelland *et al.,* 2007)

Formation of DNA adducts leads to unwinding and bending in DNA, leading to DNA distortion, the final cellular outcome being apoptosis. These distortions activate several cellular proteins involved in DNA-repair pathways which may play role in platinum-drug resistance. Other reasons involved in platinum-drug resistance will be discussed below.

Taxanes such as docetaxel and paclitaxel are semisynthetic drugs derived from the needles of yew plants. Their structure contains unique 14-member taxane ring. These drugs act by stabilizing microtubular assembly thereby blocking mitosis cell cycle. Docetaxel has more potential as compared with paclitaxel in microtubular assembly and induces apoptosis. Also, Docetaxel has been approved as drug for second-line treatment reflecting an prolonged overall survival (Page *et al.,* 2005; Shepherd *et al.,* 2000).

Irinotecan, derived from camptothecin alkaloid, is a potent topoisomerase I inhibitor thus interrupting with the elongation phase of DNA replication.

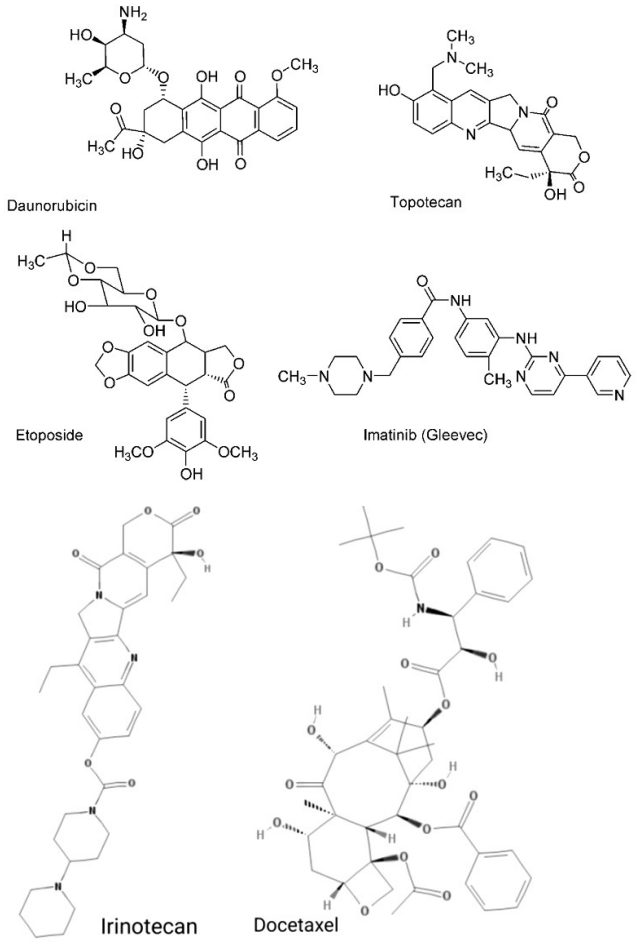


Figure 2.2 Structures of some chemotherapeutic substrates of P-gp (Source: Eckford *et al.,* 2009)

2.3.1 Drug resistance

Treatment of lung cancer often includes combination of chemotherapy drugs such as cisplatin/carboplatin, taxanes, etoposides and vinca alkaloids (Abe *et al*., 1996). However, drug resistance is the major problem which hinders the performance of cancer chemotherapy, which can either be intrinsic or extrinsic (acquired) after drug exposure. Chemoresistance is the consequence of complex cellular mechanism as shown in Figure 2.3 which renders the drug ineffective in cell killing**.** The phenomenon of drug resistance was first reported in the 1940s for alkylating agents, although various combinations of drugs proved effective, resistance remained an issue for newer targets (Cree *et al*., 2017).

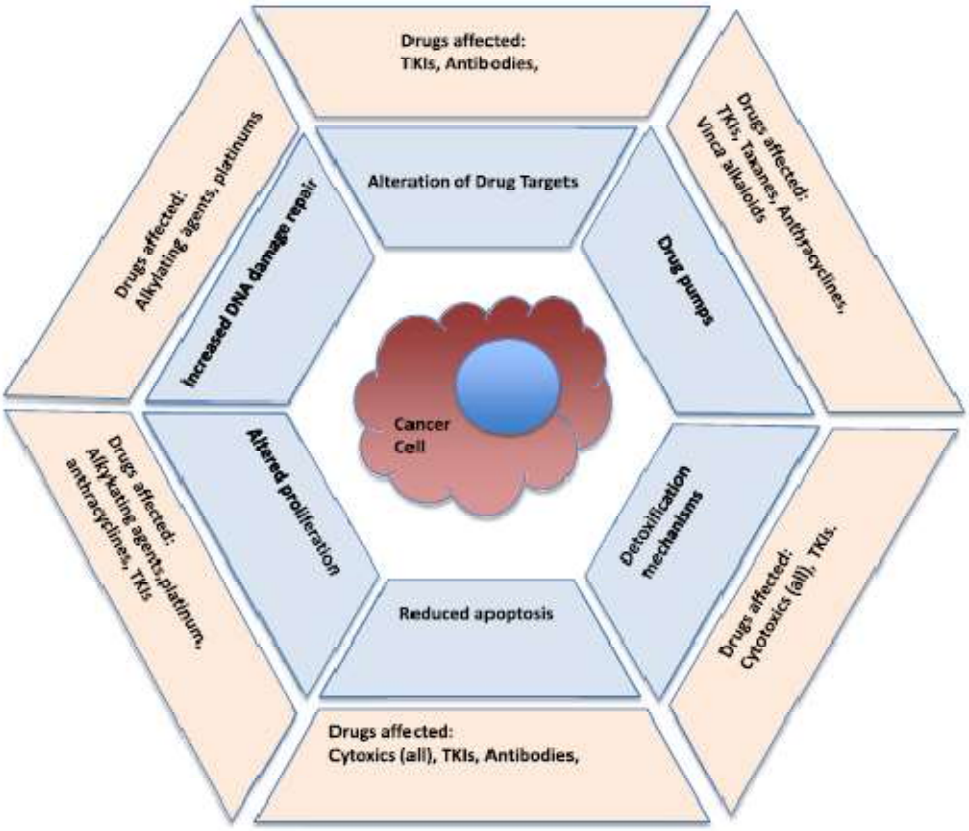


Figure 2.3 Cellular mechanisms of drug resistance (Source: Cree *et al.,* 2017)

One of the common mechanisms of chemoresistance includes over expression of membrane efflux transporters such as ATP-binding cassette (ABC) subfamily in cancer cells (Chen *et al*., 2016, Wangari *et al*., 2013). P-gp (ABCB1), MRP1 (ABCC1) and BCRP (ABCG2) are the three ABC transporter superfamily efflux transporters responsible for Chemoresistance (Chen *et al*., 2016, Sharom 2008). Apart from transporting toxins out of cells, other substrates include peptides, amino

acids, lipids, steroids, nucleotides and bile salts. Classical drug resistance is mediated by P-gp/MDR1. Although in bacteria, MDR1 helps in effluxing out toxins, in case of humans this substrate happen to be anti-cancer drugs.

Cisplatin and carboplatin is used in combination with drugs such as docetaxel, paclitaxel, gemcitabine, irinotecan and vinorelbine in metastatic disease and is much more effective as compared to single platinum-based treatment (Wangari *et al*., 2013). In platinum based chemotherapy, the platinum drugs enter the cell and form DNA adducts that activates a cascade of signaling pathways that promote p53-dependent or p53-independent cell death.

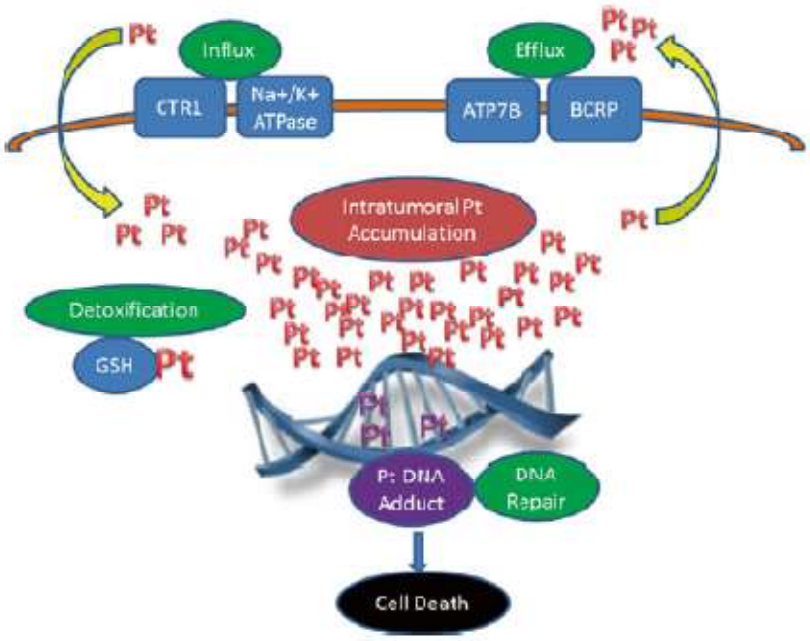


Figure 2.4 Overview of platinum resistance in lung cancer (Pt: Platinum; BCRP: Breast cancer resistance protein; CTR1: Copper transporter 1; GSH: Glutathione) (Source: Kim *et al.,* 2015)

BCRP (Breast cancer resistance protein), Glutathione (GSH) and CTR1 (Copper transporter) are involved in platinum drug resistance. GSH plays a critical role in drug detoxification. It binds to cisplatin thereby reducing the formation of platinum-DNA adducts and enhancing DNA-repair systems.

Among the several reasons responsible for platinum resistance, reduced intracellular accumulation of platinum-based drugs is one of the identified features of cisplatin resistant NSCLC cell lines. A significant relationship has been observed between percent reduction in tumor size and platinum concentrations.

In NSCLC cell lines, increased level of expression of MDR1 protein was found to be associated with increased resistance to chemotherapeutic drugs such actaxanes, vinca alkaloids and etoposide. No significant relationship was found between P-gp expression and intracellular platinum accumulation (Kim *et al., 2015*). P-gp was highly expressed in 61% of chemotherapy treated tumors highlighting the role P-gp in chemoresistance. A negative correlation was found between P-gp expression and clinical outcome in SCLC patients treated with cisplatin-etoposide (Triller *et al., 2006,* Yeh *et al., 2005*).

2.4. ABC transporters

2.4.1 Introduction

ABC (ATP-binding cassette) transporters belong to a large group of membrane-bound transport proteins involved in the influx and efflux of substrates such as ions, peptides, amino acids, xenobiotics *etc.* across the cellular/organellar membrane (ter Beek *et al*., 2014). The free energy produced on hydrolysis of phosphate bond between the γ- and β-phosphate groups of ATP (approximately -50 kJ/mol in most of the cells) is used to either accumulate the substrates in or move out of the cell or cellular compartments (ter Beek *et al*., 2014). The term ABC transporter corresponds to the cassette-like nature of the ATP-binding subunit (Wilkens *et al.*, 2015). In prokaryotes, ABC transporters are confined to the plasma membrane whereas in eukaryotes they are also found in organellar membrane (ter Beek *et al*., 2014).

Humans consist of 49 ABC protein encoding genes and based on their phylogenetic analysis, they are categorized in seven subfamilies designated from ABCA to ABCG (Sharom, 2008, Vasiliou et al., 2009) as mentioned in Table 2.1.

Table 2.1 Subfamilies of human ABC gene (Source: Vasiliou et al., 2009)

Subfamily name	Aliases	No. of genes
ABCA	ABC1	12
ABCB	MDR	11
ABCC	MRP	13
ABCD	ALD	4
ABCE	OABP	1
ABCF	GGN20	3
ABCG	White	5

Total	-	49

ABC transporters are grouped into two types: importers and exporters. Importers are further classified into Type I, type II and type III (also known as energy-coupling factor (ECF) transporters) as mentioned in Figure 2.5. Prokaryotes consist of both importers and exporters unlike eukaryotes which solely consist of exporters. As mammals mostly consist of exporters, importers class will not be discussed here (ter Beek *et al*., 2014, Wilkens *et al*., 2015).

ABC exporters are concerned with the transportation of hydrophobic compounds such as drugs, lipids, cholesterol, larger molecules such as toxins, antibiotics, bacteriocins, hydrolytic enzymes etc. As they are responsible for the export of large variety of drugs out of the cell, they are also termed as multidrug-resistant transporters (ter Beek *et al*., 2014).

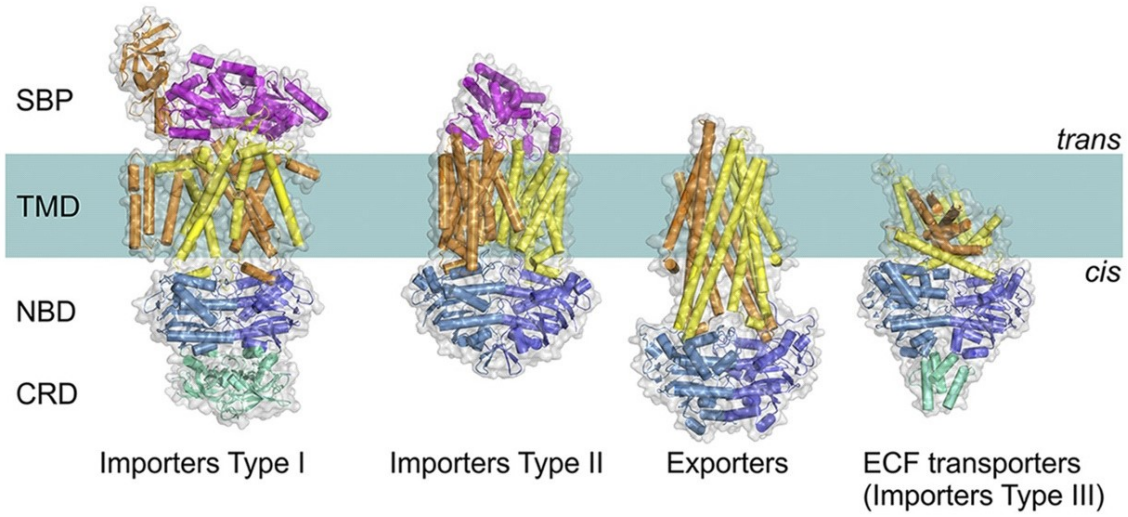


Figure 2.5 Categories of ABC transporters. Type I and Type II importers contain an additional domain called Substrate-binding proteins (SBP) situated either in periplasma or external space. C-terminal regulatory domain (CRD), a domain with regulatory function is constructed in some transporters. ECF: Energy coupling factor (Source: Ter Beek et al., 2014)

2.4.2 Structure of ABC transporters:

All ABC transporters contain a core with similar modular architecture as follows:

- Two Transmembrane binding domains (TMD) or subunits that contains 6-11 membrane spanning α-helices
- Two Nucleotide-binding domains (NBD) termed as ATPases or ABCs

NBDs are the hallmark of the ABC transporter family and are highly conserved in nature. There is 30-50% sequence identity between the NBDs of bacterial and eukaryotic exporters (Wilkens *et al.,* 2015). The existence of set of seven highly conserved motifs in the NBD subunit helps in its identification at the sequence level:

1. **A-loop:** Assists in positioning of ATP by interacting with adenine ring of ATP
2. **Walker A motif:** Phosphate binding loop
3. **Walker B motif:** Helps in coordination of Magnesium ion necessary for ATP hydrolysis
4. **D-loop:** Helps to form the ATP hydrolysis site by changing the geometry of catalytic site
5. **H-loop:** Helps in positioning of the attacking water molecule and Mg^{+2} ion
6. **Q-loop:** This loop consists of a conserved glutamine residue**.** It is the major site of interaction for the TMDs. It forms the active site of Mg-ATP and disrupts it as soon as the ATP is hydrolyzed
7. **ABC signature motif, LSGGQ:** It is the characteristic feature of ABC superfamily. It facilitates the formation of nucleotide sandwich dimer. (ter Beek *et al*., 2014)

Unlike the conserved feature of NBDs, TMDs have been found with distinct folds which also determine the mechanism for transport of substrate (ter Beek *et al*, 2014). The amino acid sequence of TMD determines the substrate specificity of each transporter (Stefkova *et al.*, 2003).

2.4.3 Molecular mechanism of action:

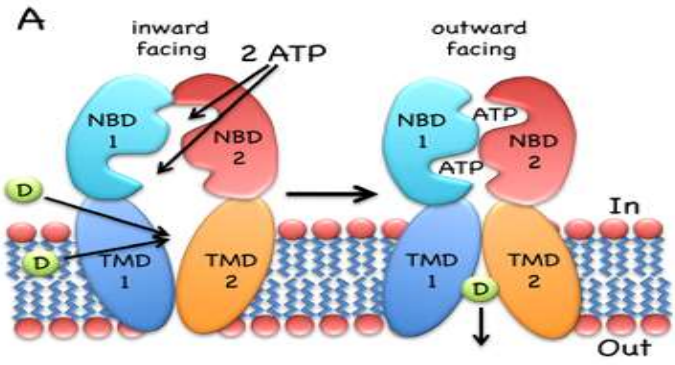


Figure 2.6 Two conformations of ABC exporters (Source: Wilkens *et al.,* 2015)

According to the ATP-switch model of Mamo *et al*., 2017, the molecular mechanism of action of ABC transporters can be divided into four basic steps:

a) Substrate binding
b) ATP binding and substrate translocation

c) ATP hydrolysis
d) Restore substrate binding conformation

When transporter is in ground/apostate, it possesses an inward-facing conformation or open confirmation, having an affinity for substrate. In this state, the substrate binding pocket is accessible by substrates from the cytosol region. The binding of substrate to the TMD induces a conformational change in NBD through the intracellular loops (ICLs), thereby increasing its affinity for ATP molecules. Two molecules of MgATP bind at the ATPase active site of NBD leading to NBD dimerization. This causes a change in TMD from inward-facing to outward-facing conformation (Figure 2.6) which reduces the affinity for the substrate, thus exporting it out of the cell. This is followed by ATP hydrolysis and release of ADP and Pi which destabilizes the NBD dimer and restores the open confirmation of TMD, thus making it accessible for other substrates. (Mamo *et al.,* 2017).

2.4.4 P-glycoprotein/ABCB1/MDR1

ABCB1, the first member of ABC transporter subfamily B, approximately 170 kDa size of membrane protein located on chromosome 7q21, is the first eukaryotic transporter discovered to be involved in cancer multidrug resistance (MDR). This glycoprotein is known to reduce drug permeability, hence the term 'Permeability-glycoprotein (P-gp)'. P-gp is expressed on the apical surface of many secretory cells of liver, intestine, kidney, placenta and luminal blood-brain barrier. MDR1 is involved in detoxification of cells by effluxing out xenobiotics and toxicants, which has been proved through various experiments performed on *mdr* knockout mice (as there is 87% sequence similarity between mouse and human P-gp). This large, polyspecific drug binding pocket of P-gp has the capacity to bind and transfer a wide array of structurally and chemically unrelated, mostly hydrophobic compounds of size ranging from 100-4000 Da such as chemotherapeutic drugs (includes docetaxel, irinotecan, Imatinib, paclitaxel, vinblastine *etc*.), antibiotics, antihistamines, HIV-protease inhibitors, analgesics, immunosuppressive agents, natural products and antiarrhythmics (Aller *et* al., 2009, Chen *et al.*, 2016, Mamo *et al*., 2017, Stefkova*et al*., 2003, Ward *et al*., 2013). P-gp is known to influence the ADME (absorption, metabolism, distribution and excretion) properties of a variety of drugs.

The transporter may either act as a 'hydrophobic vacuum cleaner' where the hydrophobic substrate initially partitions into the membrane before it is expelled out to the extracellular aqueous phase, or a 'flippase' which directly moves the substrate out to extracellular region as shown in Figure 2.7.

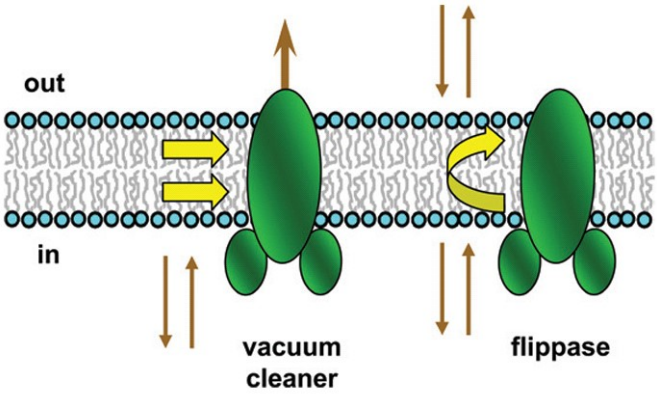


Figure 2.7 P-glycoprotein as hydrophobic vacuum cleaner or flippase (Source: Zhou, 2008)

Figure 2.7 shows the structure of P-glycoprotein containing two TMDs and two NBD. P-gp is a type full transporter consisting of 12 transmembrane spanning α-helices. As the picture depicts, the NH_2-terminal, COOH- terminal of the transporter and NBD region are located intracellularly (Zhou, 2008). It consists of two active ATP-binding sites (ATPases). The ATPase site is composed of Walker A and Walker B motifs of one NBD and signature C motif (LSGGQ) of the other NBD.

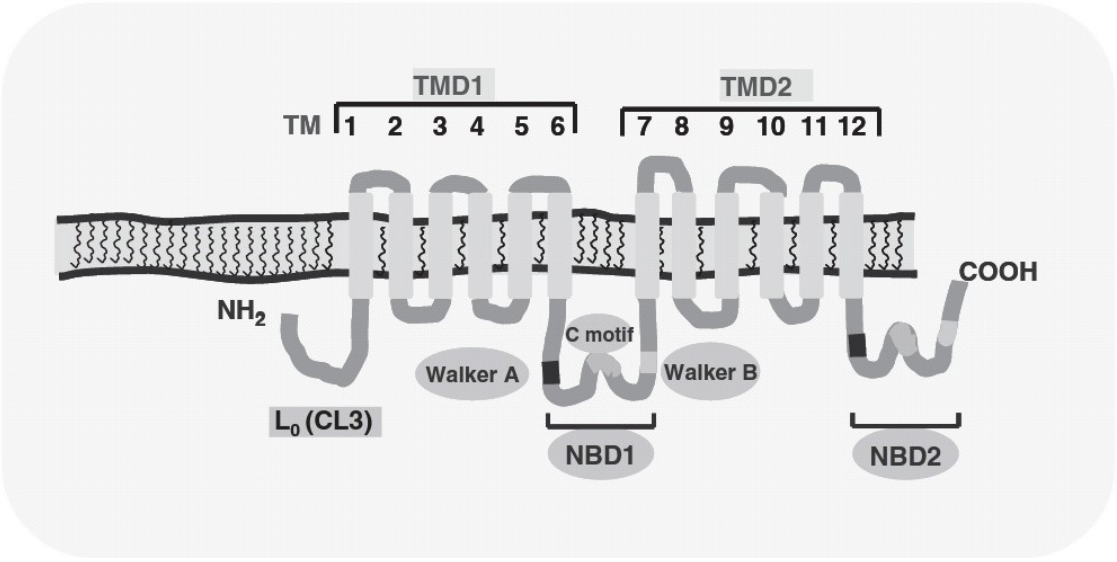


Figure 2.8 Topology of P-gp/MDR1 (Source: Zhou, 2008)

As mentioned in previous studies, P-gp has multiple substrate binding sites among which two major binding sites are located at TMD sites 5 and 6 and TMD site 11 and 12 (Zhou, 2008).

Kim *et al*., 2018 proved through their experiment of vandate trapping and photocleavage that only one ATP molecule is hydrolyzed at a time. This study also states that, that release of substrate across the membrane is driven by ATP binding and not ATP hydrolysis, which was unclear before.

2.4.5 *ABCB1* gene polymorphism

2.4.5.1 G2677T/A SNP (rs2032582) and cancer:

G2677T/A is a non synonymous SNP located in exon 21 encoding a region before TM9 (transmembrane helices) and TM10. It represents an amino acid change of Serine to either Alanine or Threonine (Ser893Ala/Thr) as three different nucleotides are found at this position (G, T or A) (Ankathil *et al.,* 2017). In African population, the genotype frequency of 2677 GG has been found to be greater than 81% in comparison with 10-32% 2677 GG genotype frequency in American Indians, Asians, Caucasians, Mexicans and Italians (Hodges *et al*., 2011). Also, in accordance with the data available from dbSNP, the allelic frequency of 2677T allele varies from 2-65% among world populations. The 2677A allele is comparatively uncommon allele ranging from 0-17% in different ethnicities (Hodges *et al*., 2011). Controversial results are observed in the functional effect of this SNP. Kim *et al.*, identified an increase in efflux function of P-gp in presence of allele 2677T while Salama *et al*., found diminished function for the same. Authors such as Wang *et al.,* and Morita *et al.,* did not find any effect. Biochemical analysis showed that the presence of 2677T or 2677A allele may alter drug transport by affecting induced drug transport activity. Further, contradictory results have been obtained regarding association of Polymorphism G2677T with overall survival in various cancers such as acute leukemia, colorectal cancer, squamous cell carcinoma (Feng *et al*., 2016, Gaikovitch E *et al*., 2004, Gervasini *et al*., 2006). Green *et al.* reported association between *ABCB1* gene polymorphism G2677T/A with chemotherapy response of ovarian cancer patients undergoing paclitaxel therapy. Patients harbouring homozygously mutated genotype showcased elevated response.

2.4.5.2 C1236T SNP (rs1128503) and cancer*:*

This synonymous SNP located in exon 12 encodes for TM6 region which is essential for substrate binding. C1236T does not exhibit an amino change at position 412 (Glycine). According to dbSNP, the allelic frequency of C allele ranges from 30-93% depending upon ethnicity. The T allele has been

found as the minor allele in African population, whereas in Asians, C allele is the minor allele (Hodges *et al.,* 2011). The SNPs C1236T and G2677T/A have been reported to be associated with altered mRNA expression levels, mRNA stability, and protein folding and influence drug pharmacokinetics (Ankathil et al., 2017). This SNP has shown to affect protein folding when found in combination with SNP C3435T due to the use of a rare codon. As observed by Salama *et al.* in an *in vitro* study. In a meta-analysis by Zu et al, a significant association between C1236T polymorphism and increasing risk of Imatinib methylate (IM) resistance in Asian Chronic myeloid leukemia (CML) patients was observed.

2.5. Computational analysis of the effects of Single nucleotide polymorphism (SNPs) substitutions in Human ABC transporters

SNPs are variations in genetic sequences that affect only one building block of the DNA segments (Adenine, guanine, cytosine or thymine), which are present in more than 1% of the population, and occur at a frequency of 100-300 nucleotides. They are responsible for the genetic differences that are observed among human population. However, these substitution may have some deleterious effect on the body, as they may alter protein/cellular function, causing serious disease like cancer, genetic disorders etc. These SNPs can either fall into the coding regions or in non-coding region or between two genes. The SNPs in coding region (coding SNPs) are considered to have major effect on the protein structure and function as they cause amino acid change in the respective protein (non-synonymous SNP). To study the impact of these non-synonymous SNPs (nsSNP), various computational tools are used such as SIFT, PROVEAN, Mutpred2, PANTHER, PMUT, SNPs&GO, Polyphen-2, MUpro, Imutant3.0, PhD-SNP and GENEMANIA. (Abdelraheem *et al.,* 2016; Manickam *et al.,* 2014)

1. **SIFT (Sorting Intolerant From Tolerant)**

   SIFT algorithm ((http://sift.jcvi.org) used to predict the effect of single Amino acid substitution on protein structure and function. In order to study this effect, the algorithm assumes that the important positions in amino acid sequence of protein have been conserved during evolution, and hence any change at this position may affect functioning of protein. The SIFT algorithm helps in sorting out intolerant amino acid substitution from the tolerant ones using sequence homology tools. The algorithm takes into account all the possible substitutions at each important position of

protein sequence and assigns score based on its effect. Positions having probabilities less than 0.05 are considered as deleterious, where probabilities more than 0.05 are tolerated.

2. **Polyphen-2 (Polymorphism Phenotyping v2)**

   Polyphen-2 software and web server is involved in prediction of potential impact of amino acid substitution on protein stability and function using comparative and structural evolutionary considerations. It measures the probability of a missense mutation being damaging by performing and combining various activities that includes functional annotation of SNPs, maps coding SNPs to gene transcripts, extraction of protein sequence annotations and structural attributes, and then final construction of conservation profiles. Input submitted to polyphen-2 software includes FASTA sequence of protein along with position of interest and new residue.

3. **MutPred2**

   MutPred2 (http://mutdb.org/mutpred) software helps in identification of harmful mutations by determining the structural and functional site changes in mutant sequence as compared with wild-type sequences. It provides us with information on probabilities of gain/loss of function of proteins structure and function by assigning g-scores. On the basis of g-score, it concludes whether the amino acid substitution is disease-associated/deleterious. Another significant value called Property score (p) determines whether this substitution has impacted any structural and functional features.

4. **PANTHER**

   PANTHER (http://pantherdb.org/tools/csnpScoreForm.jsp) software helps in determination of the possibility that a particular non-synonymous SNP accounts for functional impact on protein properties. This is done by calculation of subPSEC score (substitution position-specific evolutionary conservation) which is defined as the negative logarithm of probability ratio of wild type to mutant type at a specific position. subPSEC score value ranges from 0 (neutral) to -10 (most likely to be deleterious).

5. **PMUT**

   PMUT (http://mmb2.pcb.ub.es:8080/PMut) software uses a strong methodology to presume disease-associated mutations. The method employs the use of neural networks skilled with a large database of pathological mutations as well as neutral mutations. Output is displayed as ranging

from 0 to 1 (>0.5 signifies single pathological mutation) and a confidence index ranging from 0-9 (low to high).

6. **SNPs & GO**

   SNPs & GO (http://snps-and-go.biocomp.unibo.it/snps-and-go/) algorithm, by utilizing functional information encoded by Gene ontology, envisions the impact of protein variation. Cellular component, molecular function and biological are the three main roots included from gene ontology.

7. **PROVEAN (Protein Variation Effect Analyzer)**

   PROVEAN (http://provean.jcvi.org) software determines the impact of protein sequence variation on protein function on the basis of clustered sequences. It employs clustering methodology where BLAST hits with > 75% global sequence identity are clustered together and the final PROVEAN score is generated by averaging top 30 supporting sequence clusters. If the final PROVEAN score is below certain threshold, the variation is termed as deleterious and the score above threshold value is predicted as neutral.

8. **MUpro**

   MUpro (http://www.ics.uci.edu/~baldig/mutation.html) software functions on the basis of two machine learning programs: Neural networks and Support Vector Machines. Both of these MLP are skilled on large database of mutation and yield accuracy above 84% via 20-fold cross validation. This software is involved in prediction of impact of single AAS on protein stability. Tertiary protein structures are not required to predict protein stability changes which is a major advantage is of MUpro. The final output is displayed as energy change value and confidence score is calculated to measure the confidence of prediction (between -1 to 1). Confidence score < 0 indicated decrease in protein stability as a result of variation and score > 0 denotes variant increases in protein stability.

9. **Imutant3.0**

   Imutant3.0 http://gpcr2.biocomp.unibo.it/cgi/predictors/I-Mutant3.0/IMutant3.0.cgi) is a web server based on Support Vector Machine. It is involved in estimation of impact of single site mutations on protein stability in protein sequence and structure. We can either choose for prediction of all the possible mutation impact on protein stability or only a specific mutation. The

final output generated includes free energy change as DDG value (kcal/mol) and RI value (reliability index).

10. **PhD-SNP**

PhD-SNP (http://snps.biofold.org/phd-snp/phd-snp.html) software helps in classifying whether the single-site mutation is disease-related or a neutral polymorphism. FASTA sequences of protein along with residue changes are submitted to PhD-SNP as input for the analysis.

11. **GENEMANIA**

GENEMANIA (http://www.genemania.org/) is an online database used to predict gene function, to find other related genes to the set of genes submitted as input, and find the role of input genes in different pathways. GENEMANIA utilizes a large set of functional association data such as pathways, co-expression, co-localization, protein and genetic interactions and protein domain similarity.

12. **ELASPIC**

ELASPIC (http://elaspic.kimlab.org/many/) web server is used to observe and study the effect of single or multiple mutations on protein-protein interaction and protein folding. Effect of mutation on all the proteins present in UniProt database can be predicted using ELASPIC. The result displayed includes modelled structures of both wild and mutant type proteins.

Chapter 3

Objective

The objectives of this dissertation include:

- To investigate the effect of *ABCB1* gene polymorphism on protein structure and function, and to check whether these variation contribute in disease or not.
- To predict the overall survival of Advanced Lung Cancer Patients having polymorphism in *ABCB1* gene treated with Platinum based doublet Chemotherapy.

Chapter 4

Material and Methods

4.1. Study subjects and Sample collection

Lung Cancer patient samples were collected for this study from the Department of pulmonary Medicine, Post Graduate Institute of Medical Education and Research (PGIMER) Chandigarh, India. The present study has 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 histo-pathologically diagnosed as having NSCLC and SCLC. There was no age, gender, smoking, histological, or TNM stage restrictions, but patients with a prior history of cancer were excluded from this study. A detailed questionnaire was completed for each case by a trained interviewer. The questionnaire included information on demographic and smoking characteristics. Smokers reported tobacco habits such as smoking of cigarette and/or bidi (a native cigarette like stick of coarse tobacco hand-rolled in a dry temburni leaf). As an indication of cumulative smoking exposure, pack-years were calculated by the following formula: [(cigarettes or bidi smoked by an individual per day / 20) X No. Of years smoked]. Further, medical information of cases, such as 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 genomic DNA from whole blood of lung cancer patients was performed by employing method used by Field *et al.,* which involves Proteinase K digestion, Phenol/chloroform extraction and ethanol precipitation (Field *et al.,*. 1999).

Requirements:

- Washing buffer
- Lysis buffer
- Phenol:Chloroform:Isoamylalcohol (25:24:1)
- Chloroform:Isoamylalcohol (24:1)

- Isopropanol
- TE buffer

Preparation of Reagents and buffers:

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

4.2.1. Washing Buffer

Table 4.1. Preparation of Washing buffer

Stock concentration	Working concentration
1 M Sucrose	320 mM Sucrose
100 % Triton X - 100	1 % Triton X - 100
100 mM Magnesium chloride	5 mM Magnesium chloride
100 mM Tris-HCl (pH = 8.0)	10 mM Tris-HCl (pH = 8.0)

4.2.2. Lysis Buffer

Table 4.2. Preparation of Lysis buffer

Stock concentration	Working concentration
1 M Tris-Cl buffer (pH = 8.0)	400 mM Tris-Cl buffer (pH = 8.0)
10 % SDS	1 % SDS
0.5 M EDTA (pH = 8.0)	60 mM EDTA (pH = 8.0)
5 M NaCl	150 mM NaCl
10 mg/mL Proteinase K	10 µg/mL Proteinase K

4.2.3 5X TBE buffer

5X TBE buffer was prepared by dissolving 54g of Tris base, 27.5g of boric acid and 20 ml 0.5M EDTA. Volume was made up to 1000 mL by adding distilled water.

4.2.4 6X loading Dye

6X loading dye was prepared by dissolving 0.25% Bromophenol blue (0.05 g), 0.25% Xylene cyanol (0.05 g), 40% sucrose (8 g). Make up the final volume of 20 ml with TE buffer.

4.2.5 Procedure of DNA Isolation

- Took 5ml of blood and added 5ml of Washing Buffer 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 45 0.26ml of water) to the pellet and resuspended the pellet in the Buffer and centrifuged again (repeated this step thrice).
- Dissolved the pellet in 5ml of Lysis buffer (2 mL of 1 M Tris HCl, 0.5ml of 10% SDS, 0.6ml of 0.5 M EDTA, 0.15ml of 5M NaCl, 0.05ml of 10mg/ml Proteinase-K and 1.7ml water) followed by overnight incubation at 44°C.
- Added an equal volume of Phenol: Chloroform: Isoamyl alcohol (PCI) in 25:24:1 ratio to the above mixture and mixed the contents slowly followed by centrifugation at 8000 rpm/10 minutes/4°C.
- After centrifugation, took the upper aqueous layer and again added PCI mixture 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 DNA pellet was washed with chilled 70% Ethanol twice at 10,000 rpm for 5 minutes.
- Decant ethanol and air dried the pellet.
- Dissolved the pellet in 50µl-150µl Tris-EDTA buffer depending on the size of DNA pellet (Bartlett & White, 2003).

4.3 DNA Quantification

With the help of Thermo Scientific Nanodrop Spectrophotometer, the quality of DNA can be determined. It is much better as compared to Visible Spectrophotometer as only 1 µl of sample is required without the use of containment devices such as cuvettes. The sample is placed between two optical surfaces, which determine 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:

- Cleaned the Nanodrop using 1µl of Deionised water
- Opened the Nanodrop software and select Nucleic acid module
- Loaded 1µl of TE buffer to set blank measurement
- Loaded 1µl of DNA sample and selected 'measure' to take the reading
- Concentration and purity of sample were calculated automatically

DNA concentration can also be calculated as follows:

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

(Standard DNA sample O.D. = 1 at 50 µg/mL concentration)

Quality/purity of DNA = OD_{260} / OD_{280}

NOTE: 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. Genotyping by Polymerase chain reaction-Restriction fragment length polymorphism (PCR-RFLP)

4.4.1. Polymerase Chain Reaction of *ABCB1* genetic variants

The polymerase chain reaction (PCR) is an enzymatic process that helps in detection of specific genes within an environmental DNA sample. Oligonucleotide primers are used in PCR for amplification of required genomic DNA region. These oligonucleotide primers are short, user defined DNA sequences complementary to the target DNA sequence. 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. The two SNPs (C1236T and G2677T) of *ABCB1* gene were amplified by PCR. Briefly, the total PCR reaction mixture was of 15 µL volume and it consisted of 1x PCR buffer, 1.5 mM $MgCl_2$ with 0.5 µM of both forward and reverse primer, 200 µM of each dNTP's, 100 µg/ml bovine serum albumin (BSA) and 1U Taq polymerase (DNAzyme II DNA Polymerase, Thermo Scientific) and approximately 200 ng DNA. The amplicons obtained were run on 2 % Agarose gel containing Ethidium bromide and visualized under UV transilluminator. The primer sequence, amplified product size, restriction enzymes and cutting pattern was similar as

described by Balcerczak *et al*., 2010. Given below in Table 4.4.1 and Table 4.4.2 is the detailed information. The detail information of reagents used in PCR and PCR condition used for amplifying the two variants (C1236T (rs1128503) and G2677T (rs2032582)) of *ABCB1* gene are mentioned in Table 4.4.3 and 4.4.4, respectively.

Table 4.4.1 Details of amplified product size, restriction enzymes and cutting pattern

SNP's of *ABCB1* gene	**Amino acid substitution**	**PCR product size (bp) and Restriction Enzyme**	**Fragments Identifying Genotypes**
C1236T **(rs1128503)**	Gly412Gly	366 bp and Hae III	CC - 269, 62, 35 CT- 269, 97, 62, 35 TT - 269, 97
G2677T **(rs2032582)**	Ala893Ser	224 bp and Ban I	GG – 198, 26 GT – 224, 198, 26 TT - 224

Table 4.4.2 Details of primer sequence

SNP's of *ABCB1* gene	**Forward Primer**	**Reverse Primer**
C1236T **(rs1128503)**	5-ATCCTGTGTCTGTGAATTGCC-3	5-CCTGACTCACCACACCAATG-3
G2677T **(rs2032582)**	5-TGCAGGCTATAGGTTCCAGG-3	5-GTTTGACTCACCTTCCCAG-3

Table 4.4.3 Reagents used in PCR			
Reagent	**Stock concentration**	**Final reaction concentration**	**Quantity used**
Additive 1 (BSA)	1000 µg/mL	100 µg/mL	33 µL
PCR buffer (Mg Conc.)	10X (25 mM)	1X (1.5 mM)	33 µL
ABCB1 rs1128503 OR *ABCB1* rs2032582 Forward primer	10 µM	0.5 µM	16.50 µL
ABCB1 rs1128503 OR *ABCB1* rs2032582 Reverse primer	10 µM	0.5 µM	16.50 µL
Taq polymerase	2.0 U/ µL	1.5 mM	3.52 µL
dNTP	10 mM each	0.2 mM each	6.60 µL
PCR grade water	-	-	130.48 µL
DNA template	100 ng/µL	200 ng	2 µL

Table 4.4.4 PCR Condition*		
Steps	**Temperature**	**Time**
Initial Denaturation	95°C	5min
Denaturation	94°C	30 seconds
Annealing	55°C (*ABCB1* rs1128503) 57°C (*ABCB1* rs2032582)	45 seconds
Polymerization	72°C	30 seconds
Final Extension	72°C	5 min

*Total no. of cycle = 29

4.4.2 Restriction Fragment Length Polymorphism (RFLP)

After obtaining the desired PCR product, the amplicons of C1236T and G2677T is treated with respective restriction enzyme mentioned in Table 4.4.1. The RFLP was done using 10µl volume of amplified product, which was digested with 0.2 units of restriction enzyme. Final volume of 20µl reaction was prepared using 1X enzyme buffer. The reaction was kept for incubation at 37 ºC. After completion of incubation period reaction was stopped using 1X loading dye. After this the RFLP pattern was checked using Polyacrylamide gel electrophoresis.

Hae III: The microorganism from which it has been isolated is *Haemophilus aegypticus*. This enzyme was used to digest amplicons of *ABCB1* rs1128503. The enzyme was used along with NEB4 buffer (1X) to provide the optimum condition for reaction.

5′…….GGCC…….3′

3′…….CCGG…….5′

Ban I: The source of enzyme is *Bacillus aneurinolyticus*. It is used for digestion of *ABCB1* rs2032582 amplicons. The enzyme was used with buffer O (1X) to facilitate the reaction digestion.

5′…….GGYRCC…….3′

3′…….CCRYGG…….5′

4.4.3 Native-Polyacrylamide gel electrophoresis (N-PAGE)

PAGE gel is made up of acrylamide and bis-acrylamide. It is a highly sensitive method to check the resolution of DNA molecule by 1bp. Acrylamide molecules are linked together with each other and bis-acrlyamide further cross-links two monomeric acrylamide molecules. For performing the N-PAGE Acrylamide and bis-acrlyamide are mixed in a fixed ratio which determines the pore

size of the gel. Polymerization in N-PAGE is done with APS (ammonium per sulphate) facilitated by TEMED (N, N, N, N-tetramethylethylenediamine). Acrylamide gels are defined in terms of the total percentage of acrylamide present and the pore size in the gel can be varied by altering the the concentrations of both the acrylamide and bisacrylamide. Low percentage acrylamide gels (4 – 12%), which have large pore sizes, can be used to separate DNA according to their sizes. For C1236T and G2677T polymorphic variant 10% and 8% gels were prepared for resolution of DNA fragments. After vertical gel electrophoresis ethidium bromide staining was done to make the bands visible.
The detailed composition of N-PAGE gel is mentioned in Table 4.4.5.

Table 4.4.5 Composition of N-PAGE gel

Gel %	**30% Acrylamide**	**Distilled water**	**5X TBE**	**10% APS**	**TEMED**
8% (***ABCB1* rs1128503**)	3.2 mL	6.4 mL	2.4 mL	200 µL	10 µL
10% ***(ABCB1* rs2032582*)***	4.0 mL	5.6 mL	2.4 mL	200 µL	10 µL

4.5 Computational analysis of C1236T and G2677T polymorphic variants of *ABCB1* gene

The computation analysis for C1236T and G2677T was performed using various bioinformatics tools like SIFT, PROVEAN, POLYPHEN2, I-MUTANT, MuPro, Phd-SNP, SNP&GO, ELASPIC, PANTHER, MutPred2 and Genemania as shown in Table 4.5 (Abdelraheem *et al*., 2016; Manickam *et al*., 2014).

Table 4.5 Computational tools	
Computational Tool	**Uniform Resource Locator (URL)**
SIFT	http://sift.jcvi.org.
PROVEAN	http://provean.jcvi.org.
POLYPHEN2	http://genetics.bwh.harvard.edu/pph2/
I-MUTANT	http://gpcr2.biocomp.unibo.it/cgi/predictors/IMutant3.0/IMutant3.0.cgi.
MuPro	http://www.ics.uci.edu/~baldig/mutation.html
PhD-SNP	http://snps.biofold.org/phdsnp/phd-snp.html
SNP & GO	http://snps.biofold.org/snps-and-go/snps-andgo.html.
ELASPIC	http://elaspic.kimlab.org/many/
PANTHER	http://pantherdb.org/tools/csnpScoreForm.jsp
MutPred	http://mutdb.org/mutpred
Genemania	http://www.genemania.org/

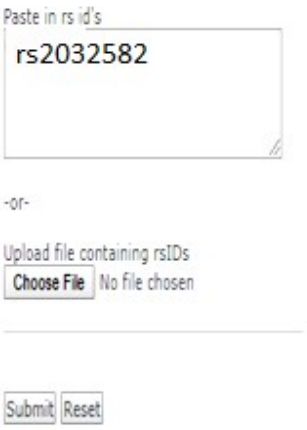
Computational analysis of Deleterious Single Nucleotide Polymorphisms (SNPs) in Human *ABCB1* gene involved in cancer using multiple softwares.

1. Prediction of structural impact of nsSNP on protein using SIFT software

a. For *ABCB1* rs2032582

Input:

rsID- rsID for *ABCB1* G2677T SNP was rs2032582, retrieved from NCBI website.



2. Prediction of functional effect of amino acid substitution using PROVEAN software

Input:

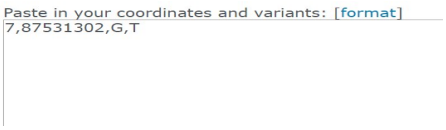
Chromosome no. and position, Nucleotide change were retrieved from NCBI website.

For *ABCB1* rs2032582,

Chromosome number: 7

Chromosome position: 87531302

Nucleotide change: G to T



3. **Prediction of deleterious SNP by POLYPHEN2 software**

a. For *ABCB1* rs2032582

Input: rs ID (rs2032582)

4. **Prediction of nsSNP impact on protein stability with the help of I-Mutant software**

Input:

Proein sequence in FASTA format (retrieved from Swiss-Prot), Position at which amino acid substitution observed and amino acid changes into which new residue (from SIFT software output), temperature and pH.

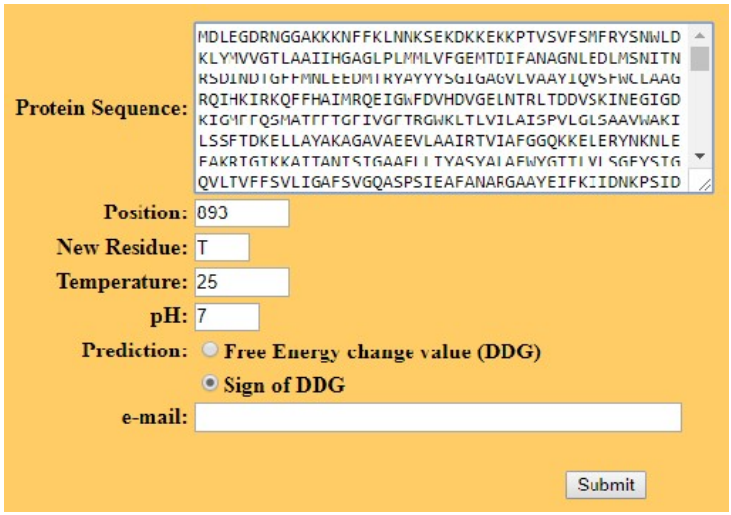
a. For *ABCB1* rs2032582

Protein sequence:
MDLEGDRNGGAKKKNFFKLNNKSEKDKKEKKPTVSVFSMFRYSNWLDKLYMVVGTL
AAIIHGAGLPLMMLVFGEMTDIFANAGNLEDLMSNITNRSDINDTGFFMNLEEDMTRYA
YYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRQFFHAIMRQEIGWFDVHDVGELNTRL

TDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLTLVILAISPVLGSAAVWAKILS
SFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNKNLEEAKRIGIKKAITANI
SIGAAFLLIYASYALAFWYGTTLVLSGEYSIGQVLTVFFSVLIGAFSVGQASPSIEAFANA
RGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKGNLEFRNVHFSYPSRKEVKILKGLNLKVQS
GQTVALVGNSGCGKSTTVQLMQRLYDPTEGMVSVDGQDIRTINVRFLREIIGVVSQEPV
LFATTIAENIRYGRENVTMDEIEKAVKEANAYDFIMKLPHKFDTLVGERGAQLSGGQKQ
RIAIARALVRNPKILLLDEATSALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVI
AGFDDGVIVEKGNHDELMKEKGIYFKLVTMQTAGNEVELENAADESKSEIDALEMSSN
DSRSSLIRKRSTRRSVRGSQAQDRKLSTKEALDESIPPVSFWRIMKLNLTEWPYFVVGVF
CAIINGGLQPAFAIIFSKIIGVFTRIDDPETKRQNSNLFSLLFLALGIISFITFFLQGFTFGKAG
EILTKRLRYMVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNI
ANLGTGIIISFIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIE
NFRTVVSLTQEQKFEHMYAQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGA
YLVAHKLMSFEDVLLVFSAVVFGAMAVGQVSSFAPDYAKAKISAAHIIMIIEKTPLIDSY
STEGLMPNTLEGNVTFGEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
QLLERFYDPLAGKVLLDGKEIKRLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVV
SQEEIVRAAKEANIHAFIESLPNKYSTKVGDKGTQLSGGQKQRIAIARALVRQPHILLLDE
ATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLL
AQKGIYFSMVSVQAGTKRQ

Position of Mutation: 893

Mutated amino acid: T

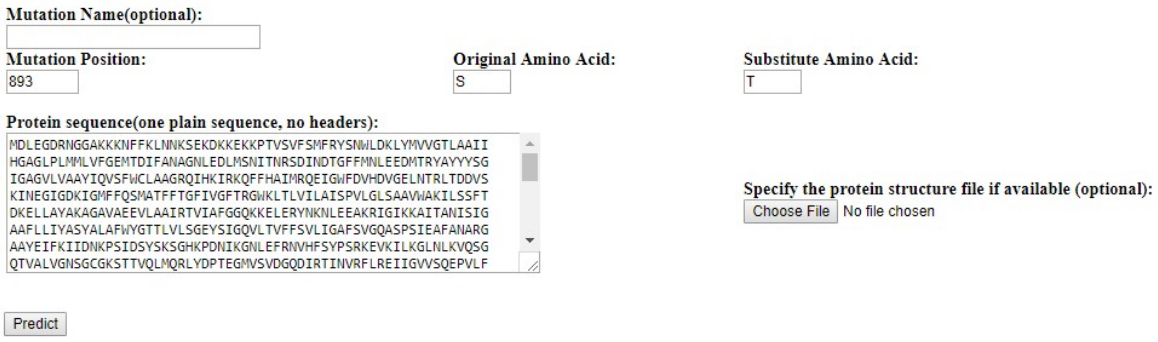


5. Prediction of protein stability on the basis of energy change by MUpro software

Input:

Mutation position, original amino acid, Mutated amino acid, protein sequence in FASTA format. Mutation position, original amino acid and mutated amino acid was fetch back from SIFT software output and protein sequence was retrieved from Swiss-Prot.

a. For *ABCB1* rs2032582



6. To predict whether the mutation is disease related or neutral

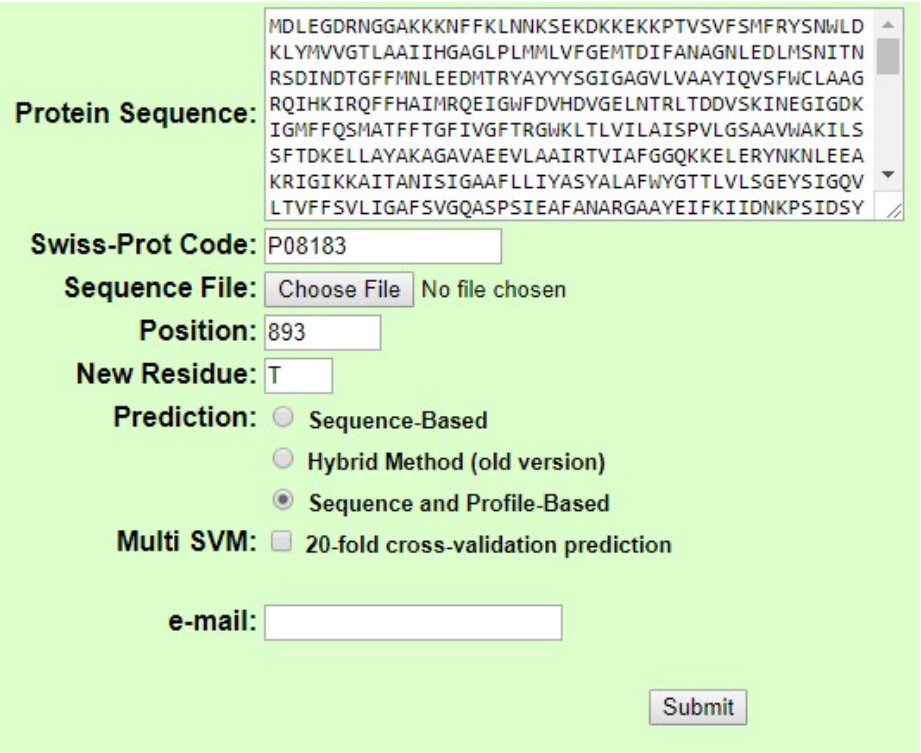
This was done by Phd-SNP software.

Input:

Protein sequence in FASTA format, Swiss-Prot Code for protein, Mutation position, Mutated amino acid and basis of prediction were entered. Protein sequence and code for protein was retrieved from swiss prot. Mutation position, Mutated amino acid was fetch back from SIFT software output.

a. For *ABCB1* rs2032582

Protein code: P08183

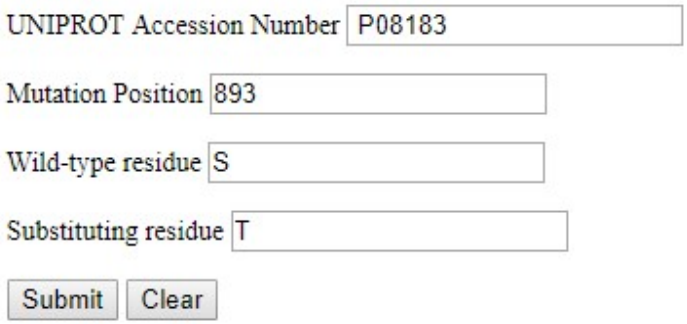


7. To predict whether mutation is related to disease or not by SNP & GO software

Input:

UNIPROT accession Number for protein, Mutation position, original amino acid, Mutated amino acid. Accession number was fetch from Swiss-Prot website and mutation position, original amino acid, mutated amino acid was retrieved from SIFT software output.

For *ABCB1* rs2032582



8. Prediction of effects of mutations on protein folding and protein-protein interactions by ELASPIC Software

INPUT: name of gene_organism_original amino acid changes at which position into what kind of new amino acid

a. For *ABCB1* rs2032582

9. Prediction of Deleterious SNP with the help of PANTHER

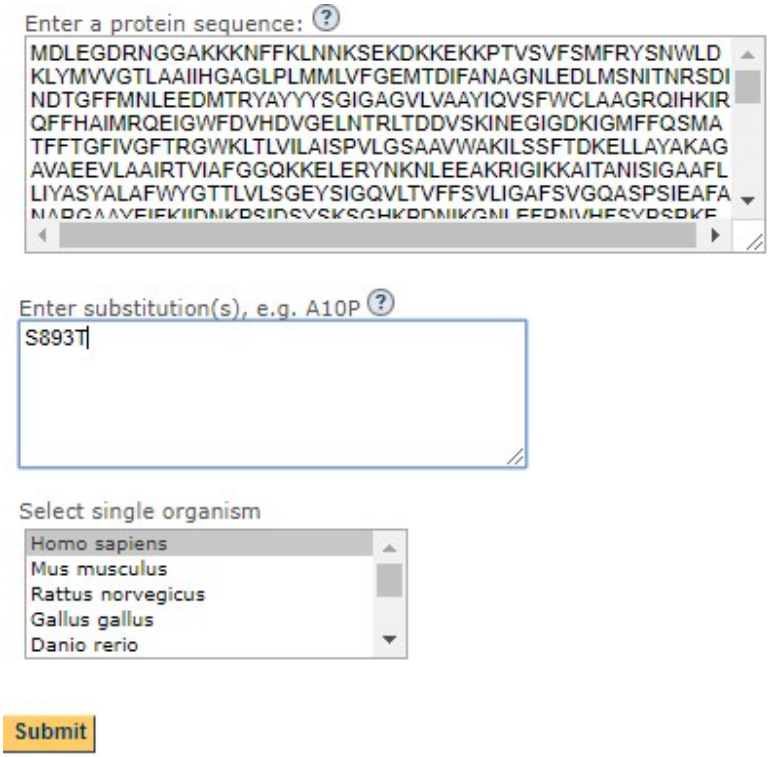
Input:

Protein sequence: retrieved from Swiss-prot database

Amino acid substitution Substitution: |wild type amino acid||position of mutation||mutated amino acid|. These were fetched from SIFT Software output.

Name of organism

For *ABCB1* rs2032582



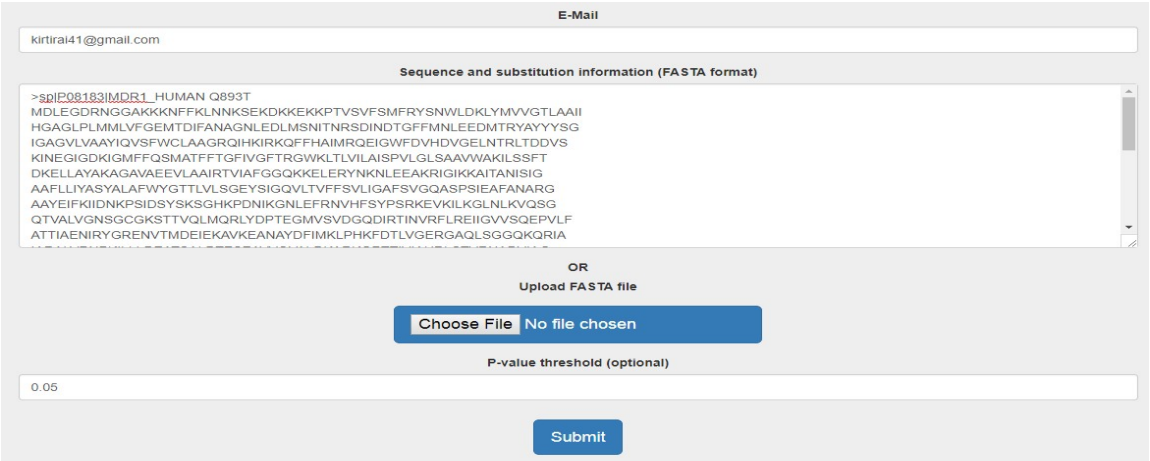
10. Prediction of harmful mutation by MutPred Software

Input:

Email ID

Sequence and Substitution Information: |Swiss-Prot protein code||sequence size||wild amino acid||position of mutation||Mutated amino acid

a. For *ABCB1* rs2032582

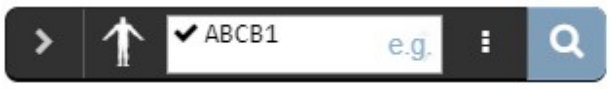


11. Prediction of interaction of gene with other genes with the help of GENEMANIA Software

Input:

For *ABCB1* rs2032582,

a. for all networks



4.6 Statistical Analysis

In order to estimate the genotype frequencies of *ABCB1* gene polymorphisms, we applied Hardy Weinberg equilibrium theory which utilizes the equation $p^2+q^2+2pq=1$ to calculate genotype frequency (where p is the frequency of wild type allele, q is the frequency of variant allele). The homozygous wild type TT as dominant model and CC as recessive model for rs1128503 variant was used as reference and for rs2032582 GG was taken as reference. Analysis of overall survival was done using unadjusted univariate method (Kaplan-Meier survival analysis). The multivariate

analysis was done by applying Cox proportional hazard regression model. Hazard ratios (HRs) and 95% CIs were calculated using Cox proportional hazards models. To demonstrate association between risk for lung cancer on the basis of clinical stage, chemotherapy response and performance status and *ABCB1* polymorphisms rs1128503 and rs2032582, adjusted Odds Ratio (ORs) along with 95% Confidence Intervals (CI) was calculated using logistic regression analysis. All tests were two-sided and statistical significance was set at $p<0.05$. All the statistical analysis was performed with Medcalc version 18.5 (Medcalc Software, Ostend, Belgium).

Chapter 5

Results

5.1. DNA isolation

Genomic DNA was isolated from 5 mL of peripheral blood of lung cancer patients and run on 0.7% agarose gel for qualitative analysis. Trans-UV-illuminator was used to confirm the presence of DNA sample as shown in Fig 5.1. The isolated DNA samples were diluted in TE buffer to a concentration of 100 ng/µL, which can further be used in Polymerase chain reaction as template.

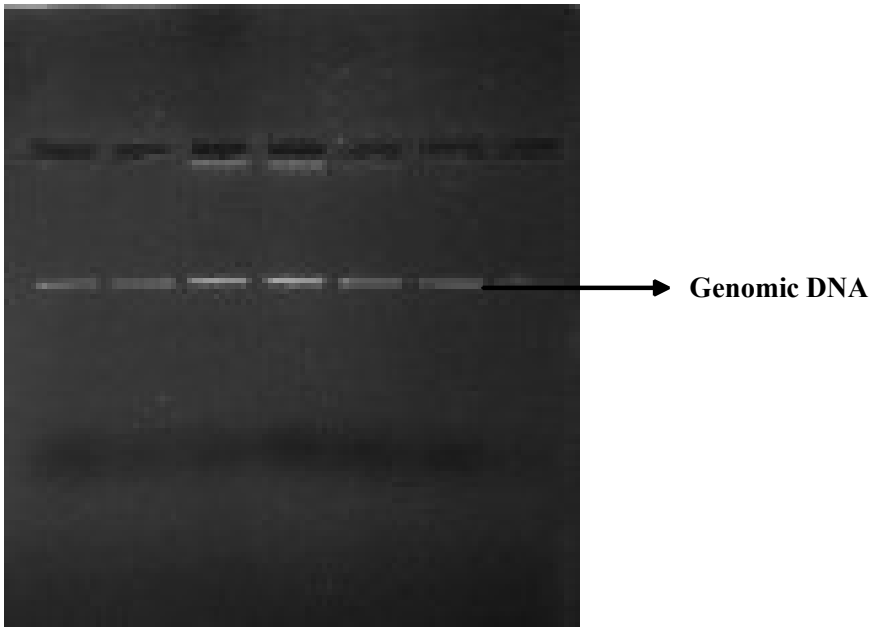


Figure 5.1 Representative example of 0.7% Agarose gel showing the presence of genomic DNA

5.2. Polymerase chain reaction (PCR) for *ABCB1* rs1128503 *and* rs2032582 polymorphism

Polymerase chain reaction of the isolated genomic DNA was performed simultaneously to amplify the required region of *ABCB1* gene using specifically designed primers. The amplicons obtained through the process were run on 1.7% agarose gel stained with ethidium bromide (EtBr) which helps in illuminating the DNA bands when observed in UV light. The respective amplicon size of *ABCB1* rs1128503 and rs2032582 obtained were 366 bp and 224 bp respectively.

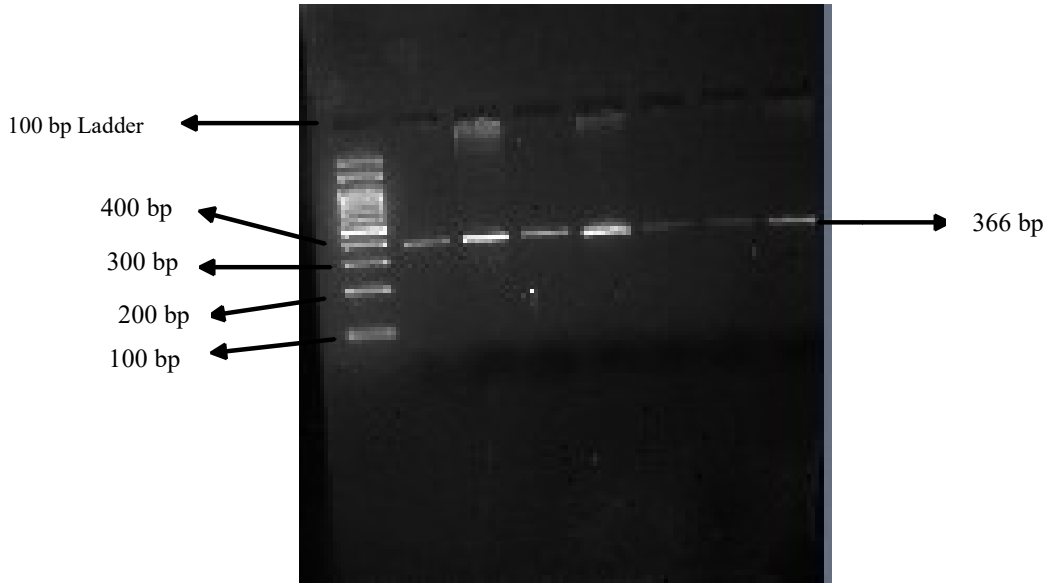


Figure 5.2 Representative example of PCR amplified DNA products of *ABCB1* rs1128503 (amplicon size 366 bp) for LC samples. Lane 1: 100 bp ladder, Lane 2,3,4,5,6,7,8: Amplified PCR products

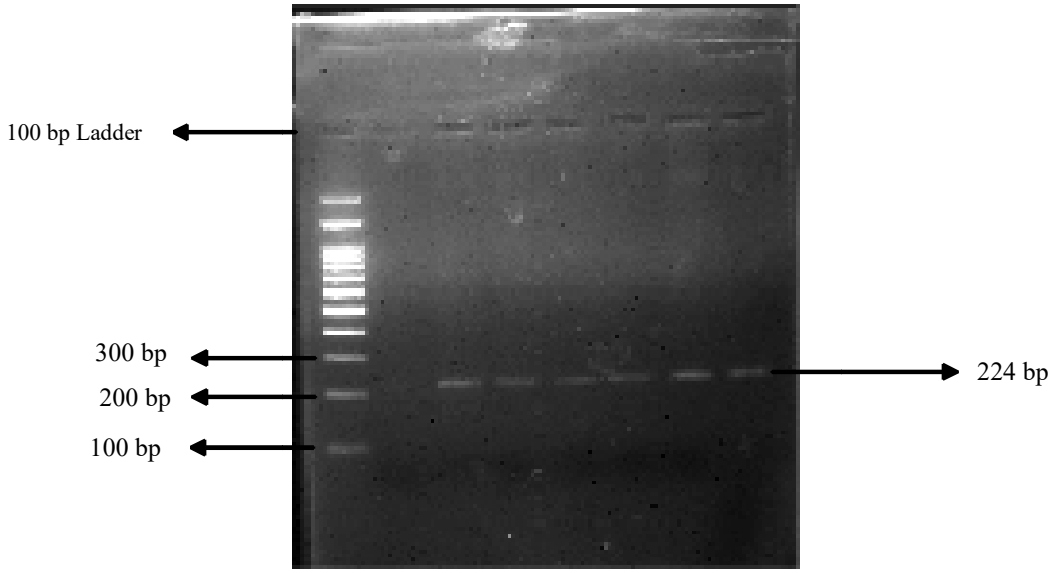


Figure 5.3 Representative example of PCR amplified DNA products of *ABCB1* rs2032582 (amplicon size 224 bp) for LC samples. Lane 1: 100 bp ladder, Lane 2, 3, 4, 5, 6, 7, 8: Amplified PCR products

5.3. Native Polyacrylamide gel electrophoresis of *ABCB1* rs1128503 and rs2032582 polymorphism amplicons

Restriction digestion of the amplicons obtained through amplification was performed using their corresponding restriction enzymes which cleave the DNA at their unique restriction sites specific for single nucleotide sequence and enzymes. The digested products were then run on 8% for ***ABCB1* rs1128503** and 10% for ***ABCB1* rs2032582** polyacrylamide gel, depending on the size of the digestion product.

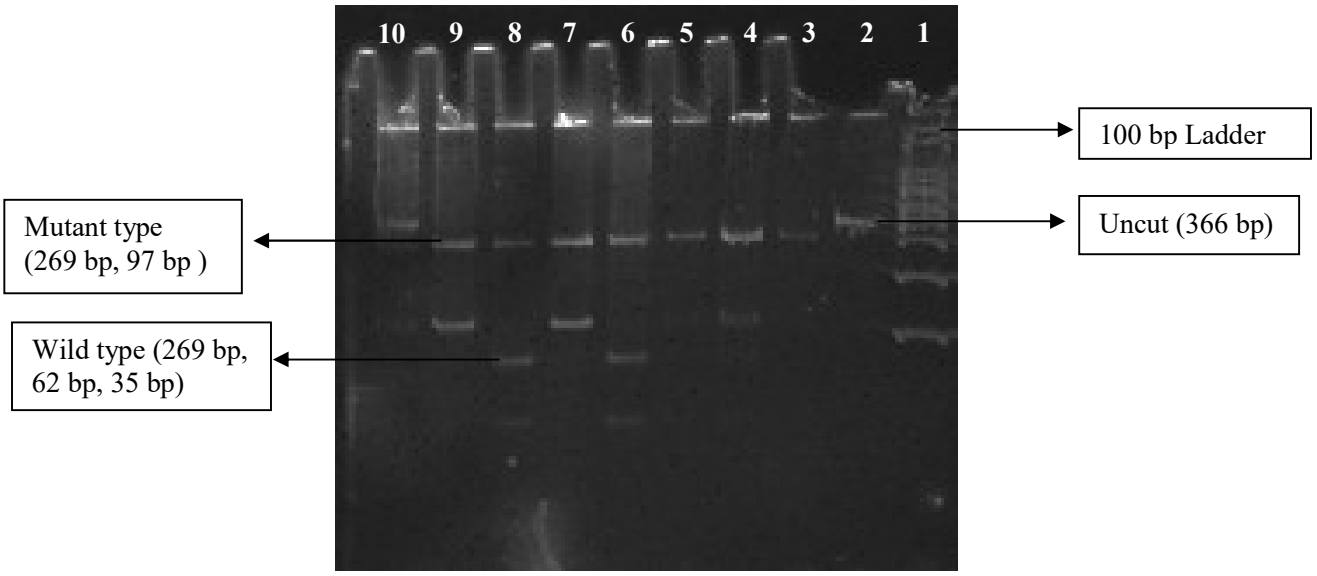


Figure 5.4 Representative sample for Restriction digestion of *ABCB1* rs1128503 PCR product showing the digested products (Wild: 269 bp, 62 bp, 35 bp; Mutant: 269 bp, 97 bp; Heterozygote: 269 bp, 97 bp, 62 bp, 35 bp). Lane 1: 100 bp Ladder; Lane 2: Uncut 366 bp; Lane 4,7,9,10: Mutant type; Lane 6,8: Wild type

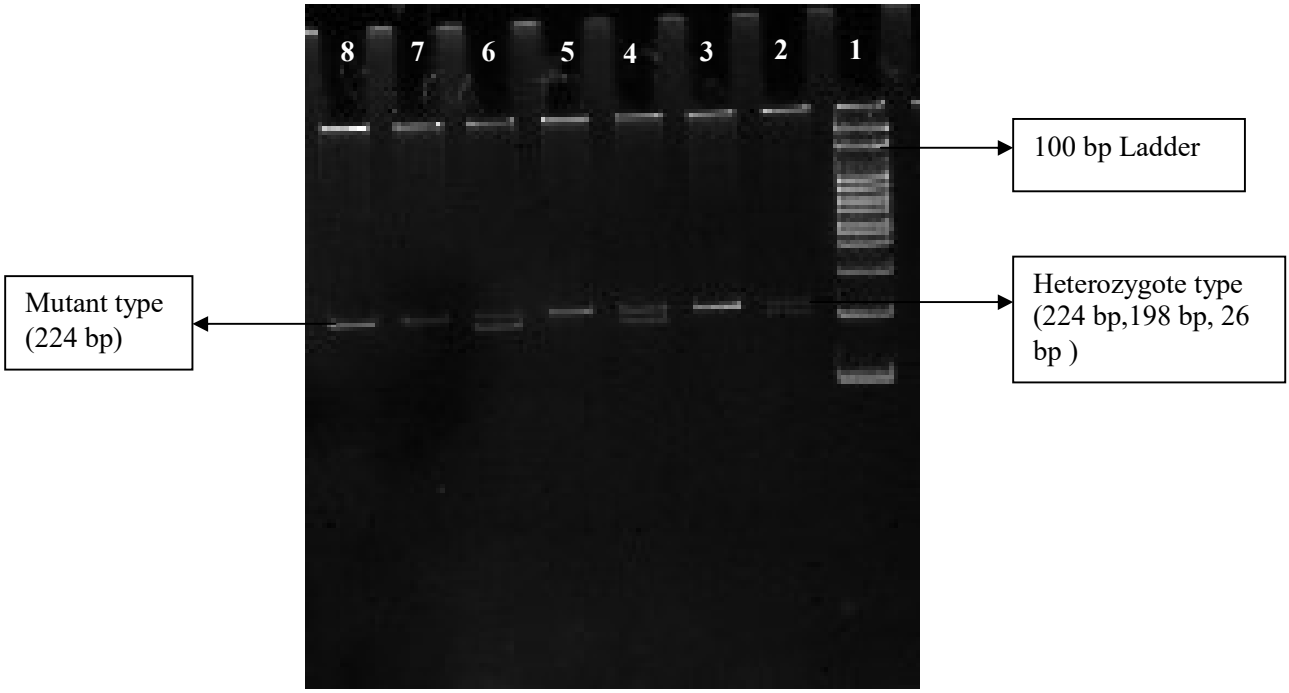


Figure 5.5 Representative sample for Restriction digestion of *ABCB1* rs2032582 PCR product showing the digested products (wild: 198 bp, 26 bp; mutant type: 224 bp; heterozygote type: 224 bp, 198 bp, 26 bp) Lane 1: 100 bp ladder; Lane 2,4,6: Heterozygote type; Lane 3,5,7,8: Mutant type

5.4. Demographic characteristics for *ABCB1* rs1128503 polymorphism

Table 5.1 Distribution of demographic features among Lung cancer patients for *ABCB1* rs1128503 polymorphism

Variable	**Cases, n(%) N = 192**	**Variable**	**Cases, n(%) N = 192**
Age (Years)		**Lymph node involvement**	
Mean±SD	58.06±9.95		
Range	28-86	N0	23 (11.97)
		N1	10 (5.20)
		N2	71 (36.97)
		N3	59 (30.72)
		N4	2 (1.04)
		Unknown	27 (14.06)
Gender		**Metastasis**	
Male	162 (84.37)	M0	95 (49.47)
Female	30 (15.62)	M1	72 (37.5)

		Unknown	27 (14.06)
Smoking status		**KPS**	
Smokers	162 (84.37)	>60	20 (10.41)
Non-smokers	26 (13.54)	70-80	90 (46.87)
		90-100	60 (31.25)
		Unknown	22 (11.45)
Pack years		**ECOG**	
Mean±SD	26.63±30.47	0	27 (14.06)
		1	62 (32.29)
		2	68 (35.41)
		3	19 (9.89)
		Unknown	16 (8.33)
Histological types		**Regimen**	
SQCC	65 (33.85)	1	35 (18.22)
ADCC	50 (28.64)	5	50 (26.04)
SCLC	74 (38.54)	6	34 (17.70)
Others	3 (1.5)	Others	11 (5.72)
		Unknown	62 (32.29)
Tumor size		**TNM staging**	
T1	7 (3.64)	I	3 (1.56)
T2	21 (10.93)	II	5 (2.60)
T3	41 (21.35)	III	82 (42.70)
T4	93 (48.43)	IV	78 (40.62)
Unknown	28 (14.58)	Unknown	24 (12.5)
Objective Response			
CR	6 (3.12)		
PR	45 (23.43)		
SD	36 (18.75)		
PD	13 (6.77)		
Unknown	92 (47.91)		

Abbreviations: ADCC = Adenocarcinoma; CR = Complete response; ECOG = Eastern Cooperative Oncology Group; KPS = Karnofsky performance status; N = Total no. of lung cancer cases; PD = Progressive disease; PR = Partial response; Regimen 1 = Docetaxel-plus-Carboplatin/Cisplatin; Regimen 5 = Irinotecan-plus-carboplatin/cisplatin; Regimen 6 = Pemetrexed-plus-carboplatin/cisplatin; SD = Stable disease; SD = Standard deviation; SCLC = Small-cell lung cancer; SQCC = Squamous cell carcinoma; TNM = Tumor node metastasis

The distribution of demographic characteristics in lung cancer patients recruited for the study of *ABCB1* rs1128503 is mentioned in Table 5.1. The parameters included are age, gender, smoking status, pack/year smoked, histology type, TNM staging and extension, tumor stage, clinical response to treatment, regimen preferred by patients and performance status. The mean age of

192 cases recruited for overall survival study was 58.06 with standard deviation of 9.65, range 28-86 years. There were more males on the study as compared to females' *viz.* 84.37% (162) and 15.62% (30) respectively. Percentage of smokers was much higher (84.37%) correlated to that of non-smokers (13.54%) which demonstrates that smoking is a significant factor in acquiring lung cancer. Pack years (calculated using formula [cigarettes or bidis smoked per day/20] x No. of years smoked) smoked was 26.63 with standard deviation of 30.47. Based on the classification of lung cancer according to histology, we found that 38.54% of patients had SCLC, followed by 33.85% of SQCC patients and 28.64% of Adenocarcinoma. Only 1.5% of patients had NSCLC and LCC. Coming to clinicopathological parameters such as Tumor Node Metastatic staging and extension, data of 165 patients was available. The size of tumor varies from T1 to T4. Majority of the patients had T4 type of tumor (48.43%) followed by 21.35% of T3 type, 10.93% and 3.64% of T2 and T1 type respectively. Among 165 cases, 95 patients (49.47%) showed no metastasis, whereas 72 (37.5%) patients showcased distant metastasis. Further, no lymph node extension was observed in 11.97% of cases. N2 and N3 type lymph node involvement was recognized in 71 (36.97%) and 59 (30.72%) cases respectively. Minimal patients reflected N4 lymph node extension of 1.04%. Significant number of cases demonstrated Stage III and Stage IV disease (42.70% and 40.62% respectively). Only 1.56% and 2.60% of patients' disease was detected in Stage I and Stage II respectively. Diseased stage of 12.5% of cases was not known. Clinical response of 92 (47.91%) cases could not be assessed. Among the 100 cases whose clinical response were noted, 3.12% showed complete response (CR), 23.43% of patients showed partial response (PR), 18.75% and 6.77% exhibited stable disease (SD) and progressive disease (PD) respectively. Additionally, on categorizing the patients on the basis of regimen received, it was observed that majority of the patients were under medication of regimen 5 which is Irinotecan-carboplatin/cisplatin (26.04%), while 18.22% of cases received Docetaxel-carboplatin/cisplatin (regimen 1) based treatment and 17.70% of patients were under Pemetrexed-carboplatin/cisplatin (regimen 6) medication. 5.72% of patients were under other medications such as gemcitabine-carboplatin/cisplatin and others. Regimen of 62 (32.29%) of cases were not known. Demographic characteristics also considered the performance status of the patients and spotted that 10.41% of cases had KPS (Karnofsky Performance status) below 70, 46.87% of patients had KPS score between 70-80 and 31.25% had this score between 90-100 showing a better functional condition of the patient. Performance status according to ECOG

(Eastern Cooperative Oncology Group) reflected 81.77% of patients with a better ECOG score of 0, 1 and 2, while 9.89% cases showed ECOG 3 and 4. For the remaining 16 (8.33%) patients, ECOG was not classified. No significance difference was observed in the genotypic distribution of *ABCB1* rs1128503 polymorphism from that expected in Hardy-Weinberg equilibrium ($p^2+2pq+q^2 = 1$, $\chi^2 = 1.4417$, P = 0.229).

5.5. Demographic characteristics for *ABCB1* rs2032582 polymorphism

Table 5.2 Distribution of demographic features among Lung cancer patients for *ABCB1* rs2032582 polymorphism

Variable	**Cases, n(%) N = 153**	**Variable**	**Cases, n(%) N = 153**
Age (Years)		**Lymph node involvement**	
Mean±SD	58.21±10.22		
Range	28-85	N0	11 (7.18)
		N1	10 (6.53)
		N2	62 (40.52)
		N3	42 (27.45)
		N4	1 (0.65)
		Unknown	27 (17.64)
Gender		**Metastasis**	
Male	129 (84.31)	M0	70 (45.75)
Female	24 (15.68)	M1	57 (37.25)
		Unknown	26 (16.99)
Smoking status		**KPS**	
Smokers	123 (80.39)	>70	16 (10.45)
Non-smokers	26 (16.99)	70-80	75 (49.01)
Unknown	4 (2.61)	90-100	50 (32.67)
		Unknown	12 (7.84)
Pack years		**ECOG**	
Mean±SD	25.25±27.15	0	25 (16.33)
		1	40 (26.14)
		2	57 (37.25)
		3	19 (12.41)
		Unknown	12 (7.84)
Histologic types		**Regimen**	
SQCC	51 (33.33)	1	38 (24.83)
ADCC	45 (29.41)	5	33 (21.56)

SCLC	50 (32.67)	6	32 (20.91)
Others	3 (1.96)	Others	8 (5.22)
Unknown	3 (1.96)	Unknown	44 (28.75)
Tumor size		**TNM staging**	
T1	9 (5.88)	I, II	4 (2.61)
T2	16 (10.45)	III	66 (43.13)
T3	28 (18.30)	IV	57 (37.25)
T4	71 (46.40)	Unknown	28 (18.30)
Unknown	29 (18.95)		
Objective Response			
CR	4 (2.61)		
PR	39 (25.49)		
SD	34 (22.22)		
PD	5 (3.26)		
Unknown	73 (47.71)		

Abbreviations: ADCC = Adenocarcinoma; CR = Complete response; ECOG = Eastern Cooperative Oncology Group; KPS = Karnofsky performance status; N = Total no. of lung cancer cases; PD = Progressive disease; PR = Partial response; Regimen 1 = Docetaxel-plus-Carboplatin/Cisplatin; Regimen 5 = Irinotecan-plus-carboplatin/cisplatin; Regimen 6 = Pemetrexed-plus-carboplatin/cisplatin; SD = Stable disease; SD = Standard deviation; SCLC = Small-cell lung cancer; SQCC = Squamous cell carcinoma; TNM = Tumor node metastasis

The distribution of demographic characteristics in lung cancer patients recruited for the study of *ABCB1* rs2032582 is mentioned in Table 5.2. The parameters included are age, gender, smoking status, pack/year smoked, histology type, TNM staging and extension, tumor stage, clinical response to treatment, regimen preferred by patients and performance status. The mean age of 153 cases recruited for overall survival study was 58.21 with standard deviation of 10.22, range 28-85 years. There were more males on the study as compared to females' *viz.* 84.31% (129) and 15.68% (24) respectively. Percentage of smokers was much higher (80.39%) correlated to that of non-smokers (16.99%) which demonstrates that smoking is a significant factor in acquiring lung cancer. Pack years (calculated using formula [cigarettes or bidis smoked per day/20] x No. of years smoked) smoked was 25.25 with standard deviation of 27.15. Based on the classification of lung cancer according to histology, we found that 33.33% of patients had SQCC, followed by 32.675% of SCLC patients and 29.41% of Adenocarcinoma. Only 1.96% of patients had NSCLC and LCC. Coming to clinicopathological parameters such as Tumor, Node and Metastasis staging and extension, size of tumor varies from T1 to T4. Majority of the patients had T4 type of tumor (46.40%) followed by 18.30% of T3 type, 10.45% and 5.88% of T2 and T1 type respectively. In 29 (18.95%) cases, the tumor size either could be detected or was not known. Among 127 cases,

70 patients (45.75%) showed no metastasis, whereas 57 (37.25%) patients showcased distant metastasis. Further, no lymph node extension was observed in 7.18% of cases. N2 and N3 type lymph node involvement was recognized in 62 (40.52%) and 42 (27.45%) cases respectively. Minimal patients reflected N4 lymph node extension (0.65%). Significant number of cases demonstrated Stage III and Stage IV disease (43.30% and 37.25% respectively). Only 2.61% of patients' disease was detected in Stage I and Stage II. Diseased stage of 18.30% of cases was not known. Clinical response of 73 (47.71%) cases could not be assessed. Among the 80 cases whose clinical response were noted, 2.61% showed complete response (CR), 25.49% of patients showed partial response (PR), 22.22% and 3.26% exhibited stable disease (SD) and progressive disease (PD) respectively. Additionally, on categorizing patients on the basis of regimen received, it was observed that majority of the patients were under medication of regimen 1 which is Docetaxel-carboplatin/cisplatin (24.83%), while 21.56% of cases received Irinotecan-carboplatin/cisplatin (regimen 5) based treatment and 20.91% of patients were under Pemetrexed-carboplatin/cisplatin (regimen 6) medication. 5.22% of patients were under other medications such as gemcitabine-carboplatin/cisplatin and others. Regimen of 44 (28.75%) of cases were not known. Demographic characteristics also considered the performance status of the patients and spotted that 10.45% of cases had KPS (Karnofsky Performance status) below 70, 49.01% of patients had KPS score 70-80 and 32.67% had this score 90-100 showing a better functional condition of the patient. Performance status according to ECOG (Eastern Cooperative Oncology Group) reflected 79.73% of patients with a better ECOG score of 0, 1 and 2, while 12.41% cases showed ECOG 3 and 4. For the remaining 12 (7.84%) patients, ECOG was not classified. Significant difference was observed in the genotypic distribution of *ABCB1* rs2032582 polymorphism from that expected in Hardy-Weinberg equilibrium ($p^2+2pq+q^2 = 1$, $\chi^2 = 19.697$, P = 0.000009).

5.6 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients

Table 5.3 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients

Dominant model (TT) in North Indian population							
***ABCB1* rs1128503 genotype**	**Dead 118 (n %)**	**Alive 21 (n %)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	49 (41.52)	14 (66.66)	10.26	1.0 (Reference)	-	1.0 (Reference)	-
TC	51 (43.22)	5 (23.8)	8.03	1.22 (0.82 to 1.81)	0.30	1.30 (0.82 to 2.07)	0.25
CC	18 (15.25)	2 (9.52)	8.40	1.13 (0.64 to 1.97)	0.65	1.27 (0.70 to 2.31)	0.42
TC+CC	69 (58.47)	7 (33.33)	8.20	1.20 (0.83 to 1.72)	0.31	1.24 (0.82 to 1.88)	0.30
Recessive model (CC) in North Indian population							
***ABCB1* rs1128503 genotype**	**Dead 118 (n%)**	**Alive 21 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	18 (15.25)	2 (9.52)	8.4	1.0 (Reference)	-	1.0 (Reference)	-
CT	51 (43.22)	5 (23.8)	8.03	1.07 (0.63 to 1.82)	0.79	1.34 (0.74 to 2.43)	0.33
TT	49 (41.52)	14 (66.66)	10.26	0.88 (0.50 to 1.54)	0.65	0.78 (0.43 to 1.41)	0.42
CT+TT	100 (84.74)	19 (90.47)	9.00	0.96 (0.58 to 1.60)	0.88	0.96 (0.57 to 1.64)	0.90

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

The strength of association between *ABCB1* rs1128503 (Dominant model TT and recessive model CC for North Indian population) and *ABCB1* rs2032582 polymorphism, demographic characteristics and *o*verall survival of lung cancer patients was evaluated using Crude HR and 95% CI obtained from Univariate Kaplan-Meier survival analysis. P-value obtained from Kaplan-Meier analysis and Log rank test was considered significant if p-value < 0.05.

Multivariate Cox proportional hazards regression analysis was applied to calculate adjusted HR and 95% CI to analyze the possible predictive variables for overall survival in lung cancer patients after adjusting for age, gender, ECOG, KPS, histology, smoking status and stage in order to study the independent impact of each variable on OS. All the calculations were performed by using software MedCalc version 18.5. The outcomes of the respective analysis for 192 cases are summarized in Table 5.3 for both dominant and recessive models. Prognostic study included evaluation of the cases from first day of diagnosis to endpoint and it was found that 15.10% (21) patients were alive whereas 84.89% patients died due to the disease.

In dominant model TT, genotype frequency as analyzed by our data obtained, 63 (45.32%) patients carrying wild genotype TT, 77.77% were dead and 22.22% were alive. Patients carrying the homozygous wild genotype TT had overall survival period of 10.26 months which was taken as reference. 14.38% (20) of the patients carried variant genotype CC with overall survival less than that of wild genotype (MST = 8.4 months). 40.28% (56) were carriers of heterozygous genotype TC which reflected a dip in their overall survival of as less as 8.033 months. As the mutant genotype was less prevalent, we combined it with heterozygous genotype as a single genotype, which was then compared with homozygous genotype TT taken as reference.

With wild genotype TT as reference having MST of 10.26 months, we found that the subjects carrying heterozygous and variant genotype comparatively had a lower survival time (MST = 8.03 and 8.40, HR = 1.22 and 1.13, P = 0.30 and 0.65 respectively). However, no significant association was found in dominant model between overall survival of the subjects and genotype of *ABCB1* rs1128503 polymorphism on evaluation with Cox proportional hazards regression analysis as observed in Figure 5.6 and 5.7 below.

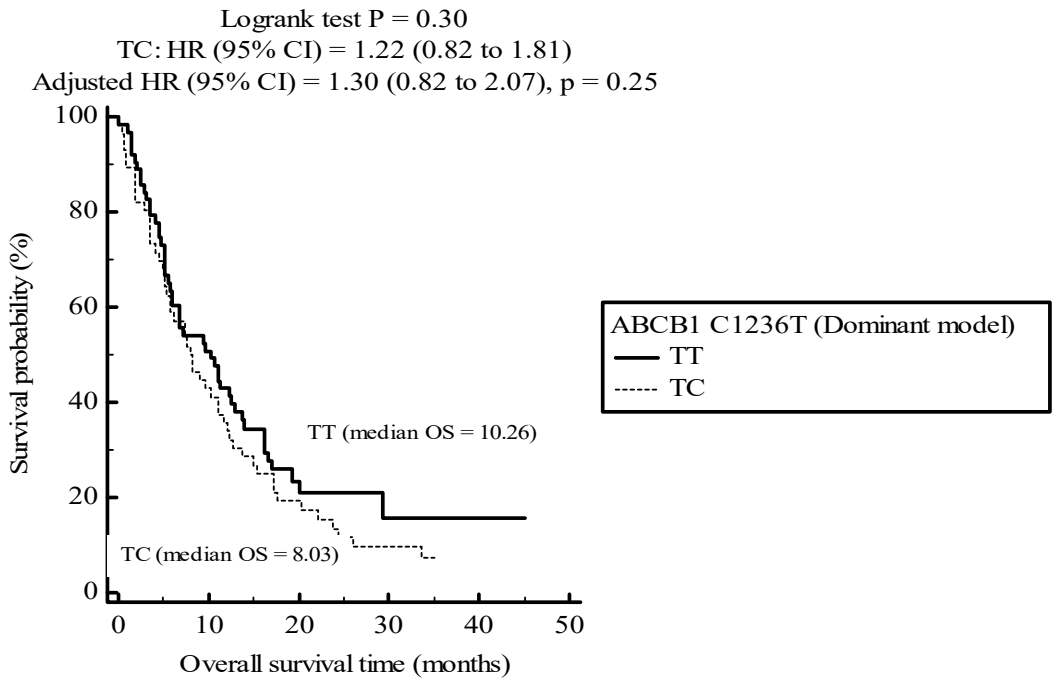


Figure 3.6 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 polymorphism (dominant model) on overall survival of lung cancer patients with heterozygous genotype (TC vs. TT)

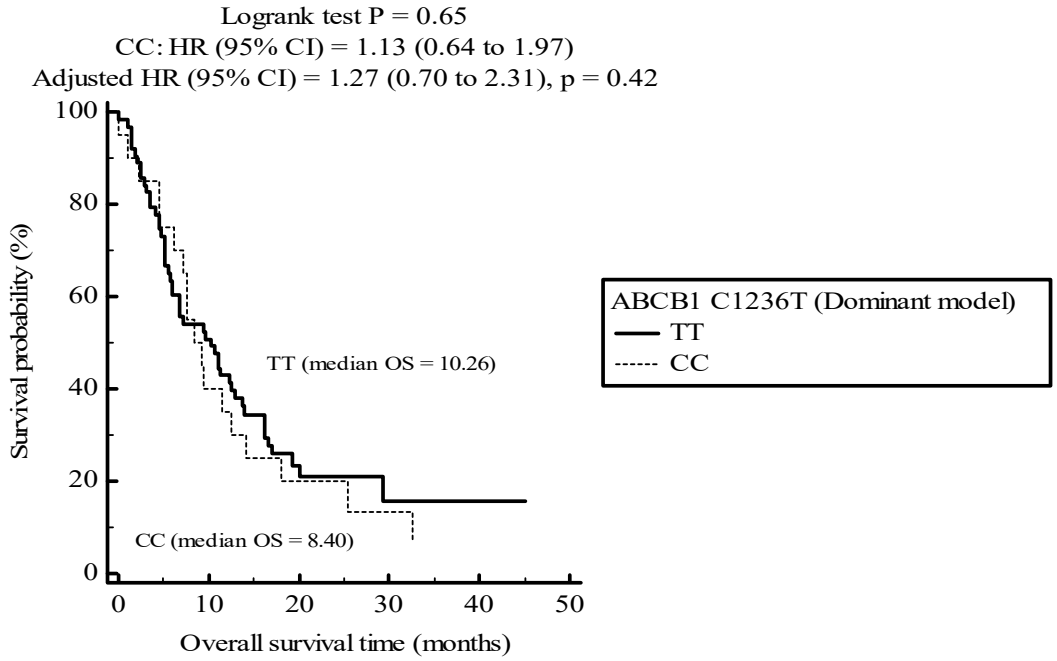


Figure 5.7 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 polymorphism (dominant model) on overall survival of lung cancer patients with mutant genotype (CC vs. TT)

In recessive model CC, genotype frequency as analyzed by our data obtained, of the 20 (14.38%) patients carrying wild genotype CC, 90% were dead and 10% were alive. Patients carrying the homozygous wild genotype CC had overall survival period of 8.4 months which was taken as reference. 45.32% (63) of the patients carried variant genotype TT with overall survival much higher than that of wild genotype (MST = 10.26 months). 40.28% (56) were carriers of heterozygous genotype TC which reflected a dip in their overall survival of as less as 8.033 months. On combining the mutant genotype with heterozygous genotype and comparing it against wild genotype CC, we found a better MST (9 months vs. 8.4 months). However, no significant association was found in recessive model between overall survival of the subjects and genotype of *ABCB1* rs1128503 polymorphism on evaluation with Cox proportional hazards regression analysis.

Table 5.4 Genotypic distribution and Relationship of *ABCB1* rs2032582 polymorphism with overall survival of lung cancer patients

***ABCB1* rs2032582 genotype**	**Dead 108**	**Alive 22**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	3	1	9.40	1.0 (Reference)	-	1.0 (Reference)	-
GT	68	10	7.33	1.43 (0.53 to 3.82)	0.53	0.39 (0.11 to 1.35)	0.14
TT	37	11	9.36	1.06 (0.34 to 3.35)	0.91	0.58 (0.15 to 2.26)	0.44
GT+TT	105	21	7.60	1.28 (0.46 to 3.57)	0.66	0.46 (0.13 to 1.54)	0.21

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Genotypic frequency of *ABCB1* Rs2032582 polymorphism as analyzed by our data obtained, 4 (3.07%) patients carrying wild genotype GG, 75% were dead and 25%% were alive. Patients carrying the homozygous wild genotype GG had overall survival period of 9.4 months which was taken as reference. 36.922% (48) of the patients carried variant genotype TT showing a overall survival similar to wild genotype (9.36 months vs. 9.4 months, HR = 1.06, P = 0.14). 60% (78) subjects were carriers of heterozygous genotype GT which showed a dip in their overall survival (7.333 vs. 9.4 months, HR = 1.43, P = 0.53). On combining the mutant genotype with

heterozygous genotype and comparing it against wild genotype (GT + TT vs. GG) , no significant improvement in MST was observed (7.6 vs. 9.4 months, HR = 1.28, P = 0.66). However, no significant association was found between overall survival of the subjects and genotype of *ABCB1* rs2032582 polymorphism on evaluation with Cox proportional hazards regression analysis as observed in Figure 5.8 and 5.9 below.

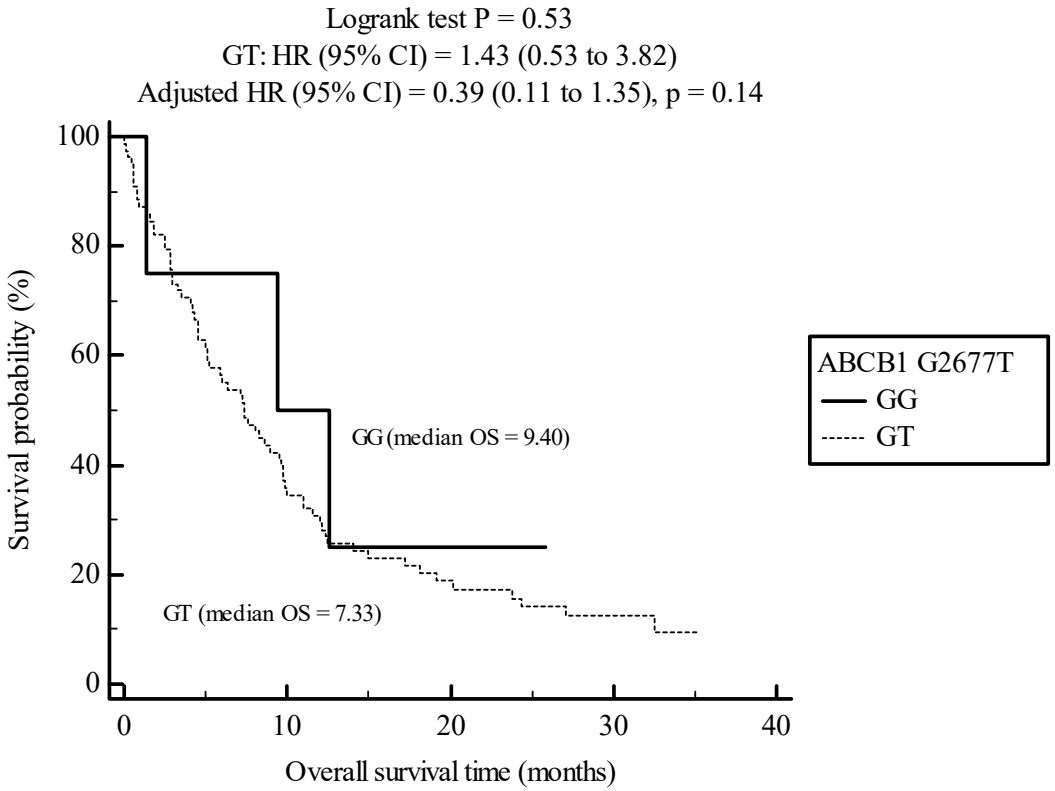


Figure 5.8 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of lung cancer patients with heterozygous genotype (GT vs. GG)

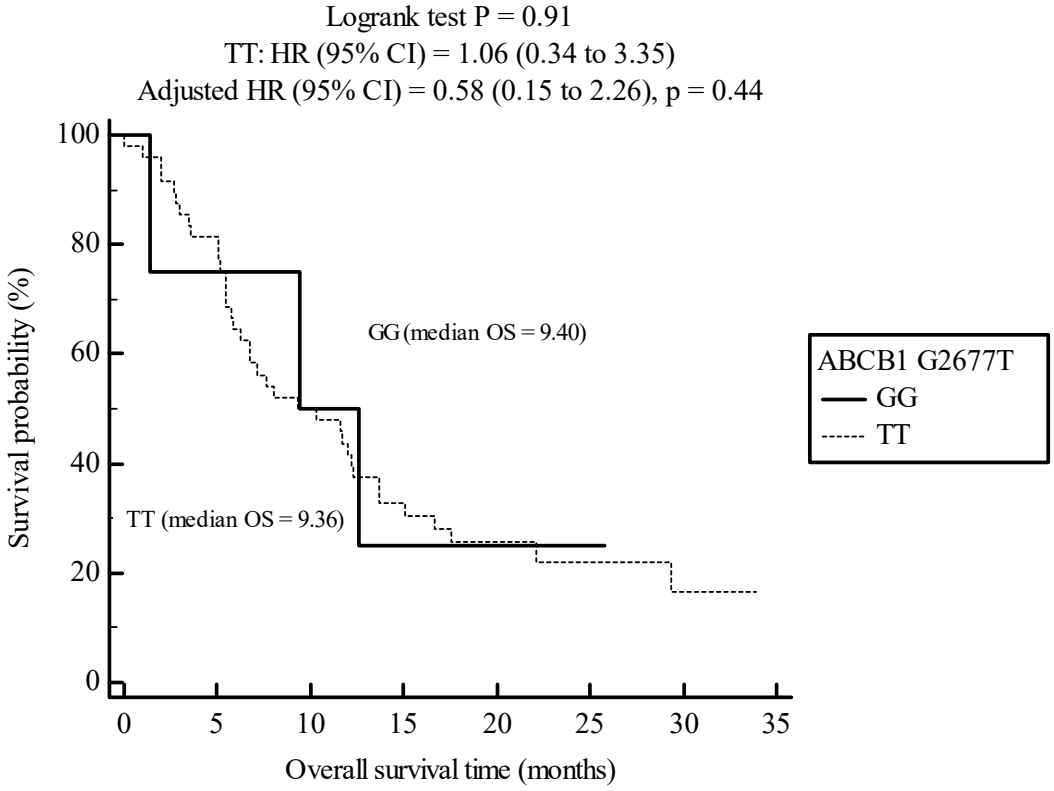


Figure 5.9 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of lung cancer patients with mutant genotype (TT vs. GG)

5.7 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients based on histological subtypes

Table 5.5 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of histological type (Dominant model (TT) for North Indian population)

ADCC							
***ABCB1* rs1128503 genotype**	**Dead 35 n%**	**Alive 6 n%**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	15 (75)	5 (25)	5.46	1.0 (Reference)	-	1.0 (Reference)	-
TC	13 (92.85)	1 (7.14)	9.56	0.90 (0.42 to 1.88)	0.77	1.16 (0.43 to 3.08)	0.75
CC	7 (100)	0 (0)	7.56	1.42 (0.54 to 3.75)	0.42	2.45 (0.60 to 10.00)	0.21
TC+CC	20 (95.23)	1 (4.76)	9.23	1.07 (0.54 to 2.08)	0.83	1.25 (0.55 to 2.82)	0.58
SCLC							
***ABCB1* rs1128503 genotype**	**Dead 44 (n%)**	**Alive 5 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	20 (90.9)	2 (9.09)	11.00	1.0 (Reference)	-	1.0 (Reference)	-
TC	18 (90)	2 (10)	5.80	1.06 (0.56 to 2.02)	0.83	0.95 (0.41 to 2.20)	0.91
CC	6 (85.71)	1 (14.28)	11.53	0.62 (0.27 to 1.43)	0.29	0.87 (0.29 to 2.60)	0.81
TC+CC	24 (88.88)	3 (11.11)	8.26	0.93 (0.51 to 1.69)	0.82	0.95 (0.46 to 1.94)	0.89
SQCC							
***ABCB1* rs1128503 genotype**	**Dead 38 (n%)**	**Alive 10 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	13 (65)	7 (35)	11.00	1.0 (Reference)	-	1.0 (Reference)	-
TC	20 (90.9)	2 (9.09)	7.33	1.69 (0.85 to 3.35)	0.13	2.30 (0.83 to 6.36)	0.10
CC	5 (83.33)	1 (16.66)	7.56	1.16 (0.39 to 3.38)	0.76	1.14 (0.26 to 4.93)	0.85
TC+CC	25 (89.28)	3 (10.71)	7.56	1.58 (0.83 to 3.00)	0.17	1.83 (0.73 to 4.53)	0.19

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage. ADCC = Adenocarcinoma; SCLC = Small cell lung cancer; SQCC = Squamous cell carcinoma.

In the cases included for overall survival association with *ABCB1* rs1128503 polymorphism, 38.54% (74) of the cases were diagnosed with SCLC, 33.85% (65) had SQCC and 28.64% (50) cases were of ADCC. Further, on classification of subjects on the basis of genotype frequency, it was found that 48.78%, 44.89% and 41.66% of ADCC, SCLC and SQCC cases respectively were carriers of homozygous wild genotype TT. 34.14%, 48.78% and 45.83% of ADCC, SCLC and SQCC patients respectively has heterozygous genotype GT. Mutant genotype was observed in 17.07%, 17.07% and 12.5% of ADCC, SCLC and SQCC population cases respectively.

In ADCC, patients with homozygous wild genotype TT showed MST of 5.467 months which was taken as a reference. Heterozygous genotype TC carrying patients showed an increased survival period (9.56 vs. 5.46 months, HR = 0.90, Log rank P = 0.77). Mutant genotype CC carrying patients had survival better than wild genotype but lower than heterozygous genotype carrying patients (7.56 vs. 5.46 months, HR = 1.42, Log rank P = 0.42). The combination of mutant and heterozygous genotype still reflected a better survival period (9.23 vs. 5.46 months, HR = 1.07, P = 0.83). According to the results obtained by Kaplan-Meier survival analysis, no significant link was observed between overall survival of patients with ADCC, SCLC and SQCC and genotype distribution.

The above data (Table 5.5) summarizes the statistical values obtained for dominant model TT for North Indian population showing the association between C/T polymorphism of *ABCB1* rs1128503 gene and overall survival of lung cancer patients on the basis of histological types. However, on performing Cox proportional hazards regression analysis, insignificant adjusted HR and 95% CI were obtained. (C1236T: ADCC: TC vs. TT, Adjusted HR = 1.16, P = 0.75; CC vs. TT: Adjusted HR = 2.45, P = 0.21; TC+CC vs. TT: adjusted HR = 1.25, P = 0.58), (SCLC: TC vs. TT, Adjusted HR = 0.95, P = 0.91; CC vs. TT: Adjusted HR = 0.87, P = 0.81; TC+CC vs. TT: adjusted HR = 0.95, P = 0.89)and (SQCC: TC vs. TT, Adjusted HR = 2.30, P = 0.10; CC vs. TT: Adjusted HR = 1.14, P = 0.85; TC+CC vs. TT: adjusted HR = 1.83, P = 0.19).

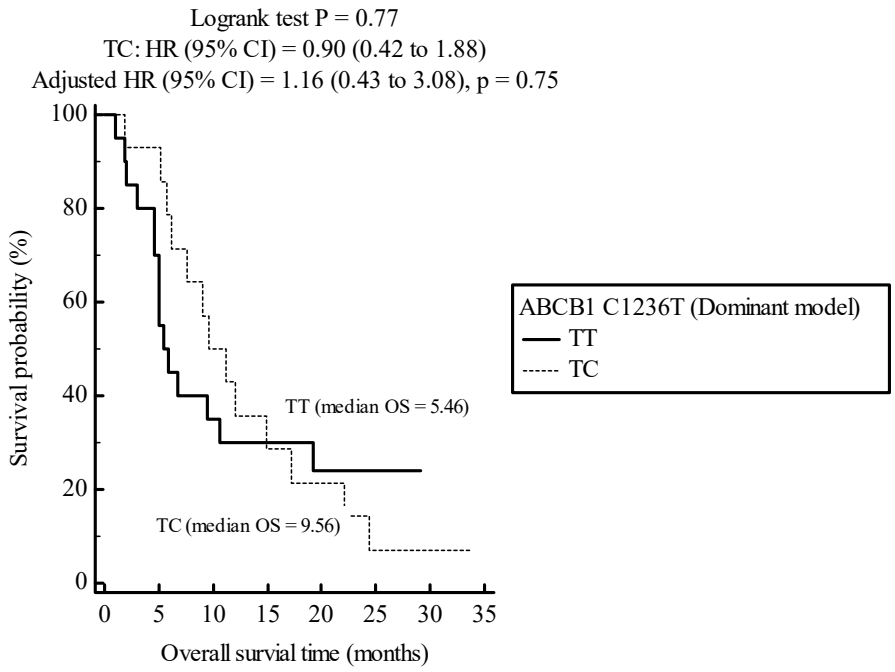


Figure 5.10 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of ADCC patients with heterozygous genotype (TC vs. TT)

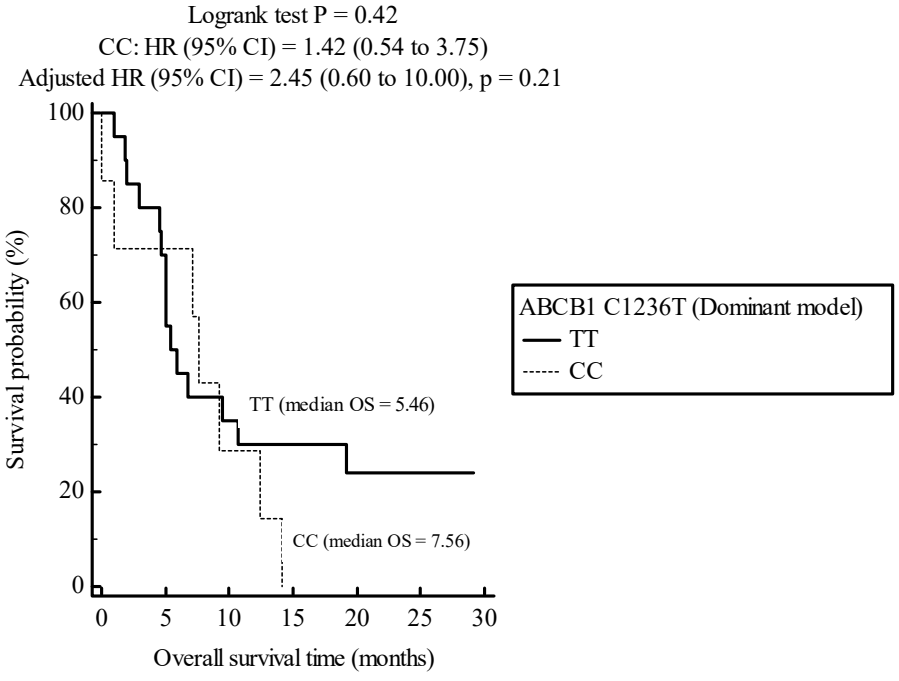


Figure 5.11 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of ADCC patients with mutant genotype (CC vs. TT)

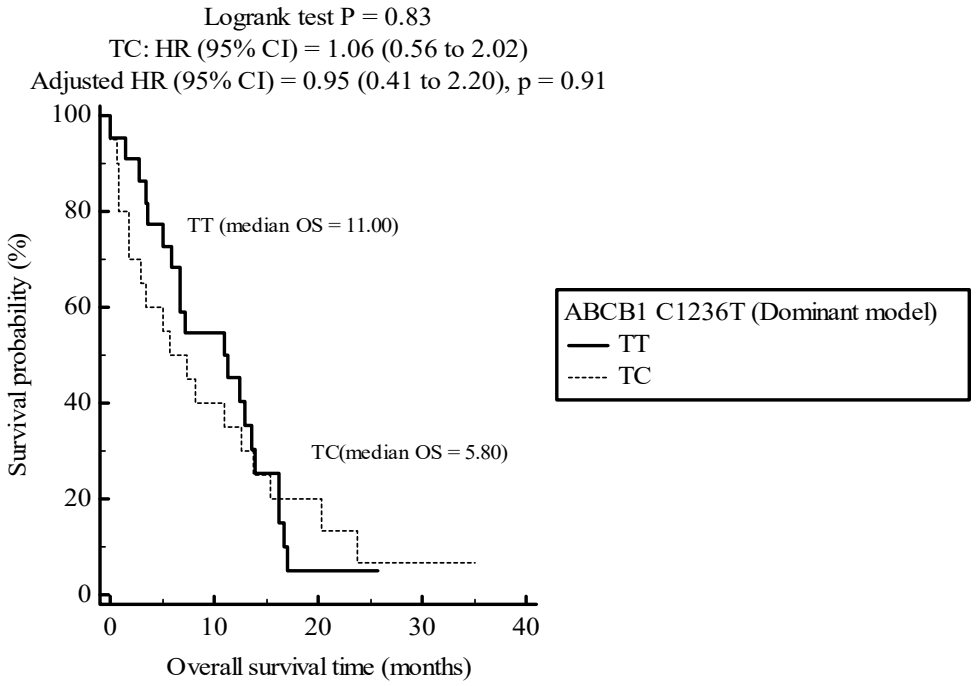


Figure 5.12 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of SCLC patients with heterozygous genotype (TC vs. TT)

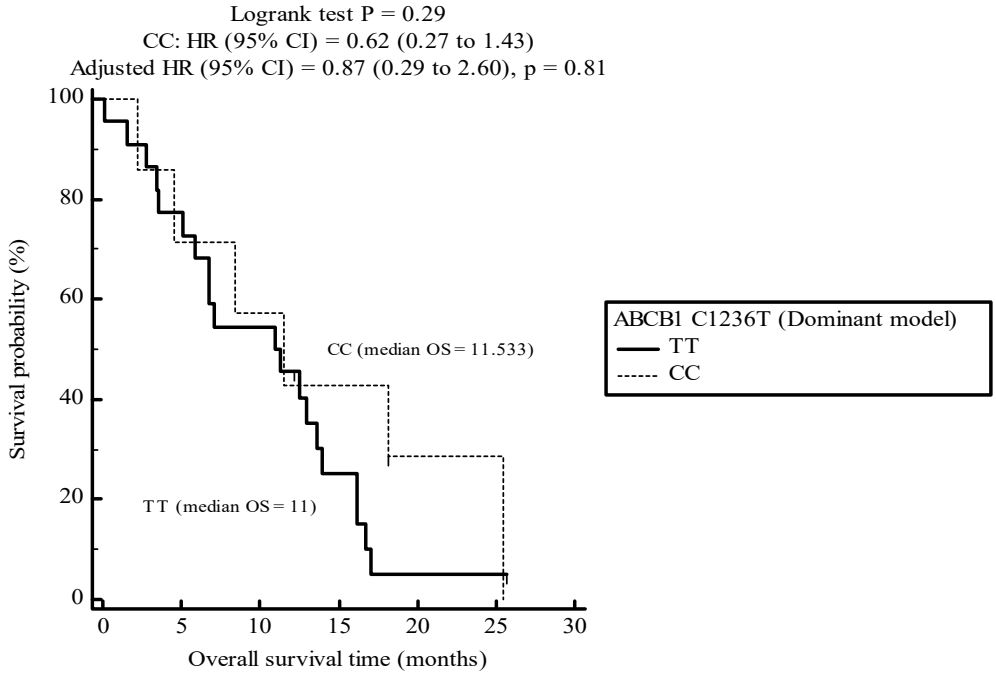


Figure 5.13 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of SCLC patients with mutant genotype (CC vs. TT)

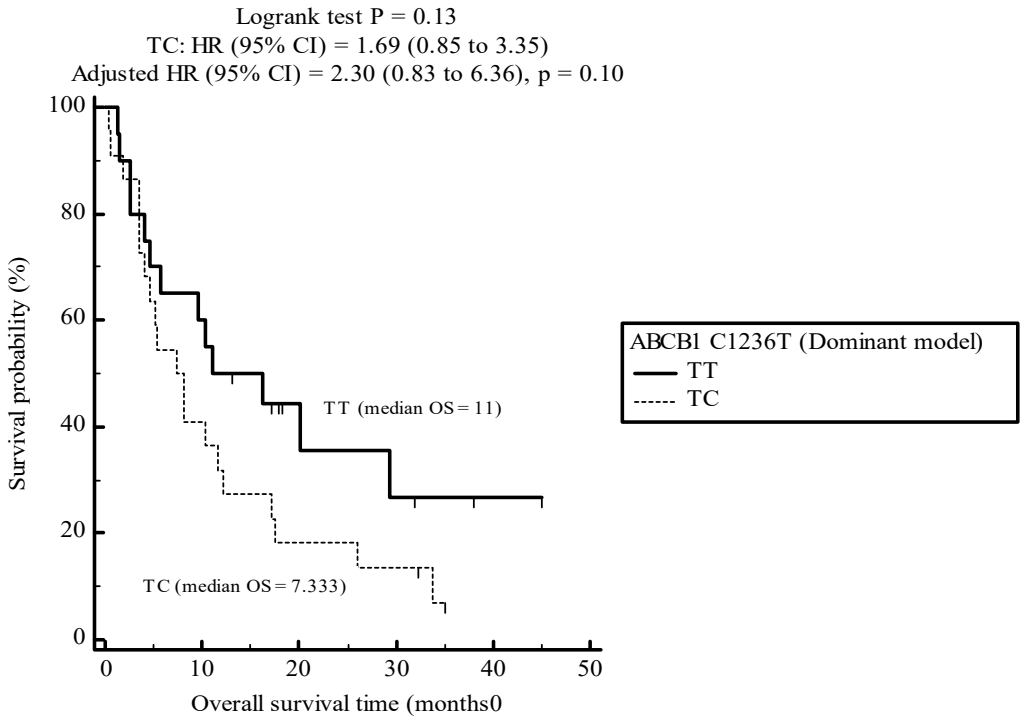


Figure 5.14 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of SQCC patients with heterozygous genotype (TC vs. TT)

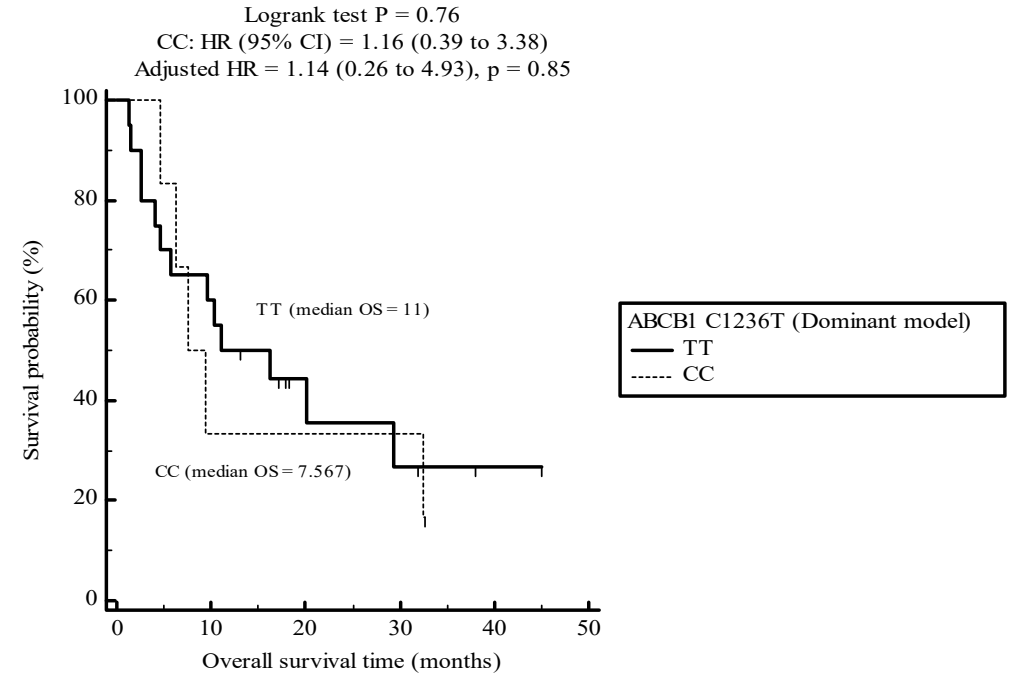


Figure 5.15 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of SQCC patients with mutant genotype (CC vs. TT)

Table 5.6 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of histological type (Recessive model CC in North Indian population)

ADCC							
ABCB1* rs1128503 *genotype	**Dead 35 (n%)**	**Alive 6 (n%)**	**Median OS months**	**HR (95% CI)**	**Log p**	**Adjusted HR (95% CI)**	**P**
CC	7	0	7.56	1.0 (Reference)	-	1.0 (Reference)	-
CT	13 (92.85)	1 (7.14)	9.56	0.50 (0.17 to 1.47)	0.12	0.59 (0.13 to 2.72)	0.50
TT	15 (75)	5 (25)	5.46	0.70 (0.26 to 1.85)	0.42	0.40 (0.09 to 1.65)	0.21
CT+TT	28 (82.35)	6 (17.64)	7.60	0.60 (0.22 to 1.61)	0.22	0.58 (0.19 to 1.71)	0.32
SCLC							
ABCB1* rs1128503 *genotype	**Dead 44 (n%)**	**Alive 5 (n%)**	**Median OS months**	**HR 95% CI**	**Log p**	**Adjusted HR 95% CI**	**P**
CC	6 (85.71)	1 (14.28)	11.53	1.0 (Reference)	-	1.0 (Reference)	-
CT	18 (90)	2 (10)	5.80	1.49 (0.63 to 3.49)	0.38	1.51 (0.42 to 5.37)	0.52
TT	20 (90.9)	2 (9.09)	11.00	1.58 (0.69 to 3.61)	0.29	1.13 (0.38 to 3.38)	0.81
CT+TT	38 (90.47)	4 (9.52)	7.33	1.52 (0.72 to 3.21)	0.32	1.29 (0.50 to 3.34)	0.58
SQCC							
ABCB1* rs1128503 *genotype	**Dead 38 (n%)**	**Alive 10 (n%)**	**Median OS months**	**HR 95% CI**	**Log p**	**Adjusted HR 95% CI**	**P**
CC	5 (83.33)	1 (83.33)	7.56	1.0 (Reference)	-	1.0 (Reference)	-
CT	20 (90.9)	2 (9.09)	7.33	1.34 (0.54 to 3.31)	0.54	2.18 (0.70 to 6.78)	0.17
TT	13 (65)	7 (35)	11.00	0.86 (0.29 to 2.50)	0.76	0.87 (0.20 to 3.77)	0.85
CT+TT	33	9	9.66	1.07	0.87	1.11	0.83

	(78.57)	(21.42)		(0.43 to 2.69)		(0.40 to 3.07)	

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage. ADCC = Adenocarcinoma; SCLC = Small cell lung cancer; SQCC = Squamous cell carcinoma.

The above data (Table 5.6) summarizes the statistical values obtained for recessive model CC for North Indian population showing the association between C/T polymorphism of *ABCB1* rs1128503 gene and overall survival of lung cancer patients on the basis of histological types. However, as shown in the table above, no significant association was observed between any of the genotypic combination stratified on the basis of histological subtypes with death rate or OS.

Table 5.7 Genotypic distribution and Relationship of *ABCB1* rs2032582 polymorphism with overall survival of lung cancer patients based on histological subtype

ADCC							
***ABCB1* rs2032582**	**Dead 34 (n%)**	**Alive 7 (n%)**	**Median OS months**	**HR 95% CI**	**Log p**	**Adjusted HR 95% CI**	**P**
GG	1	0	9.4	1.0 (Reference)	-	1.0 (Reference)	-
GT	22 (88)	3 (12)	7.56	0.88 (0.10 to 7.38)	0.90	0.02 (0.001 to 0.60)	**0.02**
TT	13 (76.47)	4 (23.52)	6.23	0.98 (0.12 to 7.67)	0.99	0.79 (0.06 to 10.46)	0.86
GT+TT	34 (82.92)	7 (17.07)	7.56	0.91 (0.11 to 7.28)	0.93	0.31 (0.02 to 3.54)	0.35
SCLC							
***ABCB1* rs2032582**	**Dead 32 (n%)**	**Alive 2 (n%)**	**Median OS months**	**HR 95% CI**	**Log p**	**Adjusted HR 95% CI**	**P**
GG	1	0	12.63	1.0 (Reference)	-	1.0 (Reference)	-
GT	21 (91.30)	2 (8.69)	5.00	1.48 (0.27 to 8.01)	0.69	1.53 (0.15 to 14.73)	0.71
TT	12 (100)	0	6.76	1.30 (0.20 to 8.19)	0.79	0.43 (0.00 to 1156.13)	0.83
GT+TT	32 (94.11)	2 (5.88)	6.73	1.41 (0.26 to 7.71)	0.72	1.33 (0.15 to 11.62)	0.79
SQCC							
ABCB1	**Dead**	**Alive**	**Median**	**HR**	**Log p**	**HR**	**P**

rs2032582	39 (n%)	11 (n%)	OS (months)	95% CI		95% CI	
GG	1	0	1.36	1.0 (Reference)	-	1.0 (Reference)	-
GT	27 (87.09)	4 (12.90)	8.03	0.13 (0.0006 to 27.19)	**0.01**	0.21 0.02 to 2.30	0.20
TT	13 (65)	7 (35)	11.7	0.046 (0.000007 to 301.98)	**<0.0001**	0.00	0.95
GT+TT	38 (77.55)	11 (22.44)	9.73	0.08 (0.00008 to 74.56)	**0.001**	0.11 0.01 to 1.19	**0.07**

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage. ADCC = Adenocarcinoma; SCLC = Small cell lung cancer; SQCC = Squamous cell carcinoma.

In the cases included for overall survival evaluation for *ABCB1* rs2032582 polymorphism, 32.67% (50) of the cases were diagnosed with SCLC, 33.33% (51) had SQCC and 29.41% (45) cases were of ADCC. Further, on classification of data on the basis of genotype frequency, it was found that 2.4%, 2.9% and 2% of ADCC, SCLC and SQCC cases respectively had homozygous wild genotype TT. 60.97%, 67.64% and 62% of ADCC, SCLC and SQCC patients respectively has heterozygous genotype GT. Mutant genotype was observed in 41.46%, 35.29% and 42% of ADCC, SCLC and SQCC population cases respectively.

Table 5.7 demonstrates that none of the genotype combinations of *ABCB1* rs2032582 polymorphism has any significant correlation for any histology. However, ADCC patients who were carriers of heterozygous genotype GT had increased OS (MST = 7.567 months) with a significant decreased death rate (Adjusted HR = 0.02, 95% CI = 0.001 to 0.60, P = 0.023) as compare to ADCC patients who are carriers of mutant genotype TT (Adjusted HR = 0.79, 95% CI = 0.06-10.46, P = 0.86) obtained by multivariate Cox proportional hazards analysis.

Further, SQCC patients who were carriers of double allelic variant/ mutant genotype TT showed a significant increased OS (MST = 11.7 months) with decreased death rate (HR = 0.04, P <0.0001) as compared to that of heterozygous genotype carriers (MST = 8.033, HR = 0.13, P = 0.017). The joint combination of genotype (GT+TT) showed significant 7.11 times increased OS

(MST = 9.333 months) and a decreased death rate (HR = 0.08, P = 0.0012) as compared to homozygous wild type carriers (MST = 1.367 months).

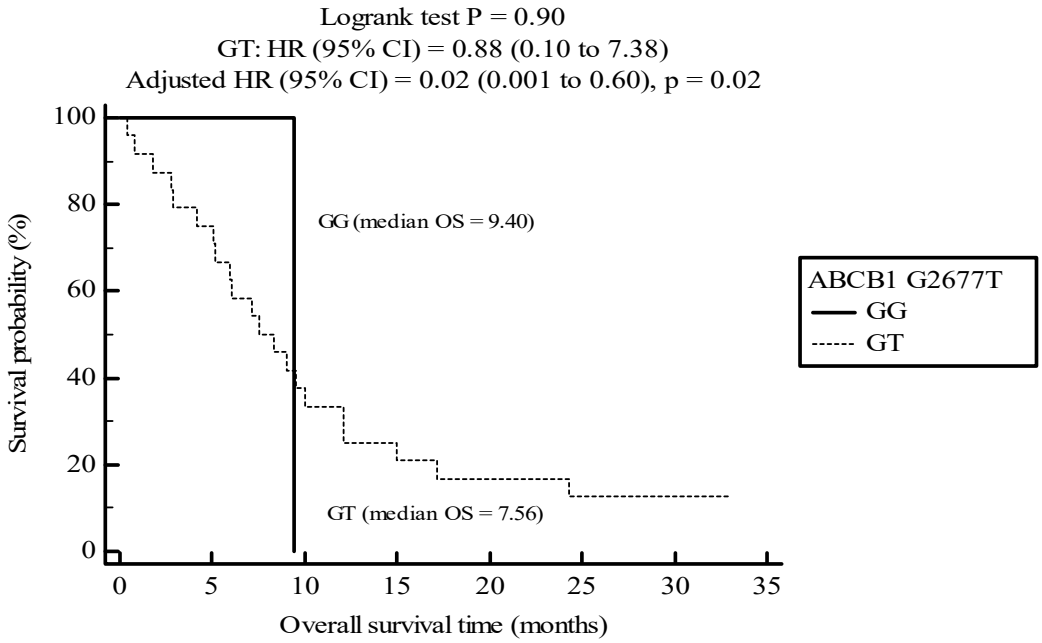


Figure 5.16 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of ADCC patients with heterozygous genotype (GT vs. GG)

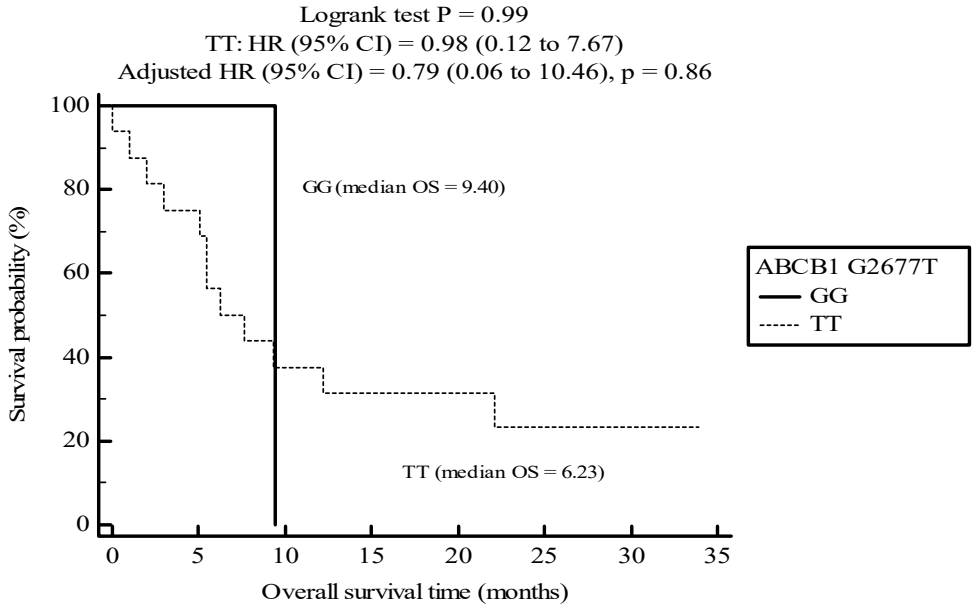


Figure 5.17 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of ADCC patients with mutant genotype (TT vs. GG)

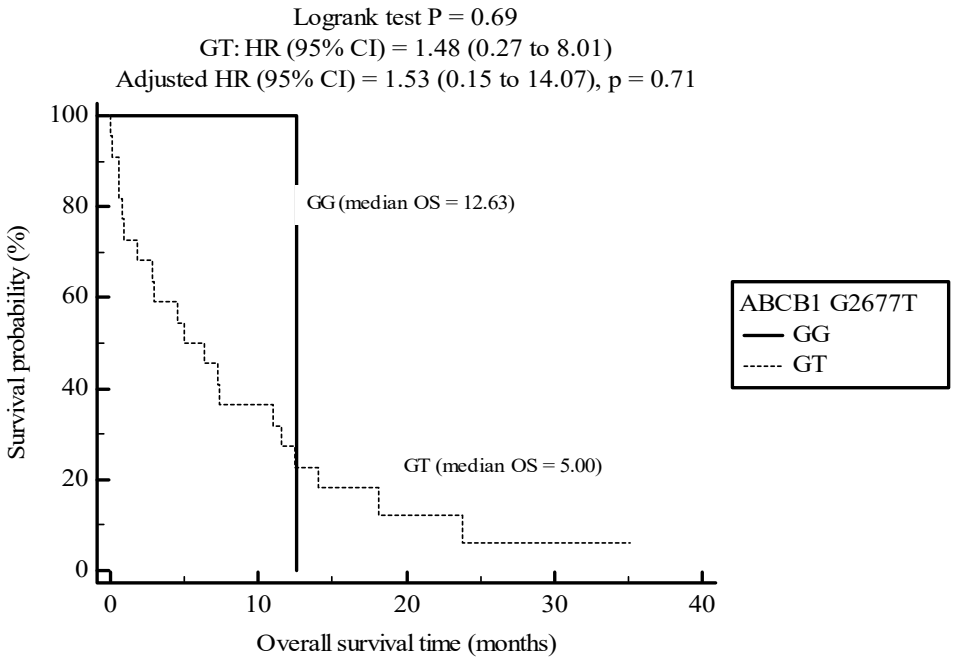


Figure 5.18 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of SCLC patients with heterozygous genotype (GT vs. GG)

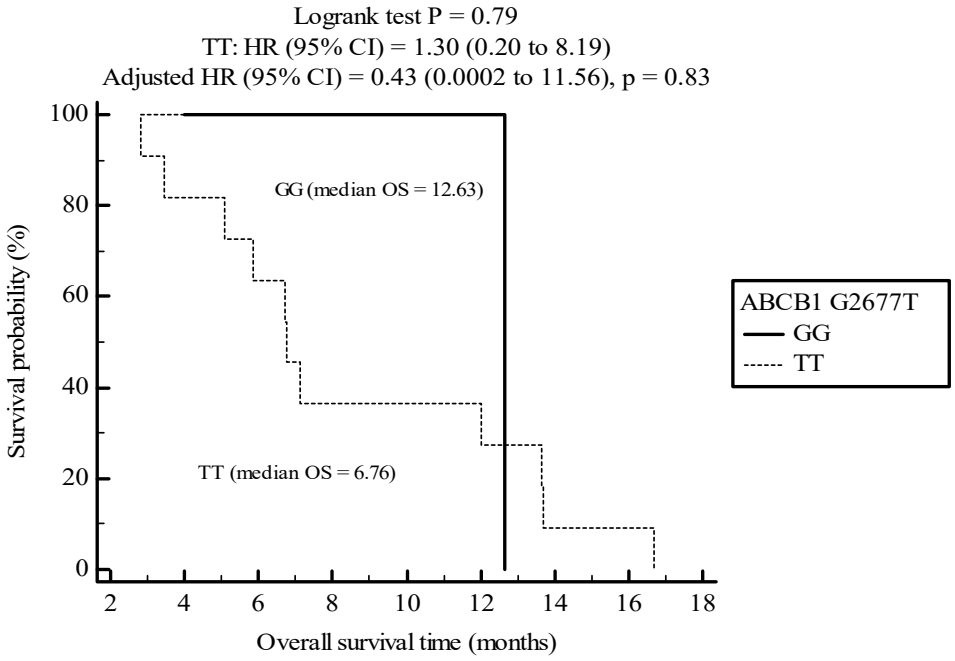


Figure 5.19 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of SCLC patients with mutant genotype (TT vs. GG)

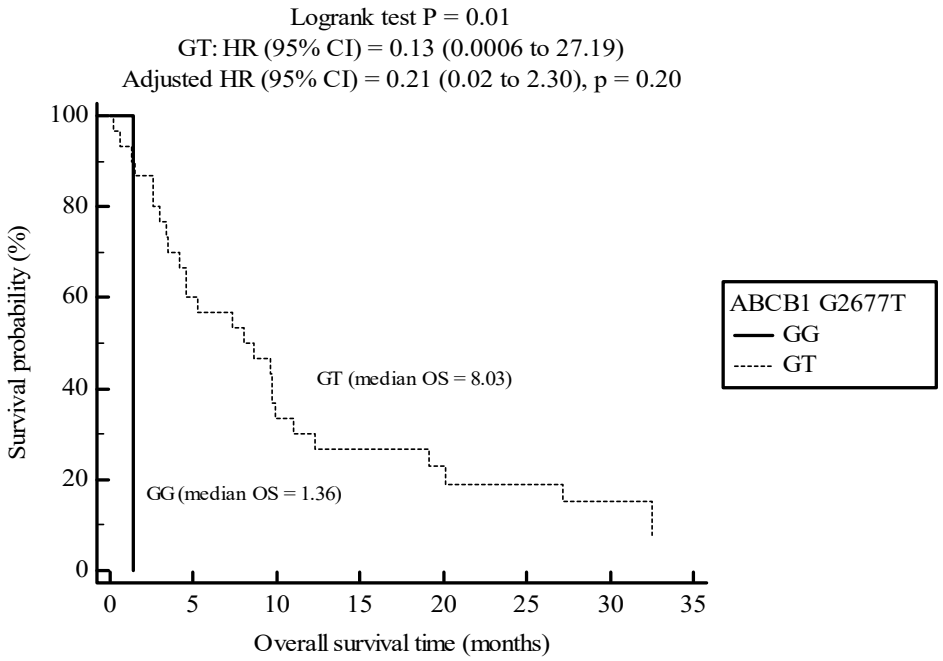


Figure 5.20 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of SQCC patients with heterozygous genotype (GT vs. GG)

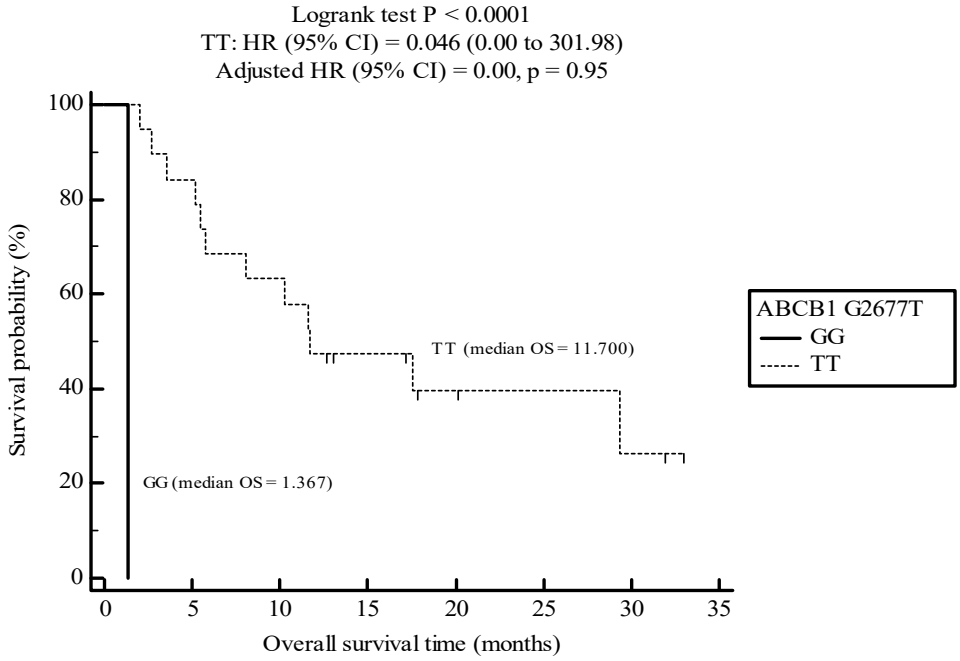


Figure 5.21 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of SQCC patients with mutant genotype (TT vs. GG)

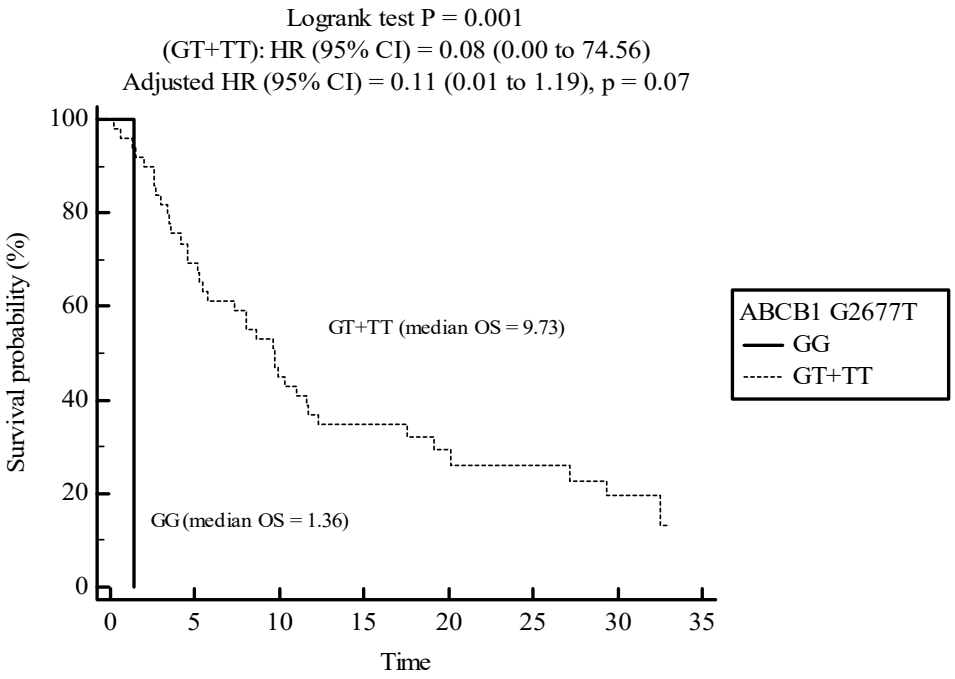


Figure 5.22 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs2032582 polymorphism on overall survival of SQCC patients with combined genotype (GT+TT vs. GG)

5.8 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients on the basis of gender

Table 5.8 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of gender (Dominant model TT in North Indian population)

Male							
***ABCB1* rs1128503 genotype**	**Alive 19 (n%)**	**Dead 97 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	12 (24)	38 (76)	11.00	1.00 (Reference)	-	1.00 (Reference)	-
TC	5 (10.63)	42 (89.36)	7.60	1.38 (0.89 to 2.16)	0.13	1.54 (0.91 to 2.60)	0.10
CC	2 (10.52)	17 (89.47)	9.23	1.17 (0.64 to 2.11)	0.58	1.12 (0.60 to 2.10)	0.70

TC+CC	7 (10.06)	59 (89.39)	8.03	1.32 (0.88 to 1.97)	0.17	1.31 (0.82 to 2.09)	0.25
Female							
***ABCB1* rs1128503 genotype**	**Alive 2 (n%)**	**Dead 21 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	2 (15.38)	11 (84.61)	6.733	1.00 (Reference)	-	1.00 (Reference)	-
TC	0	9	13.70	0.72 (0.30 to 1.75)	0.44	0.35 (0.07 to 1.81)	0.21
CC	0	1	4.50	2.49 (0.11 to 53.62)	0.35	25.64 (0.15 to 4276. 9)	0.21
TC+CC	0	10	12.06	0.79 (0.33 to 1.86)	0.56	0.68 (0.18 to 2.56)	0.57

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Table 5.9 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of gender (Recessive model CC in North Indian population)

Male							
***ABCB1* rs1128503 genotype**	**Alive 19 (n%)**	**Dead 97 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	2 (10.52)	17 (89.47)	9.23	1.00 (Reference)	-	1.00 (Reference)	-
CT	5 (10.63)	42 (89.36)	7.60	1.19 (0.69 to 2.05)	0.53	1.44 (0.78 to 2.66)	0.23
TT	12 (24)	38 (76)	11.00	0.85 (0.47 to 1.54)	0.58	0.88 (0.47 to 1.65)	0.70
CT+TT	17 (17.52)	80 (82.47)	9.00	0.99 (0.58 to 1.68)	0.98	1.09 (0.63 to 1.90)	0.73
Female							
***ABCB1* rs1128503 genotype**	**Alive 2 (n%)**	**Dead 21 (n%)**	**Median OS (months)**	**Crude HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	0	1	4.50	1.00 (Reference)	-	1.00 (Reference)	-

CT	0	9	13.70	0.31 (0.01 to 9.75)	0.22	0.005 (0.00 to 1263.4)	0.40
TT	2 (15.38)	11 (84.61)	6.73	0.40 (0.01 to 8.64)	0.35	0.03 (0.0002 to 6.50)	0.21
TC+TT	0	20	6.76	0.34 (0.01 to 9.51)	0.26	0.28 (0.01 to 4.78)	0.38

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Further stratification of the groups was done on the basis of gender and their association was evaluated with various genotype combinations of *ABCB1* rs1128503 polymorphism. Table 5.8 summarizes the statistical values obtained by performing Kaplan-Meier survival analysis and Cox proportional hazards regression analysis. Of the 192 cases undertaken for the study, 84.37% (162) were males and 15.62% (30) were females. Of the 162 male, 83.62% were dead and 16.37% were alive. While in case of 30 females, 91.30% of the females died and 8.69% were alive.

For dominant model TT for North Indian population, Male patients having homozygous wild genotype TT had a better survival period of 11 months as compared to patients carrying heterozygous allele TC (MST = 7.6 months, HR = 1.38, P = 0.13) or double variant allele CC (MST = 9.23 months, HR = 1.17, P = 0.58). In case of female patients, MST of 6.73 months was obtained statistically for female patients carrying wild genotype which was taken as reference. Female patients carrying heterozygous allele TC showed a 2-fold increase in OS (MST = 13.7 months, HR = 0.72) whereas those with mutant genotype TT showed a further decline in their OS (MST = 4.5, HR = 2.49, P = 0.35). However, insignificant values of HR and 95% CI were obtained for both multivariate and Univariate survival analysis demonstrating no association of OS or death rate of lung cancer patients with genotype distribution when stratified on the basis of gender. Similar insignificant results were obtained for recessive model CC of *ABCB1* rs1128503 polymorphism as mentioned in Table 5.9.

Table 5.10 Genotypic distribution and Relationship of *ABCB1* rs2032582 polymorphism with overall survival of lung cancer patients on the basis of gender (Male)

***ABCB1* rs2032582 genotype**	**Alive 20 (n%)**	**Dead 91 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	1 (25)	3 (75)	9.40	1.00 (Reference)	-	1.00 (Reference)	-
GT	10 (14.92)	57 (85.07)	8.03	1.35 (0.48 to 3.73)	0.60	0.46 (0.13 to 1.59)	0.22
TT	9 (22.5)	31 (77.5)	10.26	1.06 (0.33 to 3.37)	0.92	0.46 (0.11 to 1.89)	0.28
GT+TT	19 (17.75)	88 (82.24)	8.63	1.238 0.436 to 3.511	0.71	0.49 (0.14 to 1.67)	0.26

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Table 5.10 summarizes the statistical values obtained by performing Kaplan-Meier survival analysis and Cox proportional hazards regression analysis for ***ABCB1* rs2032582** polymorphism when stratified the groups in the basis of gender. Statistical data for female patients could not be evaluated as there was no reference (wild genotype GG). Of the 153 cases undertaken for the study, 84.31% (129) were males and 15.68% (24) were females. Of the 129 male, 81.98% were dead and 18.01% were alive. While in case of 24 females, 89.47% of the females died and 10.52% were alive.

Male patients having homozygous wild genotype GG had a survival period of 9.4 months which was taken as reference. Male subjects carrying double variant allele/mutant genotype TT showed increase in OS (MST = 10.267 months, HR = 1.06, P = 0.92) as compared to heterozygous genotype carrying patients (MST = 8.033, HR = 1.352, P = 0.60) and joint combination of genotype (GT+TT vs. GG) (MST = 8.633, HR = 1.238, P = 0.71) . However, insignificant values of HR and 95% CI were obtained for both multivariate and Univariate survival analysis demonstrating no association of OS or death rate of lung cancer patients with genotype distribution when stratified on the basis of gender.

5.9 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients on the basis of smoking status

Table 5.11 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of smoking status (Dominant model TT in North Indian population):

Non-smokers							
***ABCB1* rs1128503 genotype**	**Alive 2 (n%)**	**Dead 17 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	2 (25)	6 (75)	6.73	1.00 (Reference)	-	1.00 (Reference)	-
TC	0	7	12.06	0.91 (0.30 to 2.72)	0.85	0.08 (0.0006 to 11.61)	0.32
CC	0	4	0.96	2.81 (0.58 to 13.62)	0.08	1.00 (0.00 to 10.14231 E+303)	**<0.0001**
TC+CC	0	11	6.10	1.28 (0.48 to 3.37)	0.60	0.54 (0.07 to 4.10)	0.55
Smokers							
***ABCB1* rs1128503 Genotype**	**Alive 19 (n%)**	**Dead 100 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	12 (21.81)	43 (78.18)	10.66	1.00 (Reference)	-	1.00 (Reference)	-
TC	5 (10.41)	43 (89.58)	8.03	1.23 (0.80 to 1.88)	0.32	1.34 (0.81 to 2.21)	0.24
CC	2 (12.5)	14 (87.5)	9.23	0.96 (0.52 to 1.74)	0.89	1.12 (0.58 to 2.16)	0.71
TC+CC	7 (10.93)	57 (89.06)	8.26	1.15 (0.78 to 1.71)	0.46	1.22 (0.78 to 1.91)	0.37
Light-smokers							
***ABCB1* rs1128503 genotype**	**Alive 6 (n%)**	**Dead 65 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	4 (12.9)	27 (87.09)	7.13	1.00 (Reference)	-	1.00 (Reference)	-
TC	2 (6.66)	28 (93.33)	10.26	1.02 (0.60 to 1.73)	0.93	1.53 (0.77 to 3.05)	0.22
CC	0	10	7.56	1.16 (0.54 to 2.48)	0.67	1.40 (0.59 to 3.35)	0.44

TC+CC	2 (5)	38 (95)	9.33	1.05 (0.64 to 1.72)	0.82	1.51 (0.82 to 2.78)	0.18
Heavy-smokers							
***ABCB1* rs1128503 Genotype**	**Alive 10 (n%)**	**Dead 45 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	6 (26.08)	17 73.91)	10.26	1.00 (Reference)	-	1.00 (Reference)	-
TC	2 (8.69)	21 (91.30)	7.60	1.37 (0.72 to 2.60)	0.31	1.47 (0.64 to 3.38)	0.36
CC	2 (22.22)	7 (77.77)	9.23	0.88 (0.37 to 2.08)	0.78	0.78 (0.28 to 2.17)	0.64
TC+CC	4 (12.5)	28 (87.5)	8.03	1.21 (0.67 to 2.19)	0.52	1.08 (0.51 to 2.24)	0.83
Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.							

The subjects enlisted for association study of OS of lung cancer patients with genotype distribution were further categorized on the basis of smoking status as smokers and non-smokers. Of 192 cases recruited for *ABCB1* rs1128503 polymorphism association study, 84.37% of the cases included smokers whereas 13.54% cases consisted of non-smokers. Smokers were further categorized as light smokers and heavy smokers on the basis number of packs (bidi/cigarettes) smoked per day which was calculated using the formula:

$$\text{Pack years} = \frac{No.of\ Cigarretes\ or\ beedi\ smoked\ per\ day}{20}\ x\ No.of\ years\ smoked$$

Individuals who smoked more than or equal to 20 packs per day were classified as heavy smokers whereas those of smoked less than 20 packs per day were classified as light smokers. 44.79% (86) were light smokers whereas 55.20% (106) were heavy smokers.

Table 5.11 summarizes the statistical values obtained to study the association between OS of lung cancer patients stratified on the basis of smoking status with genotype distribution for dominant model for TT of *ABCB1* rs1128503 polymorphism in North Indian population.

In dominant model TT, among the non-smokers carrying the homozygous wild genotype TT, 75% were dead and 25% alive, showcasing a survival period of 6.73 months. Presence of

heterozygous genotype TC accounted for a 2-fold increase in OS (MST = 12.06, HR = 0.91, Log rank P = 0.85, Adjusted HR = 0.08, P = 0.32). However, as no significant values were obtained by Cox proportional hazard regression analysis, this is not proven statistically. Non-smokers with mutant genotype CC showcased a highly significant value pertaining to decreased overall survival of the patients (MST = 0.96, HR = 2.81, Log rank test P = 0.08, Adjusted HR = 1.00, P < 0.00001). As the number of mutant genotype carrying subjects was less, a joint combination of heterozygous and mutant genotype of subjects was compare to the wild genotype. However, no significant association was found even here (MST = 6.1 months, HR = 1.28, Log rank P = 0.6, Adjusted HR = 0.54, P = 0.55).

Further, on relating *ABCB1* rs1128503 genotype with overall survival of smokers, it was found that, out of the 55 cases carrying homozygous wild genotype, 78.18% were dead and 21.18% alive (MST = 10.66). A decline in overall survival period was observed in patients carrying heterozygous genotype (MST = 8.03 months, HR = 1.23, Log rank P = 0.32) and mutant genotype (MST = 9.23 months, HR = 0.96, Log rank P = 0.89). However, this association could not be proved either as no statistically significant value was obtained.

Among light smokers, of the 31 cases with homozygous genotype TT, 87.09% died due to the disease whereas 12.9% survived (MST = 7.13 months). The survival period was higher in patients with heterozygous genotype TC (MST = 10.26 months, HR = 1.02, Log rank P = 0.93) and mutant genotype CC (MST = 7.56 months, HR = 1.16 Log rank P = 0.67). Also, on comparing combined genotype (TC+CC) with wild genotype, an increased OS was obtained (MST = 9.33 months, HR = 1.05, Log rank P = 0.82). Despite, on performing Cox proportional hazards analysis, insignificant statistical values were obtained denying the association between OS of lung with genotype distribution in light smokers as well.

Heavy smokers containing homozygous wild genotype TT (73.91% dead and 26.08% alive) had a better survival period of 10.267 months as compared to light smokers. Nevertheless, in heavy smokers likewise, no statistically significant values were obtained suggesting no association of OS of heavy smokers with genotype distribution.

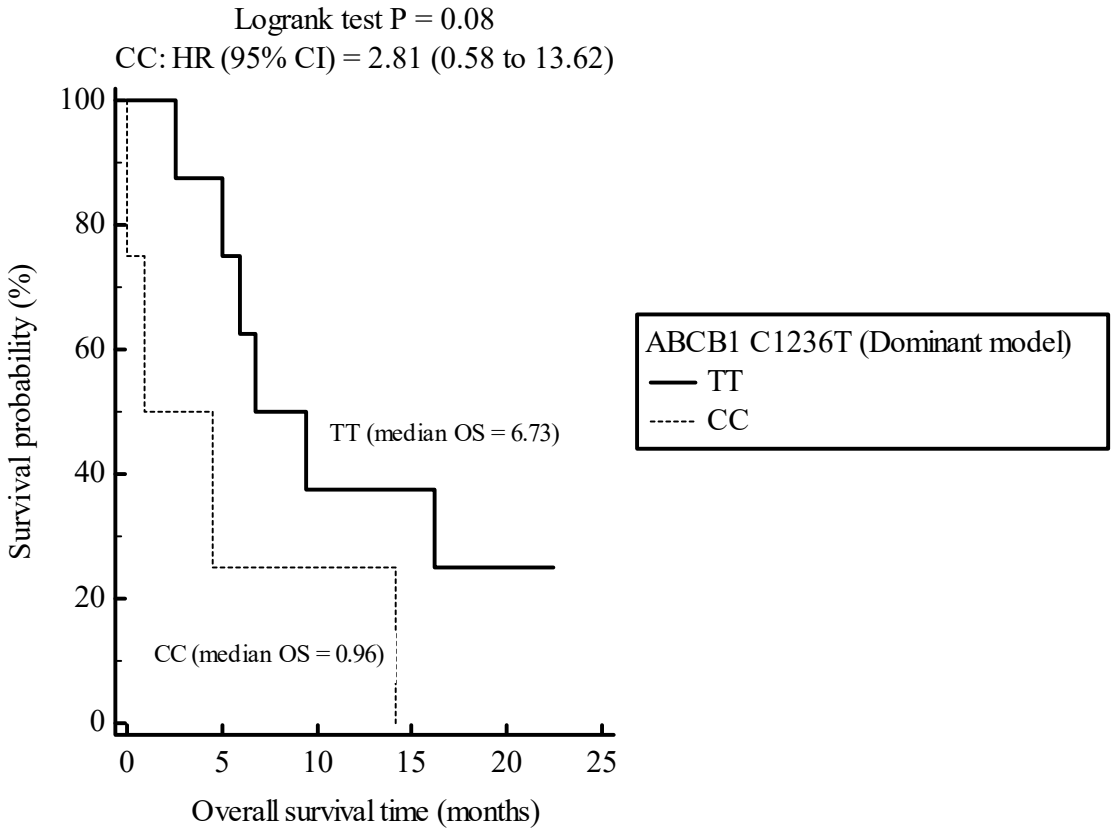


Figure 5.23 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (Dominant model) polymorphism on overall survival of lung cancer patients in non-smokers with mutant genotype (CC vs. TT)

Table 5.12 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of smoking status (Recessive model for CC in North Indian population)

Non-smokers							
***ABCB1* rs1128503 genotype**	**Alive 2 (n%)**	**Dead 17 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	0	4	0.96	1.00 (Reference)	-	1.00 (Reference)	-
CT	0	7	12.06	0.36 (0.07 to 1.74)	0.07	0.0001 (3.26E-014 to 78267.95)	0.36
TT	2 (25)	6 (75)	6.73	0.35 (0.07 to 1.71)	0.08	0.18 (0.00 to 10.14E+303)	1.00
CT+TT	2 (13.33)	13 (86.66)	9.46	0.327 (0.06 to 1.70)	**0.03**	1.74 (0.21 to 14.02)	0.60

Smokers							
ABCB1* rs1128503 *genotype	**Alive 19 (n%)**	**Dead 100 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	2 (12.5)	14 (87.5)	9.23	1.00 (Reference)	-	1.00 (Reference)	-
CT	5 (10.41)	43 (89.58)	8.03	1.28 (0.72 to 2.27)	0.40	1.60 (0.81 to 3.13)	0.17
TT	12 (21.81)	43 (78.18)	10.66	1.03 (0.57 to 1.88)	0.89	0.88 (0.46 to 1.69)	0.71
CT+TT	17 (16.5)	86 (83.49)	9.00	1.14 (0.66 to 1.95)	0.64	1.10 (0.61 to 1.98)	0.73
Light-smokers							
***ABCB1* rs1128503 genotype**	**Alive 6 (n%)**	**Dead 65 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	0	10	7.56	1.00 (Reference)	-	1.00 (Reference)	-
CT	2 (6.66)	28 (93.33)	10.26	0.82 (0.38 to 1.76)	0.60	1.08 (0.46 to 2.54)	0.85
TT	4 (12.9)	27 (87.09)	7.13	0.85 (0.40 to 1.81)	0.67	0.71 (0.29 to 1.69)	0.44
CT+TT	6 (9.83)	55 (90.16)	9.46	0.83 (0.40 to 1.70)	0.58	0.81 (0.39 to 1.68)	0.58
Heavy-smokers							
***ABCB1 rs*1128503 genotype**	**Alive 10**	**Dead 45**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	2 (22.22)	7 (77.77)	9.23	1.00 (Reference)	-	1.00 (Reference)	-
CT	2 (8.69)	21 (91.3)	7.60	1.55 (0.71 to 3.40)	0.30	2.18 (0.83 to 5.67)	0.11
TT	6 (26.08)	17 (73.91)	10.26	1.13 (0.47 to 2.67)	0.78	1.27 (0.45 to 3.55)	0.64
CT+TT	8 (17.39)	38 (82.60)	8.03	1.32 (0.63 to 2.77)	0.48	1.45 (0.62 to 3.40)	0.38

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Table 5.12 summarizes the statistical values obtained to study the association between OS of lung cancer patients stratified on the basis of smoking status with genotype distribution for recessive model for CC of *ABCB1* rs1128503 polymorphism in North Indian population. In

Recessive model for CC in North Indian population, among the non-smokers carrying the homozygous wild genotype CC, 100% dead epidemiology was obtained which may be due to the small sample size, showcasing a survival period of 0.96 months. Presence of heterozygous genotype CT accounted for a 12-fold increase in OS (MST = 12.06, HR = 0.36, Log rank P = 0.07, Adjusted HR = 0.00001, P = 0.36). Non-smokers with mutant genotype TT showcased a comparative decreased overall survival of the patients (MST = 6.73, HR = 0.35, Log rank test P = 0.08, Adjusted HR = 0.18, P = 1.00). However, as no significant values were obtained by Cox proportional hazard regression analysis, this is not proven statistically. As the number of mutant genotype carrying subjects was less, a joint combination of heterozygous and mutant genotype (CT+TT) of subjects was compared to the wild genotype. A significant association was found between the OS of patients in non-smokers with combined genotype (MST = 9.467 months, HR = 0.32, Log rank P = **0.03**, Adjusted HR = 1.74, P = 0.60).

Further, on relating *ABCB1* rs1128503 genotype with overall survival of smokers, it was found that, out of the 16 cases carrying homozygous wild genotype, 87.5% were dead and 12.5% alive (MST = 9.23). A decline in overall survival period was observed in patients carrying heterozygous genotype (MST = 8.03 months, HR = 1.28, Log rank P = 0.40) and a rise in OS period of patients carrying mutant genotype was observed (MST = 10.66 months, HR = 1.03, Log rank P = 0.89). However, this association could not be proved either as no statistically significant value was obtained.

Among light smokers, 10 cases with homozygous genotype CC showcased a overall survival of 7.56 months. The survival period was higher in patients with heterozygous genotype CT (MST = 10.26 months, HR = 0.82, Log rank P = 0.60). Patients with mutant genotype TT showed similar OS (MST = 7.13 months, HR = 0.85, Log rank P = 0.67). Also, on comparing combined genotype (CT+TT) with wild genotype, enhanced OS was obtained (MST = 9.46 months, HR = 0.83, Log rank P = 0.58). Likewise, on performing Cox proportional hazards analysis, insignificant statistical values were obtained denying the association between OS of lung with genotype distribution in light smokers as well.

Heavy smokers containing homozygous wild genotype CC (77.77% dead and 22.22% alive) had a better survival period of 9.23 months as compared to light smokers. Nevertheless, in heavy smokers likewise, no statistically significant values were obtained suggesting no association of OS of heavy smokers with genotype distribution.

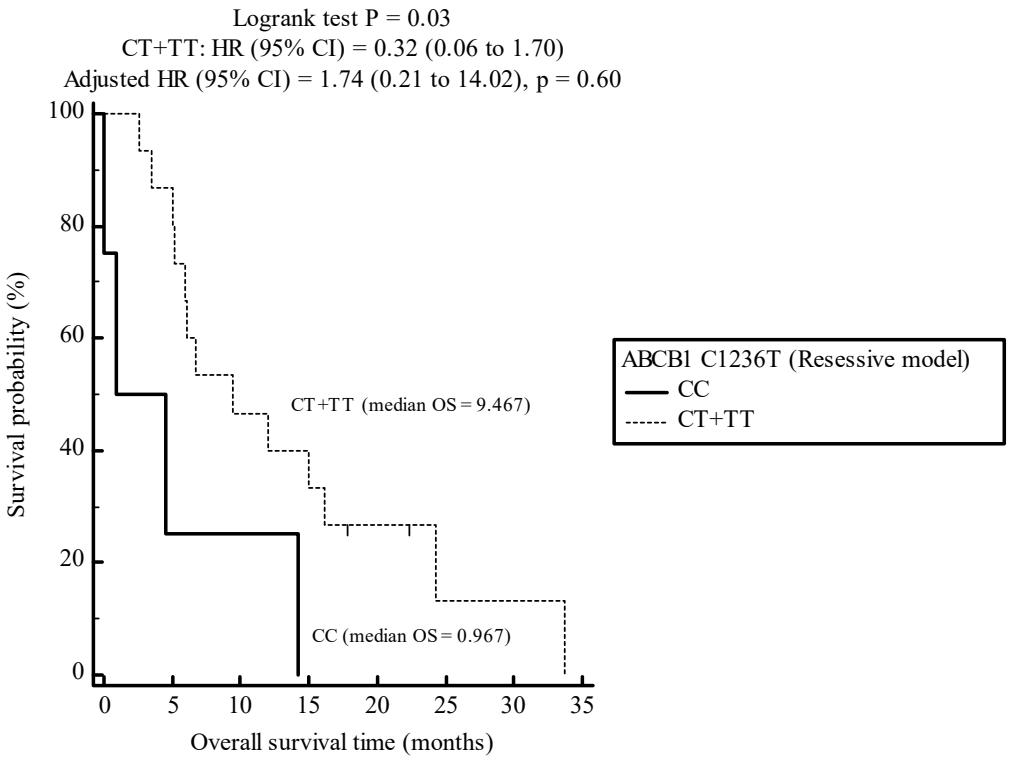


Figure 5.24 Kaplan-Meier survival curve illustrating Effect of *ABCB1* rs1128503 (recessive model) polymorphism on overall survival of lung cancer patients in non-smokers with combined genotype (CT+TT vs. CC)

Table 5.13 Genotypic distribution and Relationship of *ABCB1* rs2032582 polymorphism with overall survival of lung cancer patients based on smoking status

Non-smokers							
***ABCB1* rs2032582 genotype**	**Alive 6 (n%)**	**Dead 16 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	1 (50)	1 (50)	9.40	1.00 (Reference)	-	1.00 (Reference)	-
GT	2 (16.66)	10 (83.33)	5.13	2.52 (0.57 to 11.02)	0.35	0.03 (0.0002 to 7.11)	0.22
TT	3 (37.5)	5 (62.5)	6.23	1.96 (0.33 to 11.62)	0.52	-	-
GT+TT	5 (25)	15 (75)	5.93	2.34 (0.55 to 9.82)	0.39	0.27 (0.01 to 3.94)	0.34
Smokers							
***ABCB1* rs2032582**	**Alive 16**	**Dead 91**	**Median OS**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**

genotype	(n%)	(n%)	months				
GG	0	2	1.36	1.00 (Reference)	-	1.00 (Reference)	-
GT	8 (12.3)	57 (87.69)	7.56	0.78 (0.16 to 3.81)	0.73	0.52 (0.12 to 2.24)	0.38
TT	8 (20)	32 (80)	10.26	0.50 (0.07 to 3.55)	0.32	0.48 (0.09 to 2.43)	0.38
GT+TT	16 (15.23)	89 (84.76)	8.30	0.66 (0.11 to 3.65)	0.55	0.48 (0.11 to 2.04)	0.32
Light-smokers							
***ABCB1* rs2032582 Genotype**	**Alive 9 (n%)**	**Dead 62 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	0	1	9.40	1.00 (Reference)	-	1.00 (Reference)	-
GT	5 (11.62)	38 (88.37)	7.56	0.77 (0.08 to 7.30)	0.79	0.07 (0.005 to 1.02)	**0.05**
TT	4 (14.81)	23 (85.18)	10.26	0.66 (0.05 to 7.48)	0.68	1.02 (0.10 to 9.83)	0.98
GT+TT	9 (12.85)	61 (87.14)	8.06	0.72 (0.07 to 7.30)	0.74	0.39 (0.04 to 3.47)	0.40
Heavy-smokers							
***ABCB1* rs2032582 genotype**	**Alive 13 (n%)**	**Dead 46 (n%)**	**Median OS months**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	1 (33.33)	2 (66.66)	12.63	1.00 (Reference)	-	1.00 (Reference)	-
GT	5 (14.28)	30 (85.71)	7.33	1.89 (0.63 to 5.68)	0.36	0.70 (0.15 to 3.19)	0.65
TT	7 (33.33)	14 (66.66)	9.36	0.99 (0.22 to 4.39)	0.99	1.14 (0.10 to 12.51)	0.90
GT+TT	12 (21.42)	44 (78.57)	7.33	1.48 (0.45 to 4.87)	0.57	0.67 (0.15 to 2.97)	0.60

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-Meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Of 153 cases recruited for *ABCB1* rs2032582 polymorphism Overall survival study, 80.39% of the cases included smokers whereas 16.99% cases consisted of non-smokers.

Table 5.13 summarizes the statistical values obtained to study the association between OS of lung cancer patients stratified on the basis of smoking status with genotype distribution for G/T allele of *ABCB1 rs2033582* polymorphism. Among the non-smokers carrying the homozygous

wild genotype GG, 50% were dead and 50% alive, showcasing a survival period of 9.40 months. Presence of heterozygous genotype GT in patients showcased a dip in OS (MST = 5.13, HR = 2.58, Log rank P = 0.35, Adjusted HR = 0.03, P = 0.22). Non-smokers with mutant genotype TT showcased overall survival period of MST = 6.23 (HR = 1.96, Log rank test P = 0.52). However, as no significant values were obtained by Cox proportional hazard regression analysis, this is not proven statistically. As the number of mutant genotype was less, a joint combination of heterozygous and mutant genotype of subjects was compared to the wild genotype. However, no significant association was found even here (MST = 5.93 months, HR = 2.34, Log rank P = 0.39, Adjusted HR = 0.27, P = 0.34).

Further, on relating *ABCB1* rs2032582 genotype with overall survival of smokers, it was found that cases carrying homozygous wild genotype had a lower OS (MST = 1.367). A rise in overall survival period was observed in patients carrying heterozygous genotype (MST = 7.567 months, HR = 0.784, Log rank P = 0.734) and mutant genotype (MST = 10.267 months, HR = 0.5005, Log rank P = 0.329). Nevertheless, this association could not be proved either as no statistically significant value was obtained.

Among light smokers, subject with homozygous genotype GG had OS period of 9.400 months. The survival period was higher in patients with mutant genotype TT (MST = 10.267 months, HR = 0.66, Log rank P = 0.68). Nevertheless, this association could not be proved as no statistically significant value was obtained by Cox proportional hazards analysis. A statistically significant decline in OS period of population carrying heterozygous genotype GT was observed by performing Cox proportional hazard regression analysis (MST = 7.56 months, HR = 0.77, Log rank P = 0.79, Adjusted HR = 0.07, **P = 0.05**). Also, on comparing combined genotype (GT+TT) with wild genotype, a decreased OS was obtained (MST = 8.06 months, HR = 0.72, Log rank P = 0.74). Likewise, on performing Cox proportional hazards analysis, insignificant statistical values were obtained denying the association between OS of lung with genotype distribution in light smokers as well (Adjusted HR = 0.39, P = 0.40).

Heavy smokers containing homozygous wild genotype GG (66.66% dead and 33.33% alive) had a better survival period of 12.63 months as compared to light smokers. Nevertheless, in heavy smokers likewise, no statistically significant values were obtained suggesting no association of OS of heavy smokers with genotype distribution.

5.10 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients based on clinical stage

Table 5.14 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to clinical stage of lung cancer

Genotype	Clinical stage		AOR (95% CI)	P-value
	Stage III	Stage IV		
***ABCB1* rs1128503 (Dominant model TT)**	N= 79 (n%)	N= 81 (n%)		
TT	32 (40.5)	34 (41.9)	Reference	-
TC	33 (41.77)	35 (43.2)	1.19 (0.57 to 2.46)	0.63
CC	14 (17.7)	12 (14.8)	0.57 (0.18 to 1.76)	0.33
TC+CC	47 (59.49)	47 (58.02)	1.00 (0.50 to 2.00)	0.98
***ABCB1* rs1128503 (Recessive model CC)**	N= 79 (n%)	N= 81 (n%)		
CC	32 (40.5)	34 (41.9)	Reference	-
CT	33 (41.77)	35 (43.2)	1.78 (0.80 to 3.95)	0.15
TT	14 (17.7)	12 (14.8)	1.74 (0.56 to 5.39)	0.33
CT+TT	47 (59.49)	47 (58.02)	2.00 (0.73 to 5.48)	0.17
***ABCB1* rs2032582**	N= 66 (n%)	N= 57 (n%)		
GG	2 (3.03)	1 (1.75)	Reference	-
GT	41 (62.12)	36 (63.15)	1.96 (0.15 to 25.57)	0.60
TT	23 (34.84)	20 (35.08)	1.85 (0.12 to 27.48)	0.65
GT+TT	64 (96.96)	56 (98.24)	2.05 (0.16 to 25.89)	0.57

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 *GG* genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for clinical stage III and stage IV. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better than control arm provided significant P-value

Table 5.14 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to clinical stage of the disease with genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated included patients from stage III and stage IV.

Out of 192 cases included for study of *ABCB1* rs1128503 polymorphism, 42.71% (81) cases were at stage III and 40.62% (79) were at stage IV of the disease. As mentioned in the table above, of the 79 cases of Stage III patients, 40.5%, 41.77% and 17.7% were carriers of homozygous wild, heterozygous and mutant genotypes respectively. Whereas in case of 81 cases of stage IV patients, 41.9% were homozygous wild genotype, 43.2% had heterozygous genotype and 14.8% consisted of mutant genotype. However, association between genotypic distribution and associated risk with overall survival according to clinical stage cannot be proved as no significant statistical values were obtained by logistic regression analysis.

Out of 153 cases included for study of *ABCB1* rs2032582 polymorphism, 43.13% (66) cases were at stage III and 37.25% (57) were at stage IV of the disease. As mentioned in the table above, of the 66 cases of Stage III patients, 3.03%, 62.12% and 34.84% were carriers of homozygous wild, heterozygous and mutant genotypes respectively. Whereas in case of 57 cases of stage IV patients, 1.75% were homozygous wild genotype, 63.15% had heterozygous genotype and 35.08% consisted of mutant genotype. However, association between genotypic distribution and associated risk with overall survival according to clinical stage cannot be proved as no significant statistical values were obtained by logistic regression analysis.

5.11 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients based on their performance status after receiving chemotherapy

Table 5.15 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to KPS

Genotype	KPS		AOR (95% CI)	P-value
	KPS 90-100	KPS < 90		
***ABCB1* rs1128503 (Dominant model TT)**	N= 60 (n%)	N=110 (n%)		
TT	28 (46.66)	44 (40)	Reference	-
TC	22 (36.66)	51 (46.36)	1.57 (0.76 to 3.21)	0.21
CC	10 (16.66)	15 (13.63)	0.87 (0.32 to 2.35)	0.79
TC+CC	32 (53.33)	66 (60)	1.39 (0.72 to 2.67)	0.32
***ABCB1* rs1128503 (Recessive model CC)**	N= 60 (n%)	N=110 (n%)		
CC	10 (16.66)	15 (13.63)	Reference	-
CT	22 (36.66)	51 (46.36)	1.52 (0.54 to 4.24)	0.42
TT	28 (46.66)	44 (40)	1.14 (0.42 to 3.06)	0.79
CT+TT	50 (83.33)	95 (86.36)	1.22 (0.48 to 3.05)	0.66
***ABCB1* rs2032582**	N= 50 (n%)	N= 91 (n%)		
GG	1 (2)	3 (3.29)	Reference	-
GT	28 (56)	58 (63.73)	0.60 (0.04 to 7.48)	0.69
TT	21 (42)	30 (32.96)	0.22 (0.01 to 3.49)	0.28
GT+TT	49 (98)	88 (96.7)	0.46 (0.04 to 5.16)	0.53

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 GG genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for KPS. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm

is better than control arm provided significant P-value

Table 5.14 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to Karnofsky performance status of the patients with genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated included patients KPS score 90-100 who are considered as individuals with good performance status and KPS score 70-80 and <60 who are patients with decreased performance status.

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in KPS group 90-100 and KPS < 90 consisted of 46.66% and 40% of homozygous wild genotype respectively. 36.66% and 46.36% of patients had heterozygous genotype in KPS 90-100 group and KPS < 90 group respectively (OR = 1.57, P = 0.21). 16.66% and 13.63% individual were carriers of mutant genotype in KPS 90-100 and KPS < 90 groups respectively (OR = 0.87, p = 0.79). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of KPS score cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in KPS group 90-100 and KPS < 90 consisted of 2% and 3.29% of homozygous wild genotype respectively, muss less than that of *ABCB1* rs1128503 polymorphism. 56% and 63.73% of patients had heterozygous genotype in KPS 90-100 group and KPS < 90 group respectively (OR = 0.60, P = 0.69). 42% and 32.96% individual were carriers of mutant genotype in KPS 90-100 and KPS < 90 groups respectively (OR = 0.22, p = 0.28). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of KPS score cannot be proved.

Table 5.16 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to ECOG

Genotype	ECOG		AOR (95% CI)	P-value
	ECOG 0+1	ECOG 2+3+4		
***ABCB1* rs1128503 (Dominant model TT)**	N= 89 (n%)	N= 87 (n%)		
TT	35 (39.32)	38 (43.67)	Reference	-
TC	40 (44.94)	38 (43.67)	0.88 (0.46 to 1.71)	0.72
CC	14 (15.73)	11 (12.64)	0.70 (0.26 to 1.91)	0.49
TC+CC	54 (60.67)	49 (56.32)	0.87 (0.47 to 1.63)	0.68
***ABCB1* rs1128503 (Recessive model CC)**	N= 89 (n%)	N= 87 (n%)		
CC	14 (15.73)	11 (12.64)	Reference	-
CT	40 (44.94)	38 (43.67)	0.99 (0.37 to 2.65)	0.99
TT	35 (39.32)	38 (43.67)	1.41 (0.52 to 3.83)	0.49
CT+TT	75 (84.26)	76 (87.35)	1.12 (0.45 to 2.75)	0.79
***ABCB1* rs2032582**	N= 65 (n%)	N= 76 (n%)		
GG	3 (4.61)	1 (1.31)	Reference	-
GT	40 (61.53)	46 (60.52)	2.57 (0.21 to 30.38)	0.45
TT	22 (33.84)	29 (38.15)	2.00 (0.16 to 24.84)	0.58
GT+TT	62 (95.38)	75 (98.68)	2.39 (0.22 to 25.96)	0.47

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 GG genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for ECOG. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better than control arm provided significant P-value

Table 5.15 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to performance status based on ECOG score of the patients with genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated included patients ECOG value (0-1) in one group who are considered as individuals with good performance status and ECOG value (2, 3 and 4) who are patients with declined performance status.

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in ECOG (0-1) and ECOG (2-3-4) consisted of 39.32% and 43.67% of homozygous wild genotype respectively. 44.94% and 43.67% of patients had heterozygous genotype in ECOG (0-1) and ECOG (2-3-4) group respectively (OR = 0.88, P = 0.72). 15.73% and 12.64% individual were carriers of mutant genotype in ECOG (0-1) and ECOG (2-3-4) groups respectively (OR = 0.7, p = 0.49). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of ECOG value cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in ECOG (0-1) and ECOG (2-3-4) groups consisted of 4.61% and 1.31% of homozygous wild genotype respectively. 61.53% and 60.52% of patients had heterozygous genotype in ECOG (0-1) and ECOG (2-3-4) group respectively (OR = 2.57, P = 0.45). 33.85% and 38.15% individual were carriers of mutant genotype in ECOG (0-1) and ECOG (2-3-4) groups respectively (OR = 2.003, p = 0.58). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of ECOG value cannot be proved.

5.12 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with overall survival of lung cancer patients on the basis of different regimen

Table 5.17 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of regimen (Dominant model TT in North Indian population)

Docetaxel-plus-Carboplatin/Cisplatin (regimen 1)							
***ABCB1* rs1128503 genotype**	**Alive 4 (n%)**	**Dead 26 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	3 (18.75)	13 (81.25)	5.76	1.00 (Reference)	-	1.00 (Reference)	-
TC	0	9	10.26	0.91 (0.39 to 2.13)	0.83	1.18 (0.38 to 3.65)	0.77
CC	1 (20)	4 (80)	7.56	0.84 (0.28 to 2.48)	0.77	0.41 (0.11 to 1.46)	0.17
TC+CC	1 (7.14)	13 (92.85)	8.20	0.89 (0.41 to 1.93)	0.77	0.66 (0.27 to 1.61)	0.37
Irinotecan-plus-Carboplatin/Cisplatin (regimen 5)							
***ABCB1* rs1128503 genotype**	**Alive 6 (n%)**	**Dead 31 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log rank P**	**Adjusted HR (95% CI)**	**P**
TT	3 (20)	12 (80)	13.66	1.00 (Reference)	-	1.00 (Reference)	-
TC	2 (11.76)	15 (88.23)	7.33	1.51 (0.70 to 3.22)	0.27	1.85 (0.64 to 5.33)	0.25
CC	1 (20)	4 (80)	25.43	0.53 (0.19 to 1.48)	0.26	0.46 (0.08 to 2.44)	0.36
TC+CC	3 (13.63)	19 (86.36)	8.26	1.16 (0.56 to 2.36)	0.68	1.51 (0.56 to 4.11)	0.41
Pemetrexed-plus-Carboplatin/Cisplatin (regimen 6)							
***ABCB1* rs1128503 genotype**	**Alive 4 (n%)**	**Dead 25 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
TT	3 (21.42)	11 (78.57)	5.06	1.00 (Reference)	-	1.00 (Reference)	-
TC	1 (10)	9 (90)	9.56	0.82 (0.34 to 1.98)	0.65	0.39 (0.09 to 1.53)	0.18
CC	0	5	9.23	1.67 (0.50 to 5.53)	0.30	35.37 (2.35 to 530.90)	**0.01**
TC+CC	1 (6.66)	14 (93.33)	9.33	1.03 (0.47 to 2.27)	0.93	0.89 (0.30 to 2.61)	0.83

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR,

95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

In dominant model for TT of *ABCB1* rs1128503 polymorphism in North Indian population, among 30 patients treated with Docetaxel-Carboplatin/cisplatin, 86.66% of the cases were dead and 13.33% were alive as summarized in Table 5.17. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type TT genotype, 81.25% were dead and 18.75% were alive (MST = 5.76 months). Overall survival was highest in individuals carrying heterozygous genotype TC (MST = 10.26, HR = 0.91, Log rank P = 0.83) and mutant genotype CC (MST = 7.56, HR = 0.84, Log rank P = 0.77). On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild genotype we found that OS period was much better than individuals with wild genotype (MST = 8.2 months, HR = 0.89, Log rank P = 0.77). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis thus denying the association between OS of lung cancer patients on the basis of regimen with genotypic distribution.

Among 37 patients treated with Irinotecan-Carboplatin/cisplatin, 83.78% of the cases were dead and 16.21% were alive as summarized in Table 5.17. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type TT genotype, 80% were dead and 20% were alive (MST = 13.66 months). Overall survival was highest in individuals carrying mutant genotype CC (MST = 25.43, HR = 0.53, Log rank P = 0.26) and lowest in individuals carrying heterozygous genotype TC (MST = 7.33, HR = 1.51, Log rank P = 0.27). On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild genotype we found that OS period decreased (MST = 8.26 months, HR = 1.16, Log rank P = 0.68). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis thus denying the association between OS of lung cancer patients on the basis of regimen with genotypic distribution.

Among 29 patients treated with Pemetrexed-Carboplatin/cisplatin, 86.20% of the cases were dead and 13.79% were alive as summarized in Table 5.17. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type TT genotype, 78.57% were dead and 21.42% were alive (MST = 5.067 months). Overall survival was highest in individuals carrying heterozygous genotype TC (MST = 9.56, HR = 0.82, Log rank P = 0.65). On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild

genotype we found that OS period was much better than individuals with wild genotype (MST = 9.33 months, HR = 1.03, Log rank P = 0.93). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis. Individuals carrying mutant genotype CC showcased statistically significant increased survival period then subjects with wild genotype (MST = 9.23 months, HR = 1.67, **P = 0.01**).

Table 5.18 Genotypic distribution and Relationship of *ABCB1* rs1128503 polymorphism with overall survival of lung cancer patients on the basis of regimen (Recessive model CC in North Indian population)

Docetaxel-plus-Carboplatin/Cisplatin (regimen 1)							
***ABCB1* rs1128503 genotype**	**Alive 4 (n%)**	**Dead 26 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	2 (33.33)	4 (66.66)	7.56	1.00 (Reference)	-	1.00 (Reference)	-
CT	0	9	10.26	0.90 (0.27 to 3.02)	0.87	2.00 (0.47 to 8.57)	0.34
TT	2 (13.33)	13 (86.66)	5.76	1.17 (0.40 to 3.45)	0.77	2.41 (0.68 to 8.55)	0.17
CT+TT	2 (8.33)	22 (91.66)	8.20	1.08 (0.38 to 3.05)	0.88	2.41 (0.74 to 7.85)	0.14
Irinotecan-plus-carboplatin/cisplatin (regimen 5)							
***ABCB1* rs1128503 genotype**	**Alive 6 (n%)**	**Dead 31 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	1 (20)	4 (80)	25.43	1.00 (Reference)	-	1.00 (Reference)	-
CT	2 (11.76)	15 (88.23)	7.33	2.13 (0.84 to 5.44)	0.15	4.71 (0.85 to 25.84)	0.07
TT	3 (20)	12 (80)	13.66	1.85 (0.67 to 5.07)	0.26	2.14 (0.40 to 11.21)	0.36
CT+TT	5 (15.62)	27 (84.38)	9.00	1.96 (0.84 to 4.55)	0.19	1.79 (0.52 to 6.09)	0.35
Pemetrexed-plus-Carboplatin/Cisplatin (regimen 6)							
***ABCB1* rs1128503 genotype**	**Alive 4 (n%)**	**Dead 25 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
CC	0	5	9.23	1.00 (Reference)	-	1.00 (Reference)	-
CT	1 (10)	9 (90)	9.56	0.49 (0.13 to 1.76)	0.17	0.43 (0.07 to 2.38)	0.33
TT	3	11	5.06	0.59	0.30	0.02	**0.01**

	(21.42)	(78.57)		(0.18 to 1.97)		(0.001 to 0.42)	
CT+TT	4 (30.28)	20 (83.33)	9.46	0.54 (0.16 to 1.83)	0.20	0.33 (0.09 to 1.21)	0.09
Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.							

In recessive model for CCof *ABCB1* rs1128503 polymorphism in North Indian population, among 30 patients treated with Docetaxel-Carboplatin/cisplatin, 86.66% of the cases were dead and 13.33% were alive as summarized in Table 5.18. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type CC genotype, 66.66% were dead and 33.33% were alive (MST = 7.56 months). Overall survival was highest in individuals carrying heterozygous genotype CT (MST = 10.26, HR = 0.90, Log rank P = 0.87) and lowest in individuals with mutant genotype TT (MST = 5.76 months, HR = 1.17, Log rank P = 0.77). On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild genotype we found that OS period was much better than individuals with wild genotype (MST = 8.2 months, HR = 1.08, Log rank P = 0.88). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis thus denying the association between OS of lung cancer patients on the basis of regimen 1 with genotypic distribution.

Among 37 patients treated with Irinotecan-Carboplatin/cisplatin, 83.78% of the cases were dead and 16.21% were alive as summarized in Table 5.18. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type CC genotype, 80% were dead and 20% were alive (MST = 25.43 months). A decline in Overall survival was observed in individuals carrying mutant genotype TT (MST = 13.66, HR = 1.85, Log rank P = 0.26) and lowest in individuals carrying heterozygous genotype TC (MST = 7.33, HR = 2.13, Log rank P = 0.15) with increased risk of death rate. On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild genotype we found that OS period decreased (MST = 9 months, HR = 1.96, Log rank P = 0.19). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis thus denying the association between OS of lung cancer patients on the basis of regimen with genotypic distribution.

Among 29 patients treated with Pemetrexed-Carboplatin/cisplatin, 86.20% of the cases were dead and 13.79% were alive as summarized in Table 5.18. Based on genotypic distribution, individuals carrying homozygous wild type TT genotype showcased OS of 9.23 months. Overall survival was highest in individuals carrying heterozygous genotype TC (MST = 9.56 months, HR = 0.49, Log rank P = 0.17). On combining heterozygous and mutant genotype into a single genotype and comparing it against the wild genotype we found that OS period was similar to individuals with wild genotype (MST = 9.46 months, HR = 0.54, Log rank P = 0.20). However, no statistically significant values were obtained with multivariate Cox proportional hazards regression analysis. Individuals carrying mutant genotype CC showcased statistically significant dip in their survival period marking an increased risk of death rate (MST = 5.067 months, HR = 0.59, Log rank P = 0.30, Adjusted HR = 0.02, **P = 0.01)**.

Table 5.19 Genotypic distribution and Relationship of *ABCB1* rs2032582 polymorphism with overall survival of lung cancer patients on the basis of regimen

Docetaxel-plus-Carboplatin/Cisplatin (regimen 1)							
***ABCB1* rs2032582 genotype**	**Alive 7 (n%)**	**Dead 30 (n%)**	**Median OS (months)**	**HR (95% CI)**	**Log P**	**Adjusted HR (95% CI)**	**P**
GG	1 (50)	1 (50)	1.36	1.00 (Reference)	-	1.00 (Reference)	-
GT	3 (14.28)	18 (85.71)	9.66	1.64 (0.32 to 8.34)	0.61	0.07 (0.003 to 1.70)	0.10
TT	3 (21.42)	11 (78.57)	11.60	1.21 (0.18 to 8.04)	0.85	0.000 3.0768E-198 to 1.9571E+186	0.95
GT+TT	6 (17.14)	29 (82.85)	9.86	1.50 (0.28 to 7.85)	0.68	0.06 (0.003 to 1.09)	**0.05**

Hazard ratios (HR), 95% Confidence intervals (CI) and their respective p-values were estimated using Kaplan-meier survival analysis after adjusting for remission and overall survival (months). Adjusted HR, 95% CI and their respective P-values were determined after adjusting for age, ECOG, gender, KPS, smoking status and stage.

Among 37 patients treated with Docetaxel-Carboplatin/cisplatin, 81.08% of the cases were dead and 18.91% were alive as summarized in Table 5.19. Based on genotypic distribution, we observed that among individuals carrying homozygous wild type GG genotype, OS period of as low as 1.36 months was observed. No significant values were observed by performing Cox proportional hazards regression analysis confirming no association between OS of lung cancer

patients on the basis of regimen with genotypic distribution, except, for joint combination of genotypes (GT+TT). It was observed that individuals carrying heterozygous-mutant genotype combination showcased a significantly increased overall survival of 9.867 months (Adjusted HR = 0.06, **P = 0.05**).

The above tests could not be performed for regimen 5 (Irino-Carbo/Cis) and regimen 6 (Pemetrexed-Carbo/Cis) due to absence of homozygous wild genotype (genotypic distribution not complying with HWE).

5.13 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with lung cancer susceptibility on the basis of chemotherapy response

Table 5.20 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to response to chemotherapy

Genotype	Response		AOR (95% CI)	P-value
	CR/PR	SD/PD		
***ABCB1* rs1128503 (Dominant model TT)**	N= 51 (n%)	N= 48 (n%)		
TT	26 (50.98)	21 (43.75)	Reference	-
TC	17 (33.33)	17 (35.41)	1.35 (0.52 to 3.51)	0.52
CC	8 (15.68)	10 (20.83)	2.02 (0.61 to 6.73)	0.24
TC+CC	25 (49.01)	27 (56.25)	1.45 (0.62 to 3.37)	0.37
***ABCB1* rs1128503 (Recessive model CC)**	N= 51 (n%)	N= 48 (n%)		
CC	8 (15.68)	10 (20.83)	Reference	-
CT	17 (33.33)	17 (35.41)	0.93 (0.26 to 3.34)	0.91
TT	26 (50.98)	21 (43.75)	0.49 (0.14 to 1.63)	0.24
CT+TT	43 (84.31)	38 (79.16)	0.62 (0.20 to 1.86)	0.39
***ABCB1* rs2032582**	N= 43 (n%)	N= 39 (n%)		

GG	0 (0)	1 (2.56)	Reference	-
GT	30 (69.76)	16 (41.02)	1.00	0.73
TT	13 (30.23)	22 (56.41)	0.00	0.99
GT+TT	43 (100)	38 (97.43)	0.00	0.99

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 *GG* genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for chemotherapy response. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better than control arm provided significant P-value.

Chemotherapy response by patients is stratified as complete response (CR), partial response (PR), stable disease (SD) and progressive disease (PD). CR and PR is considered in the category of good responders whereas SD/PD are categorized as bad responders.

Table 5.20 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to chemotherapy response of patients to the treatment based on genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated included patients CR/PR in one group and SD/PD responding patients in another group.

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in CR/PR and SD/PD group consisted of 50.98% and 43.75% of homozygous wild genotype respectively. 33.33% and 35.41% of patients had heterozygous genotype in CR/PR and SD/PD group respectively (OR = 1.35, P = 0.52). 15.68% and 20.83% individual were carriers of mutant genotype in CR/PR and SD/PD groups respectively (OR = 2.02, p = 0.24). No statistically significant values were obtained by logistic regression analysis again; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of chemotherapy response cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in CR/PR and SD/PD group consisted of 0% and 2.56% of homozygous wild genotype respectively. 69.76% and 41.02% of patients had heterozygous genotype in CR/PR and SD/PD group respectively (OR = 1, P = 0.73). 33.85% and 38.15% individual were carriers of mutant genotype in CR/PR and SD/PD groups respectively (OR = 0.00, p = 0.99). Nevertheless, no significant statistical values were obtained

by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of ECOG value cannot be proved.

5.14 Association of *ABCB1* rs1128503 and rs2032582 polymorphisms genotypic distribution with clinicopathological parameters

Table 5.21 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to type of Tumor

Genotype	Tumor size and invasion		AOR (95% CI)	P-value
	T0+T1+T2	T3+T4		
***ABCB1* rs1128503 (Dominant model TT)**	N= 33 (n%)	N= 134 (n%)		
TT	11 (33.33)	57 (42.43)	Reference	-
TC	16 (48.48)	57 (42.53)	0.68 (0.28 to 1.66)	0.40
CC	6 (18.18)	20 (14.92)	0.67 (0.19 to 2.33)	0.53
TC+CC	22 (66.66)	77 (57.46)	0.66 (0.29 to 1.51)	0.33
***ABCB1* rs1128503 (Recessive model CC)**	N= 33 (n%)	N= 134 (n%)		
CC	6 (18.18)	20 (14.92)	Reference	-
CT	16 (48.48)	57 (42.53)	1.07 (0.34 to 3.35)	0.89
TT	11 (33.33)	57 (42.43)	1.47 (0.42 to 5.08)	0.54
CT+TT	27 (81.81)	114 (85.07)	1.16 (0.40 to 3.35)	0.78
***ABCB1* rs2032582**	N= 28 (n%)	N= 99 (n%)		
GG	1 (3.57)	1 (1.01)	Reference	-
GT	17 (60.71)	62 (62.62)	5.16 (0.25 to 125.18)	0.27
TT	10 (35.71)	36 (36.36)	2.34 (0.11 to 46.60)	0.57
GT+TT	27 (96.42)	98 (98.98)	3.34 (0.18 to 61.97)	0.41

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1*

rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 *GG* genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for tumor size extent of invasion. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better then control arm provided significant P-value.

Table 5.21 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to tumor size and invasion extent based on genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated according to genotypic distribution included patients with Tumor T0+T1+T2 in one group giving zero or minimal tumor growth and invasion and Tumor T3+T4 in another group showcasing extensive tumor size growth and invasion.

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in T0+T1+T2 and T3+T4 group consisted of 33.33% and 42.43% of homozygous wild genotype respectively. 48.48% and 42.53% of patients had heterozygous genotype in T0+T1+T2 and T3+T4 group respectively (OR = 0.68, P = 0.40). 18.18% and 14.92% individual were carriers of mutant genotype in T0+T1+T2 and T3+T4 groups respectively (OR = 0.678, p = 0.53). No statistically significant values were obtained by logistic regression analysis again; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of tumor size and extent of invasion cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in T0+T1+T2 and T3+T4 group consisted of 3.57% and 1.01% of homozygous wild genotype respectively. 60.71% and 62.62% of patients had heterozygous genotype in T0+T1+T2 and T3+T4 group respectively (OR = 5.16, P = 0.27). 35.71% and 36.36% individual were carriers of mutant genotype in T0+T1+T2 and T3+T4 groups respectively (OR = 2.34, p = 0.57). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of tumor size and extent of invasion cannot be proved.

Table 5.22 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to Lymph node involvement

Genotype	Lymph node		AOR (95% CI)	P-value
	N0+N1	N2+N3		
***ABCB1* rs1128503 (Dominant model TT)**	N= 35 (n%)	N= 132 (n%)		
TT	15 (42.85)	53 (40.15)	Reference	-
TC	18 (51.42)	55 (41.66)	0.84 (0.37 to 1.90)	0.69
CC	2 (5.71)	24 (18.18)	4.49 (0.81 to 24.94)	0.08
TC+CC	20 (57.14)	79 (59.84)	1.05 (0.48 to 2.30)	0.88
***ABCB1* rs1128503 (Recessive model CC)**	N= 35 (n%)	N= 132 (n%)		
CC	2 (5.71)	24 (18.18)	Reference	-
CT	18 (51.42)	55 (41.66)	0.28 (0.05 to 1.37)	0.11
TT	15 (42.85)	53 (40.15)	0.22 (0.04 to 1.23)	0.08
CT+TT	33 (94.28)	108 (81.81)	0.27 (0.05 to 1.25)	0.09
***ABCB1* rs2032582**	N= 21 (n%)	N= 106 (n%)		
GG	1 (4.76)	1 (0.94)	Reference	-
GT	10 (47.61)	69 (65.09)	17.88 (0.49 to 649.54)	0.11
TT	10 (47.61)	36 (33.96)	4.64 (0.22 to 94.57)	0.31
GT+TT	20 (95.23)	105 (99.05)	6.98 (0.36 to 134.21)	0.19

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 *GG* genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for lymph node involvement. . OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better then control arm provided significant P-value.

Table 5.22 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on stratification of groups according to lymph node involvement based on genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two group to be evaluated included patients Node N0+N1 in one group showcasing absence or minimal lymph node metastasis and Node N2+N3 showcasing distant lymph node metastasis in another group .

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in N0+N1 and N2+N3 group consisted of 42.85% and 40.15% of homozygous wild genotype respectively. 51.42% and 41.66% of patients had heterozygous genotype in N0+N1 and N2+N3 group respectively (OR = 0.84, P = 0.69). 5.71% and 18.18% individual were carriers of mutant genotype in N0+N1 and N2+N3 groups respectively (OR = 4.49, p = 0.08). Nevertheless, no statistically significant values were obtained by logistic regression analysis again; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of lymph node metastasis cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in N0+N1 and N2+N3 group consisted of 4.76% and 0.94% of homozygous wild genotype respectively. 47.61% and 65.09% of patients had heterozygous genotype in N0+N1 and N2+N3 group respectively (OR = 17.8, P = 0.11). 47.61% and 33.96% individual were carriers of mutant genotype in N0+N1 and N2+N3 groups respectively (OR = 4.64, p = 0.31). However, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of lymph node metastasis cannot be proved.

Table 5.23 Genotypic distribution of *ABCB1* rs1128503 and *ABCB1* rs2032582 polymorphisms and risk associated with overall survival of lung cancer according to Metastasis

Genotype	Metastasis		AOR (95% CI)	P-value
	M0	M1		
***ABCB1* rs1128503 (Dominant model TT)**	N= 95 (n%)	N= 72 (n%)		
TT	39 (41.05)	29 (40.27)	Reference	-
TC	42 (44.21)	31 (43.05)	1.20 (0.59 to 2.47)	0.60
CC	14 (14.73)	12 (16.66)	1.04 (0.35 to 3.05)	0.94
TC+CC	56 (58.94)	43 (59.72)	1.18 (0.59 to 2.33)	0.63
***ABCB1* rs1128503 (Recessive model CC)**	N= 95 (n%)	N= 72 (n%)		
CC	14 (14.73)	12 (16.66)	Reference	-
CT	42 (44.21)	31 (43.05)	1.40 (0.48 to 4.08)	0.53
TT	39 (41.05)	29 (40.27)	0.96 (0.32 to 2.81)	0.94
CT+TT	81 (85.26)	60 (83.33)	1.14 (0.43 to 2.99)	0.78
***ABCB1* rs2032582**	N= 70 (n%)	N= 57 (n%)		
GG	1 (1.42)	1 (1.75)	Reference	-
GT	43 (61.42)	36 (63.15)	0.94 (0.04 to 20.34)	0.96
TT	26 (37.14)	20 (35.08)	0.78 (0.03 to 20.02)	0.88
GT+TT	69 (98.57)	56 (98.24)	1.01 (0.04 to 22.03)	0.99

Calculated Adjusted Odds ratio(AOR) and 95% CI by logistic regression analysis with *ABCB1* rs1128503 TT genotype (dominant model) and CC genotype (recessive model), *ABCB1* rs2032582 *GG* genotype respectively, as reference group after adjusting for age, gender, histology and smoking status for metastasis. OR > 1 implies that control arm is better than cases and OR < 1 implies that case arm is better then control arm provided significant P-value.

Table 5.23 demonstrates the statistical values obtained by logistic regression analysis to evaluate the risk associated with overall survival of lung cancer patients on classification of groups on the

basis of extent of metastasis based on genotype distribution for polymorphisms *ABCB1* rs1128503 (dominant model TT and recessive model CC) and *ABCB1* rs2032582. The two groups to be evaluated included individuals with Metastasis M0 in one group showcasing no metastasis and Metastasis M1 patients in another group showing distant metastasis.

In dominant model for TT of *ABCB1* rs1128503 polymorphism, individuals in Metastasis M0 and M1 group consisted of 41.05% and 40.27% of homozygous wild genotype respectively. 44.21% and 43.05% of patients had heterozygous genotype in M0 and M1 group respectively (OR = 1.20, P = 0.60). 14.73% and 16.66% individual were carriers of mutant genotype in CR/PR and SD/PD groups respectively (OR = 1.041, p = 0.94). No statistically significant values were obtained by logistic regression analysis again; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of extent of metastasis cannot be proved.

In *ABCB1* rs2032582 polymorphism OS study, individuals in M0 and M1 group consisted of 1.42% and 1.75% of homozygous wild genotype respectively. 61.42% and 63.15% of patients had heterozygous genotype in M0 and M1 group respectively (OR = 0.94, P = 0.96). 37.14% and 35.08% individual were carriers of mutant genotype in M0 and M1 groups respectively (OR = 0.78, p = 0.88). Nevertheless, no significant statistical values were obtained by logistic regression analysis; hence association between various genotypic combinations with associated risk of OS in lung cancer patients on the basis of extent of metastasis cannot be proved.

5.16 RESULT OF COMPUTATIONAL ANALYSIS OF EFFECT OF SINGLE NUCLEOTIDE POLYMORPHISM IN ABC TRANSPORTERS

Table 5.24 contains the SNPs of various ABC gene subfamily selected for computational analysis.

Table 5.24 Details of rs IDs submitted to SIFT algorithm

rs ID	SNP	ABC transporter type
rs2032582	G2677T	ABCB1
rs2032582	G2677A	ABCB1
rs41395947	G128C	ABCC1
rs41494447	C218T	ABCC1
rs4148356	G2168A	ABCC1
rs41410450	G3173A	ABCC1
rs8187710	G4544A	ABCC2
rs2231142	C421A	ABCG2

The impact of the given nsSNPs on structure and function of respective protein was analyzed using various softawares such as GENEMANIA, SIFT, Polyphen-2, MutPred2, PMUT, PROVEAN, PhD-SNP, SNPs & GO, Imutant3, MUpro, PANTHER and ELASPIC.

5.16.1 GENEMANIA

Figure 5.25 depicts the functional association between various ABC transporter subfamily gene and its other related gene. Genemania demonstrated many important functions of ABC transporters such as it is a primary active transmembrane transporter; it has ATPase activity, hydrolase activity, present on apical part of cell, involved interaction with drugs thus affecting drug response rate, in response to stress, it regulates transcription from RNA polymeras II promoters, positively regulates glycolysis etc.

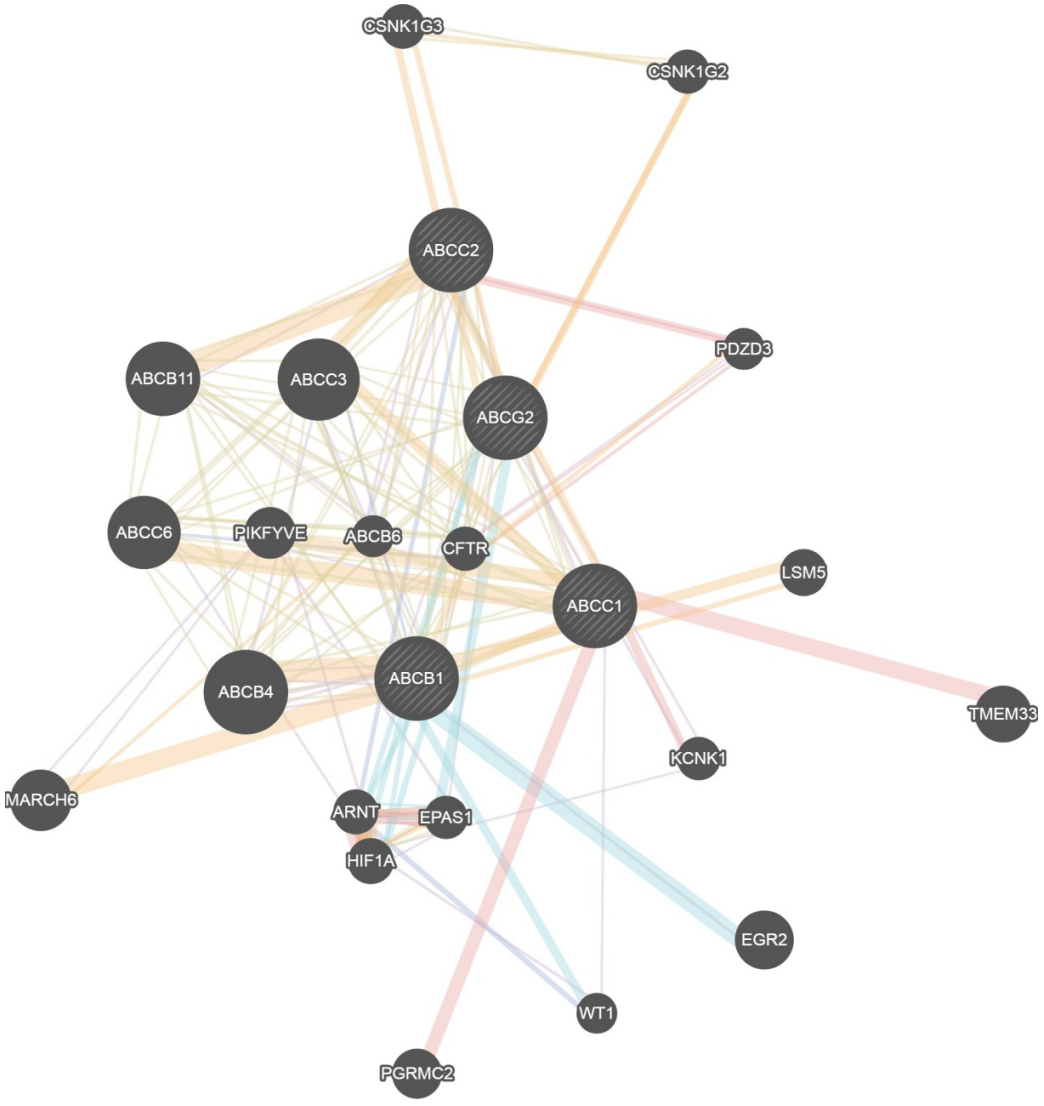


Figure 5.25 Functional relation between different ABC transporter gene and its related genes

5.16.2 Result prediction of SIFT and Polyphen-2 algorithm

Table 5.25 Showing SNPs predicted by SIFT and Polyphen-2

SNP rsID	Protein ID	Nucleotide change	Amino acid change		SIFT score	SIFT prediction	Polyphen-2 score	Polyphen-2 prediction
rs2032582	NP_000918	G/T	S893A	S A	0.51 1	Tolerated Tolerated	0.002	Benign
rs2032582	NP_000918	G/A	S893T	S T	0.51 0.07	Tolerated Tolerated	0.07	Benign
rs41395947	NP_004987	G/C	C43S	C S	1 0.1	Tolerated Tolerated	0.997	**Probably damaging**
rs41494447	NP_004987	C/T	T73I	T I	0.45 0.15	Tolerated Tolerated	0.075	Benign
rs4148356	NP_004987	G/A	R723Q	R Q	1 0.3	Tolerated Tolerated	0.013	Benign
rs41410450	NP_004987	G/A	R1058Q	R Q	1 0.01	Tolerated **Damaging**	0.917	**Probably damaging**
rs8187710	NP_000383	G/A	C1515Y	C Y	0.04 0.22	**Damaging** Tolerated	0	Benign
rs2231142	NP_004818	C/A	Q141K	Q K	0.36 0.34	Tolerated Tolerated	0.216	Benign

Table 5.25 shows the predicted results of SIFT and Polyphen-2 algorithms. SIFT result was obtained by submitting batch rsIDs of the SNP to be studied. Out of the 8 SNPs submitted to SIFT server, 2 SNPs with a SIFT score below 0.05 were forecasted as damaging (rs41410450 and rs8187710). Amino acid substitution in other SNPs were predicted as tolerated by SIFT algorithm (SIFT score > 0.05). Polyphen-2 predictions were obtained by submitting FASTA sequences of proteins obtained from UniProt database (UniProt IDs: P08183, P33527, Q92887 and Q9UNQ0). rs41395947 and rs41410450 SNP were determined as probably damaging by obtaining high scores (greater the score, greater is the damaging effect).

5.16.3 Using MutPred2 software to predict harmful non-synonymous SNPs

Table 5.26 Functional effect of nsSNPs on respective ABC transporters as predicted by MutPred2

rsID	Amino acid substitution	MutPred2 probability score	MutPred2 prediction
rs2032582	S893A	0.058	Neutral
rs2032582	S893T	0.508	**Harmful mutation**
rs41395947	C43S	0.8	**High confidence**
rs41494447	T73I	0.51	**Harmful mutation**
rs4148356	R723Q	0.159	Neutral
rs41410450	R1058Q	0.428	Neutral
rs8187710	C1515Y	0.087	Neutral
rs2231142	Q141K	0.397	Neutral

MutPred2 score helps in determination of whether an amino acid changes is neutral or associated with a disease. MutPred2 predictions are made as follows: MutPred2 score > 0.5 (harmful such as rs2032582 and rs41494447); MutPred2 score > 0.75 (High confidence harmful such as rs41395947); and MutPred2 score < 0.5 (neutral). (Table 5.26)

5.16.4 Recognition of Disease-associated nsSNPs using softwares PMUT, PROVEAN, PhD-SNP and SNPs & GO

Table 5.27 showing result predicted by PMUT, PROVEAN, PhD-SNP and SNPs & GO software

SNP rsID	Amino acid substitution	PMUT	PROVEAN		PhD-SNP	SNPs & GO
			Score	Prediction (cut off= -2.5)		
rs2032582	S893A	Neutral	1.665	Neutral	Neutral	Neutral
rs2032582	S893T	Neutral	-1.1	Neutral	**Disease**	Neutral
rs41395947	C43S	**Disease**	-6.226	**Deleterious**	**Disease**	Neutral
rs41494447	T73I	Neutral	-0.523	Neutral	Neutral	Neutral
rs4148356	R723Q	Neutral	-0.36	Neutral	**Disease**	Neutral
rs41410450	R1058Q	**Disease**	-2.846	**Deleterious**	Neutral	Neutral
rs8187710	C1515Y	Neutral	4.086	Neutral	Neutral	Neutral
rs2231142	Q141K	Neutral	-1.588	Neutral	Neutral	Neutral

PMUT software is used to analyze the functional effect nsSNP on proteins. Similarly, PROVEAN helps to predict the impact of mutation on biological activity of the protein. The variant is categorized as disease-associated if the PROVEAN score of below threshold value (cut off = -2.5). According to PMUT and PROVEAN server, out of 8 SNPs submitted, 2 (rs41395947 and rs41410450) were predicted as disease associated while the rest were labeled as neutral.

Contradictory results were obtained with PhD-SNP and SNPs & GO softwares. SNPs & GO software predicted all amino acid substitutions submitted in the server as neutral. In PhD-SNP software, 3 SNPs were forecasted as deleterious (rs2032582 S→T, rs41395947 C→S and rs4148356 R→Q).

5.16.5 Impact of nsSNPs on stability of protein as predicted by Imutant3 and MUpro

Table 5.28 showing results predicted by I-mutant3 and MUpro showcasing impact of non-synonymous SNP on stability of protein

Amino acid substitution	DDG (Kcal/mol)	I-mutant3 prediction		DDG value (Kcal/mol)	MUpro prediction
		SVM2	SVM3		
S893A	-0.17	Decrease	Neutral	-1.2707	Decrease stability
S893T	0.22	Increase	Neutral	-1.2354	Decrease stability
C43S	-0.74	Decrease	Large decrease	-0.7523	Decrease stability
T73I	-0.24	Decrease	Large decrease	-1.0115	Decrease stability
R723Q	-0.71	Decrease	Large decrease	-0.9465	Decrease stability
R1058Q	-1.14	Decrease	Large decrease	-0.8016	Decrease stability
C1515Y	-0.3	Decrease	Large decrease	-0.736	Decrease stability
Q141K	-0.46	Decrease	Large decrease	-1.4482	Decrease stability

According to the nsSNPs predicted by I-Mutant3, alteration in protein stability were observed as a consequence of single point mutation in 6 SNPS (C→S, T→I, R→Q, R→Q, C→Y and Q→K). Also, all the nsSNPs submitted were associated with decreased stability of proteins as predicted by MUpro (Table 5.28).

5.16.6 Result prediction of PANTHER

Table 5.29 showing prediction results of PANTHER			
SNP rsID	**Amino acid substitution**	**Preservation time (million years)**	**PANTHER prediction**
rs2032582	S893A	6	Probably benign
rs2032582	S893T	6	Probably benign
rs41395947	C43S	797	**Probably damaging**
rs41494447	T73I	30	Probably benign
rs4148356	R723Q	221	**Probably damaging**
rs41410450	R1058Q	797	**Probably damaging**
rs8187710	C1515Y	30	Probably benign
rs2231142	Q141K	-	-

Protein FASTA sequence along with amino acid substitution position were submitted as input in PANTHER to estimate the impact of deleterious nsSNPs on function of protein. PSEP (position-specific evolutionary preservation) score is the time in millions of years a particular position has been preserved during evolution. The longer a position is preserved; greater is the probability of substitution being deleterious. Out of the 8 nsSNPs, 3 were predicted as probably damaging whereas 4 were labeled as probably benign. (Table 5.29)

5.16.7 Prediction of results by ELASPIC

Table 5.30 showing protein stability results predicted by ELASPIC on domain cores and domain-domain interface:						
SNP rsID	**Amino acid change**	**UniProt ID**	**Protein type**	**ΔG_{wt}**	**ΔG_{mut}**	**$\Delta\Delta G$**
rs2032582	S893A	P08183	Core	253.017	251.566	-2.179
rs2032582	S893T	P08183	Core	253.017	254.118	-0.936
rs41395947	C43S	P33527	-	-	-	-
rs41494447	T73I	P33527	-	-	-	-
rs4148356	R723Q	P33527	Core	28.5622	29.1107	0.231
rs41410450	R1058Q	P33527	Core	956.947	956.818	0.17
rs8187710	C1515Y	Q92887	Core	189.961	192.112	0.361
rs2231142	Q141K	Q9UNQ0	Core	336.203	344.078	0.767

ELASPIC software forecasted 6 out of 8 mutations in the core of ABC family proteins while 2 mutations were not available on ELASPIC (Table 5.29). Figure 5.25 to Figure 5.30 below depicts changes in core protein having wild-type residue and new residue.

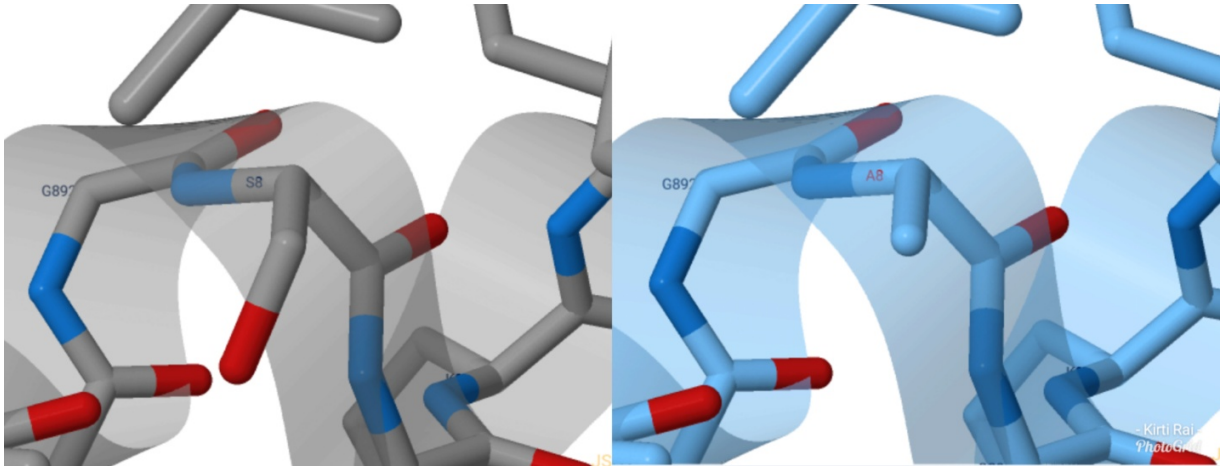


Figure 5.26 ELASPIC result of rs2032582 (S893A): wild residue (left) and mutant residue (right)

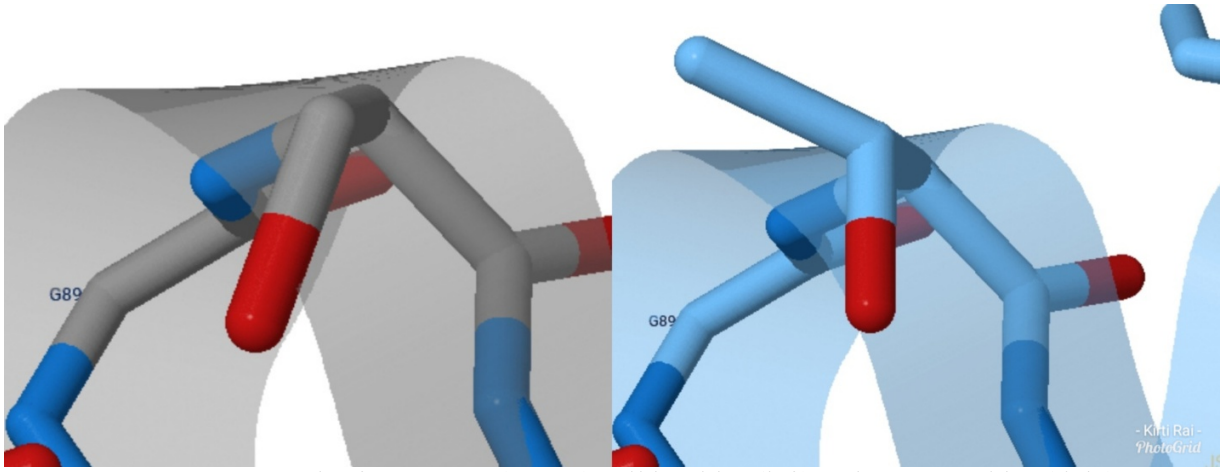


Figure 5.27 ELASPIC result of rs2032582 (S893T): wild residue (left) and mutant residue (right)

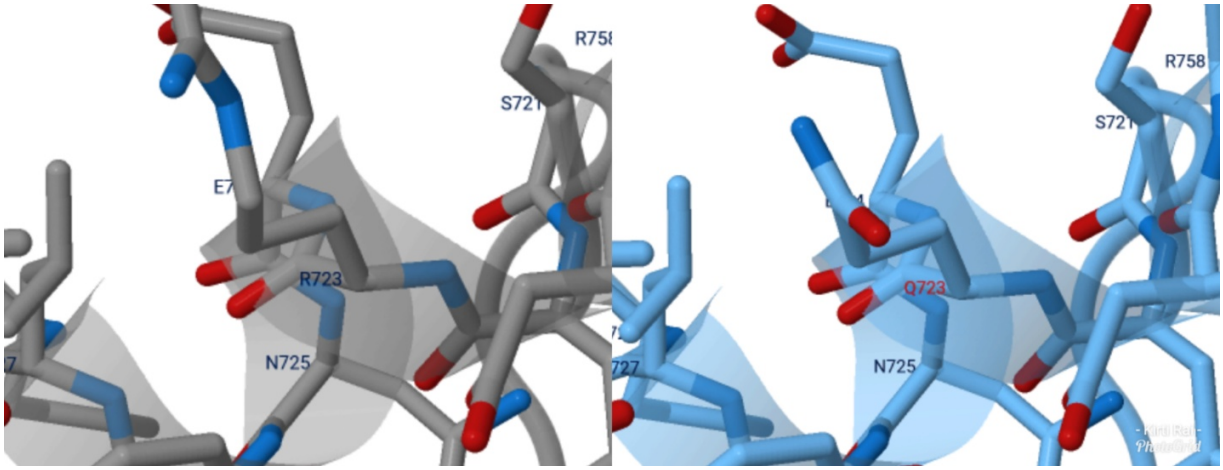


Figure 5.28 ELASPIC result of rs4148356 (R723Q): wild residue (left) and mutant residue (right)

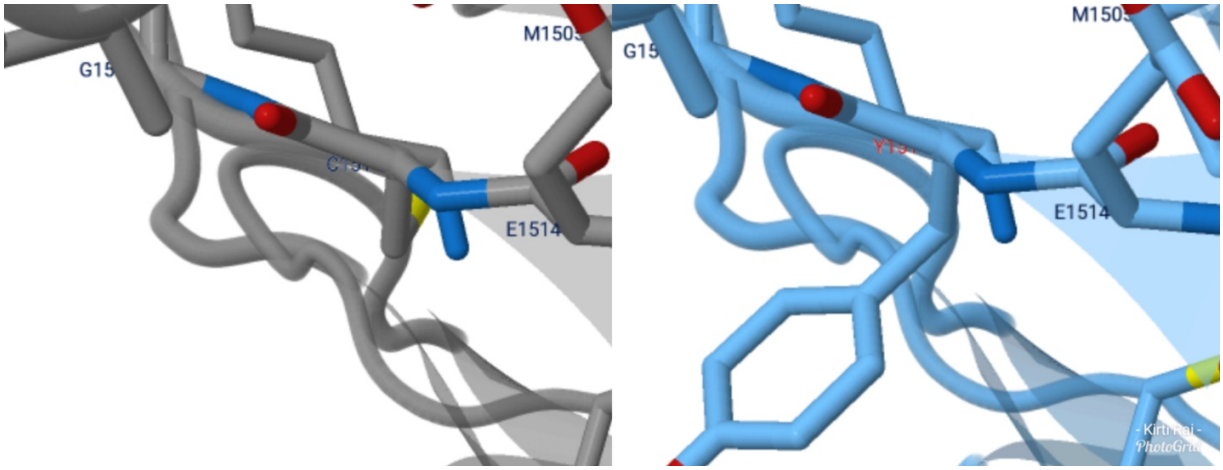


Figure 5.29 ELASPIC result of rs8187710 (C1515Y): wild residue (left) and mutant residue (right)

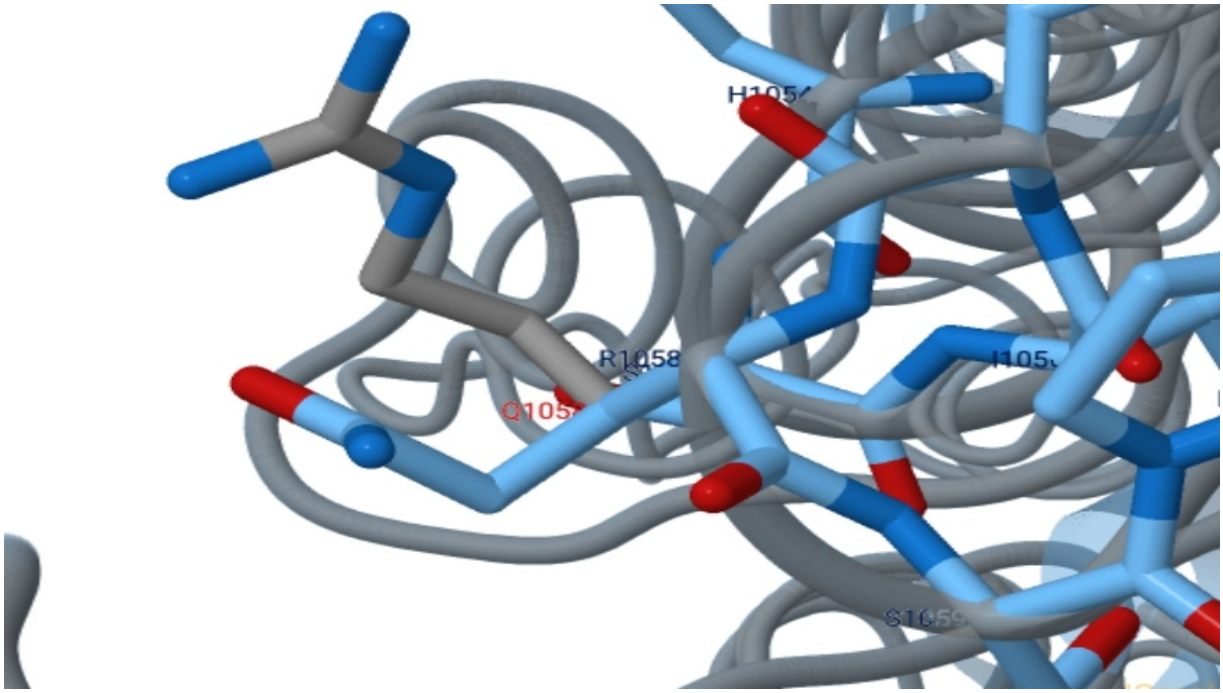


Figure 5.30 ELASPIC result of rs41410450 (R1058Q): showing both wild residue and mutant residue

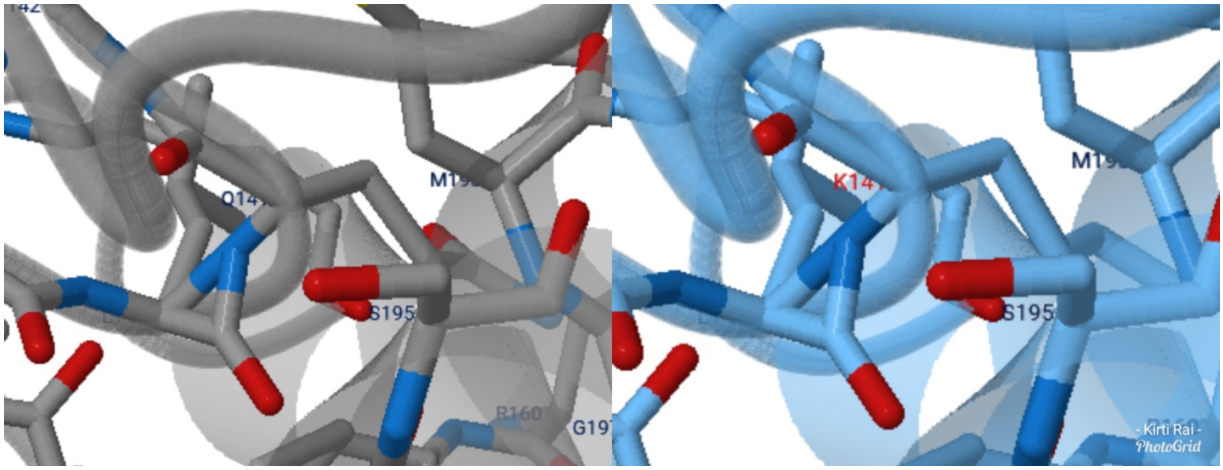


Figure 5.31 ELASPIC result of rs2231142 (Q141K): wild residue (left) and mutant residue (right)

Chapter 6

Discussion

Chemotherapy is contemplated as the vital element in cancer therapeutics. However, drug resistance to chemotherapy has been comprehended as the root cause of therapeutic inadequacy and human adversity, particularly in cases with poor prognosis (Chen *et al.,* 2010). *ABCB1* gene, expressing P-gp, which is a drug efflux protein expressed on luminal epithelial cell surfaces, has been extensively studied for multidrug resistance (MDR). MDR restricts the potency of diverse cytotoxic drugs employed in the treatment of various types of lung cancer. It has been recognized that polymorphisms of ABCB1 protein play a substantial part in influencing chemotherapy outcomes (Knez *et al.,* 2012). In the present study, we scrutinized the correlation of *ABCB1* gene polymorphisms, rs1128503 (C1236T) and rs2032582 (G2677T), with overall survival of North Indian lung cancer patients, individually, on the basis of gender, smoking status, histological type, clinicopathological parameters and their response to doublet platinum-based chemotherapy.

Very constrained information exists asserting the role of individual *ABCB1* genetic variants in overall survival of lung cancer patients after diagnosis. Besides, contradictory results have been found in studying the association of *ABCB1* genetic variants with overall survival in different types of cancer. In a study conducted by Weissfeld *et al.,* reported a statistically significant role of rs1128503 in overall survival of stage III-IV lung cancer patients after chemotherapy. Also, in a study chaperoned in Brazilian acute myeloid lymphoma (AML) patients by Scheiner *et al*., reported that AML subjects harboring CC genetic variants of C1236T polymorphism exhibited a better 5-year overall survival and 5- year Event-free survival rates as compared with individuals harboring other genotype. On the contrary, another study performed by Green *et al.* which included Swedish AML patients, demonstrated a poorer overall survival of subjects harboring CC genotype of C1236T and GG genotype of G2677T polymorphisms. Megias-Vericat *et al.* stated that subjects carrying variant allele of C1236T and G2677T/A showcased a significantly higher OS.

Contrasting outcomes have been noted in breast cancer patients. Chang *et al.* stated no relationship between G2677T/A and overall survival in metastatic breast cancer patients treated with paclitaxel-based chemotherapy. Nevertheless, Wu *et al.* demonstrated no significant

correlation between overall survival of breast cancer patients treated with anthracycline-based chemotherapy and G2677T/A and C1236T polymorphism. In the present study, our results are in sync with the later findings suggesting no significant association between genotypic distribution of *ABCB1* gene polymorphisms rs1128503 and rs2032582 and overall survival of lung cancer patients in North Indian population.

To evaluate the clinical relevance of *ABCB1* gene polymorphism (rs1128503 and rs2032582) with overall survival of lung cancer patients, on the basis of clinicopathological parameters such as extent of tumor invasion, lymph node involvements and metastasis, we compared the results with earlier studies. Balcerczak *et al.* conducted a study in colorectal cancer patients and found that the genotype CT and TT of C1236T single nucleotide polymorphism were more prevalent in T1/T2 groups as compared to T3/T4 groups. Further, in the same study group, statistically insignificant association was obtained in colorectal patients with reference to lymph node metastasis and C1236T polymorphism of *ABCB1* gene. In addition to this, study conducted with reference to extent of metastasis of colorectal cancer in the same group, Balcerczak *et al.,* revealed that subjects with no metastases (M0 group) harbored CT and TT genotypes of C1236T polymorphism in higher frequency as compared to M1 group whereas subjects with distant metastases M1 groups showcased greater frequency of CC genotype of C1236T polymorphism. Henceforth, this study demonstrated a significant relationship between genotypic distribution of C1236T polymorphism and clinicopathological parameter tumor invasion and metastases in colorectal cancer patients.

In contrast, Knez et al., illustrated no significant association of *ABCB1* gene polymorphism G2677T and overall survival of SCLC Caucasian patients. In another study published by Wu *et al.,* in 2012 which compised of Chinese breast carcinoma subjects, demonstrated association of *ABCB1* polymorphism G2677T/A with clinicopathological parameters and found no statistically significant association between G2677T/A genotypes and tumor size and invasion, clinical stage, histology and lymph node metastasis. Our findings are in agreement with the later study signifying no significant role of *ABCB1* polymorphisms (rs1128503 and rs2032582) in overall survival of lung cancer patients stratified on the basis of tumor invasion, lymph node involvement and metastases.

Further, similar association study was examined between effects of *ABCB1* polymorphism with chemotherapy response and it was found that NSCLC subjects harboring wild genotype GG of G2677T polymorphism were classified as good-responders as compared to subjects holding other genotypes, when treated with Docetaxel-cisplatin chemotherapy (Pan *et al.,*. 2009). In a metaanalysis performed by Yin *et al.* in Asian population diagnosed with NSCLC concluded that subjects who are carriers of GG genotype showcased better response to platinum-based chemotherapy. In another study executed by Green *et al.* in ovarian cancer patients concluded a significant correlation between homozygously mutated patients carrying TT genotype and good chemotherapy response; whereas contrasting results have been observed by Marsh *et al.* signifying no association between G2677T polymorphism and chemotherapy response in ovarian cancer subjects treated with carboplatin-plus-docetaxel/paclitaxel.

Instead, contrasting results were obtained in a study performed by Sohn *et al.* which comprised of SCLC patients undergoing Etoposide-cisplatin chemotherapy, which illustrated no significant association amidst G2677T polymorphism and chemotherapy response. Additionally, Chang *et al.* reported no correlation between G2677T polymorphism and response rate in subjects diagnosed with metatstatic breast cancer undergoing paclitaxel treatment. Further, Chen *et al.* also reported that patients housing wild genotype GG were good responders as compared to those who are carriers of variant and heterozygous genotype, however, this data was not proven significantly. This study harmonizes with later findings stating no significant association between G2677T and chemotherapy response.

When same association was analyzed for *ABCB1* C1236T polymorphism, we found that cancer patients treated with irinotecan, who are carriers of TT genotype showcased good chemotherapy response (Mathijssen *et al.,* 2003). Also, in a pooled and subgroup meta-analysis conducted by Zhou *et al.* comprising of breast cancer and osteosarcoma patients illustrated that individuals who are carriers of at least one T allele in their genotype (CT or TT) showcased an enhanced chemotherapy response in comparison to those harboring CC genotype. Yang *et al.* found that TT genotype of C1236Tpolymorphism is related with good chemotherapy response in osteosarcoma patients. However, in our study we found contradictory results stating no correlation between C1236T polymorphism and response to chemotherapy.

Knez *et al.,* demonstrated no statistically significant correlation between genotype distribution of *ABCB1 rs2032582* polymorphism (G2677T/A) with age, gender, performance status (ECOG) and stage of disease in subjects diagnosed with SCLC. Our findings are in sync with this study manifesting no significant dependency between this polymorphism and performance status (ECOG and KPS) of the subjects, clinical stage of the disease and gender. Additionally, not much study has been done in *ABCB1* gene polymorphism (rs1128503 and rs2032582) with regard to various cancer types.

Also, our findings indicate a significant relationship between G2677T polymorphism and overall survival of lung cancer patients diagnosed with Squamous cell carcinoma in North Indian population, evaluated by Kaplan-Meier survival analysis. According to this study, homozygously and heterozygously mutated individuals harboring TT and GT genotype respectively of G2677T polymorphism showcased an increased overall survival (HR = 0.046 and 0.1314, **Logrank P < 0.0001 and Logrank P = 0.017** respectively) as compared to the wild genotype carrying individuals.

However, any divergence observed in this study may be firstly, a consequence of small sample size and secondly population was restricted to North India.

Apart from evaluating statistical significance of genotype frequency on overall survival of lung cancer patients, we additionally demonstrated the impact of various non-synonymous single nucleotide polymorphisms of ABC superfamily genes on protein structure, function and stability by applying *in silico* tools such as SIFT, Polyphen-2, MutPred2, PMUT, PROVEAN, SNPs & GO, PhD-SNP, PANTHER, MUpro and Imutant3.

Out of the 8 nsSNPs submitted to SIFT software, 2 SNPs with rs ID (rs41410450, R→Q and rs8187710, C→Y) were predicted to have damaging effects on the protein with SIFT score < 0.05. Polyphen2 labeled SNPs (rs41395947 and rs41410450) as probably damaging. According to MutPred2, rs2032582 (S→T), rs41395947 (C→S) and rs41494447 (T→I) were forecasted as damaging and affect functional activity of the respective protein. PMUT and PROVEAN predicted rs41395947 and rs41410450 as deleterious. Further, out of 8 SNPs submitted to PhD-SNP, 3 were obtained to be deleterious: rs2032582, rs41395947 and rs4148356. All the 8 nsSNPs given in SNPs & GO were predicted as neutral. In the present study, it was found the rs41410450 were predicted by 4 softwares to have damaging effects (SIFT, Polyphen2, PMUT and

PROVEAN). SNP rs41395947 was predicted by PhD-SNP, PMUT, PROVEAN and MutPred2 as damaging.

Chapter 7

Conclusion

This case study appertains to patients visiting PGIMER, Chandigarh. PGIMER is a referral centre for patients belonging to Union territory Chandigarh, and states of Punjab, Haryana, Himachal Pradesh, Uttar Pradesh and Jammu & Kashmir. The following points can be concluded from our study:

- *ABCB1* gene polymorphisms, rs1128503 and rs2032582, individually, were not found to be associated with overall survival of lung cancer patients belonging to North India.
- Statistically significant association was marked between *ABCB1* rs2032582 single nucleotide polymorphism and overall survival of SQCC patients.
- Another evident observation in our study includes a remarkably increased median survival time of patients, treated with platinum-based chemotherapy in combination with Docetaxel/Irinotecan/Pemetrexed, harboring heterozygous and mutant genotypes of *ABCB1* rs1128503 (Dominant model TT) and rs2032582 polymorphisms as compared to patients with wild genotypes. However, this data was not proven statistically.
- The combined genotype of *ABCB1* rs1128503 polymorphism (Recessive model CC) in lung cancer patients was found to be statistically associated with better overall survival in non-smokers.
- *ABCB1* gene variants included in this study were not found to be associated with chemotherapy response in patients treated with platinum-based doublet chemotherapy.
- Different bioinformatics softwares predicted 6 nsSNPs of the few selected from *ABC* gene subfamily to have damaging effects on their respective ABC transporter. *ABCB1* rs2032582 (S→T) SNP was predicted by MutPred2 to affect the functional activity of the protein. Also, PhD-SNP classified this amino acid substitution as disease-associated.

In conclusion, the present case study demonstrated that the single nucleotide polymorphisms of ABCB1 drug efflux transporter may considerably contribute in estimating the various factors affecting the prognosis and treatment outcome of North Indian patients suffering from lung

cancer. These genetic variants may prove to be as useful biomarkers for examining and designing prospective diagnostic and therapeutic tools.

Bibliography

- *Aller, S. G., Yu, J., Ward, A., Weng, Y., Chittaboina, S., Zhuo, R., ... & Chang, G. (2009). Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science, 323(5922), 1718-1722.*
- *Ambudkar, S. V., Kim, I. W., & Sauna, Z. E. (2006). The power of the pump: mechanisms of action of P-glycoprotein (ABCB1). european journal of pharmaceutical sciences, 27(5), 392-400.*
- *Ankathil, R. (2017). ABCB1 genetic variants in leukemias: current insights into treatment outcomes. Pharmacogenomics and personalized medicine, 10, 169.*
- *Azzoli, C. G., Giaccone, G., & Temin, S. (2010). American Society of Clinical Oncology clinical practice guideline update on chemotherapy for stage IV non–small-cell lung cancer. Journal of oncology practice, 6(1), 39-43.*
- *Balcerczak, E., Panczyk, M., Piaskowski, S., Pasz-Walczak, G., Sałagacka, A., & Mirowski, M. (2010). ABCB1/MDR1 gene polymorphisms as a prognostic factor in colorectal cancer. International journal of colorectal disease, 25(10), 1167-1176.*
- *Berliner, N., Teyra, J., Çolak, R., Lopez, S. G., & Kim, P. M. (2014). Combining structural modeling with ensemble machine learning to accurately predict protein fold stability and binding affinity effects upon mutation. PloS one, 9(9), e107353.*
- *Binder, D., Hackenthal, M., Graseck, L., Schweisfurth, H., Schäper, C., Krüll, M., ... & Hellriegel, K. P. (2009). A phase II study of weekly docetaxel-cisplatin as first-line treatment for advanced non-small cell lung cancer. Journal of Thoracic Oncology, 4(9), 1144-1147.*
- *Bray, F., Jemal, A., Grey, N., Ferlay, J., & Forman, D. (2012). Global cancer transitions according to the Human Development Index (2008–2030): a population-based study. The lancet oncology, 13(8), 790-801.*
- *Chang, H., Rha, S. Y., Jeung, H. C., Im, C. K., Ahn, J. B., Kwon, W. S., ... & Chung, H. C. (2008). Association of the ABCB1 gene polymorphisms 2677G> T/A and 3435C> T with clinical outcomes of paclitaxel monotherapy in metastatic breast cancer patients. Annals of oncology, 20(2), 272-277.*

- *Chen, S., Huo, X., Lin, Y., Ban, H., Lin, Y., Li, W., ... & Xu, X. (2010). Association of MDR1 and ERCC1 polymorphisms with response and toxicity to cisplatin-based chemotherapy in non-small-cell lung cancer patients. International journal of hygiene and environmental health, 213(2), 140-145.*
- *Chen, Z., Shi, T., Zhang, L., Zhu, P., Deng, M., Huang, C., ... & Li, J. (2016). Mammalian drug efflux transporters of the ATP binding cassette (ABC) family in multidrug resistance: A review of the past decade. Cancer letters, 370(1), 153-164.*
- *Cree, I. A., & Charlton, P. (2017). Molecular chess? Hallmarks of anti-cancer drug resistance. BMC cancer, 17(1), 10.*
- *Croce, C. M. (2008). Oncogenes and cancer. New England Journal of Medicine, 358(5), 502-511.*
- *Delbaldo, C., Michiels, S., Syz, N., Soria, J. C., Le Chevalier, T., & Pignon, J. P. (2004). Benefits of adding a drug to a single-agent or a 2-agent chemotherapy regimen in advanced non–small-cell lung cancer: a meta-analysis. Jama, 292(4), 470-484.*
- *Douillard, J. Y., Rosell, R., De Lena, M., Carpagnano, F., Ramlau, R., Gonzáles-Larriba, J. L., ... & Clary, C. (2006). Adjuvant vinorelbine plus cisplatin versus observation in patients with completely resected stage IB–IIIA non-small-cell lung cancer (Adjuvant Navelbine International Trialist Association [ANITA]): a randomised controlled trial. The lancet oncology, 7(9), 719-727.*
- *Eckford, P. D., & Sharom, F. J. (2009). ABC efflux pump-based resistance to chemotherapy drugs. Chemical reviews, 109(7), 2989-3011.*
- *Global Burden of Disease Cancer Collaboration. (2015). The global burden of cancer 2013. JAMA oncology, 1(4), 505.*
- *Green, H., Söderkvist, P., Rosenberg, P., Horvath, G., & Peterson, C. (2006). mdr-1 single nucleotide polymorphisms in ovarian cancer tissue: G2677T/A correlates with response to paclitaxel chemotherapy. Clinical Cancer Research, 12(3), 854-859.*
- *Group, N. M. A. C. (2008). Chemotherapy in addition to supportive care improves survival in advanced non–small-cell lung cancer: a systematic review and meta-analysis of individual*

patient data from 16 randomized controlled trials. Journal of Clinical Oncology, 26(28), 4617.

- *Hanahan, Douglas, and Robert A. Weinberg. (2011). "Hallmarks of cancer: the next generation." Cell, 144.5, 646-674*
- *Hodges, L. M., Markova, S. M., Chinn, L. W., Gow, J. M., Kroetz, D. L., Klein, T. E., & Altman, R. B. (2011). Very important pharmacogene summary: ABCB1 (MDR1, P-glycoprotein). Pharmacogenetics and genomics, 21(3), 152.*
- *International Adjuvant Lung Cancer Trial Collaborative Group. (2004). Cisplatin-based adjuvant chemotherapy in patients with completely resected non–small-cell lung cancer. New England Journal of Medicine, 350(4), 351-360.*
- *Kim, E. S. (2016). Chemotherapy resistance in lung cancer. In Lung Cancer and Personalized Medicine . Springer, Cham. 189-209.*
- *Kim, Y., & Chen, J. (2018). Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. Science, eaar7389.*
- *Knez, L., Košnik, M., Ovčariček, T., Sadikov, A., Sodja, E., Kern, I., & Cufer, T. (2012). Predictive value of ABCB1 polymorphisms G2677T/A, C3435T, and their haplotype in small cell lung cancer patients treated with chemotherapy. Journal of cancer research and clinical oncology, 138(9), 1551-1560.*
- *Lawania, S., Singh, N., Behera, D., & Sharma, S. (2017). Xeroderma pigmentosum complementation group D polymorphism toward lung cancer susceptibility survival and response in patients treated with platinum chemotherapy. Future Oncology, 13(29), 2645-2665.*
- *Lilenbaum, R. C., Herndon, J. E., List, M. A., Desch, C., Watson, D. M., Miller, A. A., ... & Weeks, J. C. (2005). Single-agent versus combination chemotherapy in advanced non–small-cell lung cancer: The Cancer and Leukemia Group B (study 9730). Journal of Clinical Oncology, 23(1), 190-196.*
- *Madrid-Paredes, A., Cañadas-Garre, M., Sánchez-Pozo, A., Expósito-Ruiz, M., & Calleja-Hernández, M. Á. (2017). ABCB1 gene polymorphisms and response to chemotherapy in breast cancer patients: A meta-analysis. Surgical oncology, 26(4), 473-482.*

- *Mamo, G., & Pandi, A. (2017). A Review on ATP Binding Cassette (ABC) Transporters. Int J Pharma Res Health Sci, 5(2), 1607-1615.*
- *Marsh, S., Paul, J., King, C. R., Gifford, G., McLeod, H. L., & Brown, R. (2007). Pharmacogenetic assessment of toxicity and outcome after platinum plus taxane chemotherapy in ovarian cancer: the Scottish Randomised Trial in Ovarian Cancer. Journal of Clinical Oncology, 25(29), 4528-4535.*
- *Mathijssen, R. H., Marsh, S., Karlsson, M. O., Xie, R., Baker, S. D., Verweij, J., ... & McLeod, H. L. (2003). Irinotecan pathway genotype analysis to predict pharmacokinetics. Clinical Cancer Research, 9(9), 3246-3253.*
- *Page, R., & Takimoto, C. (2004). Principles of chemotherapy. Cancer Management: A Multidisciplinary Approach Medical, Surgical & Radiation Oncology. Editors: R Pazdur, LR Coia, WJ Hoskins, LD Wagman. PRR, New York, 21-38.*
- *Pallis, A. G. (2012). A review of treatment in non-small-cell lung cancer. Eur. Oncol. Haematol, 8(4), 208-12.*
- *Pan, J. H., Han, J. X., Wu, J. M., Huang, H. N., Yu, Q. Z., & Sheng, L. J. (2009). MDR1 single nucleotide polymorphism G2677T/A and haplotype are correlated with response to docetaxel-cisplatin chemotherapy in patients with non-small-cell lung cancer. Respiration, 78(1), 49-55.*
- *Rosell, R., Perez-Roca, L., Sanchez, J. J., Cobo, M., Moran, T., Chaib, I., ... & Diz, P. (2009). Customized treatment in non-small-cell lung cancer based on EGFR mutations and BRCA1 mRNA expression. PloS one, 4(5), e5133.*
- *Rowell, N. P., & Williams, C. J. (2001). Radical radiotherapy for stage I/II non-small cell lung cancer in patients not sufficiently fit for or declining surgery (medically inoperable): a systematic review. Thorax, 56(8), 628-638.*
- *Scagliotti, G. V., Fossati, R., Torri, V., Crino, L., Giaccone, G., Silvano, G., ... & Tonato, M. (2003). Randomized study of adjuvant chemotherapy for completely resected stage I, II, or IIIA non–small-cell lung cancer. Journal of the National Cancer Institute, 95(19), 1453-1461.*
- *Scheiner, M. A. M., da Cunha Vasconcelos, F., da Matta, R. R., Figueira, R. D. B., & Maia, R. C. (2012). ABCB1 genetic variation and P-glycoprotein expression/activity in a cohort of*

Brazilian acute myeloid leukemia patients. Journal of cancer research and clinical oncology, 138(6), 959-969.

- *Sharma, D. C. (2016). Cancer data in India show new patterns. The Lancet Oncology, 17(7), e272.*
- *Sharom, F. J. (2008). ABC multidrug transporters: structure, function and role in chemoresistance*
- *Sharom, F. J. (2011). The P-glycoprotein multidrug transporter. Essays in biochemistry, 50, 161-178.*
- *Shepherd, F. A., Dancey, J., Ramlau, R., Mattson, K., Gralla, R., O'rourke, M., ... & Coughlin, S. (2000). Prospective randomized trial of docetaxel versus best supportive care in patients with non–small-cell lung cancer previously treated with platinum-based chemotherapy. Journal of Clinical Oncology, 18(10), 2095-2103.*
- *Shewach, D. S., & Kuchta, R. D. (2009). Introduction to cancer chemotherapeutics.*
- *Singh, Chinnappan Ravinder, and Kandasamy Kathiresan. 2014. "Molecular understanding of lung cancers–A review." Asian Pacific journal of tropical biomedicine 4, S35-S41.*
- *Sohn, J. W., Lee, S. Y., Lee, S. J., Kim, E. J., Cha, S. I., Kim, C. H., ... & Park, J. Y. (2006). MDR1 polymorphisms predict the response to etoposide–cisplatin combination chemotherapy in small cell lung cancer. Japanese journal of clinical oncology, 36(3), 137-141.*
- *Štefková, J., Poledne, R., & Hubáček, J. A. (2004). ATP-binding cassette (ABC) transporters in human metabolism and diseases. Physiol Res, 53, 235-243.*
- *Ter Beek, J., Guskov, A., & Slotboom, D. J. (2014). Structural diversity of ABC transporters. The Journal of general physiology, 143(4), 419-435.*
- *Triller N, Korosec P, Kern I, Kosnik M, Debeljak A. (2006). Multidrug resistance in small cell lung cancer: expression of P-glycoprotein, multidrug resistance protein 1 and lung resistance protein in chemo-naive patients and in relapsed disease. Lung Cancer 54(2), 235–240*
- *Vasiliou, V., Vasiliou, K., & Nebert, D. W. (2009). Human ATP-binding cassette (ABC) transporter family. Human genomics, 3(3), 281.*

- *Wangari-Talbot, J., & Hopper-Borge, E. (2013). Drug resistance mechanisms in non-small cell lung carcinoma. Journal of cancer research updates, 2(4), 265.*
- *Ward, A. B., Szewczyk, P., Grimard, V., Lee, C. W., Martinez, L., Doshi, R., ... & Swartz, D. J. (2013). Structures of P-glycoprotein reveal its conformational flexibility and an epitope on the nucleotide-binding domain. Proceedings of the National Academy of Sciences, 110(33), 13386-13391.*
- *Weissfeld, J. L., Diergaarde, B., Nukui, T., Buch, S., Pennathur, A., Socinski, M. A., ... & Romkes, M. (2014). Inherited variation in the ATP-binding cassette transporter ABCB1 and survival after chemotherapy for stage III–IV lung cancer. Journal of Thoracic Oncology, 9(9), 1264-1271.*
- *Wilkens, S. (2015). Structure and mechanism of ABC transporters. F1000prime reports, 7.*
- *Winton, T., Livingston, R., Johnson, D., Rigas, J., Johnston, M., Butts, C., ... & Fry, W. (2005). Vinorelbine plus cisplatin vs. observation in resected non–small-cell lung cancer. New England Journal of Medicine, 352(25), 2589-2597.*
- *Witvliet, D. K., Strokach, A., Giraldo-Forero, A. F., Teyra, J., Colak, R., & Kim, P. M. (2016). ELASPIC web-server: proteome-wide structure-based prediction of mutation effects on protein stability and binding affinity. Bioinformatics, 32(10), 1589-1591.*
- *Yeh JJ, Hsu NY, Hsu WH, Tsai CH, Lin CC, Liang JA (2005) Comparison of chemotherapy response with P-glycoprotein, multidrug resistance-related protein-1, and lung resistance related protein expression in untreated small cell lung cancer. Lung 183(3), 177–183.*
- *Yin, J. Y., Huang, Q., Zhao, Y. C., Zhou, H. H., & Liu, Z. Q. (2012). Meta-analysis on pharmacogenetics of platinum-based chemotherapy in non small cell lung cancer (NSCLC) patients. PloS one, 7(6), e38150.*
- *Zappa, C., & Mousa, S. A. (2016). Non-small cell lung cancer: current treatment and future advances. Translational lung cancer research, 5(3), 288.*
- *Zhou, S. F. (2008). Structure, function and regulation of P-glycoprotein and its clinical relevance in drug disposition. Xenobiotica, 38(7-8), 802-832.*

- *Zhou, Z., Chen, Q., Zuo, D., Wang, H., Hua, Y., & Cai, Z. (2015). ABCB1 (rs1128503) polymorphism and response to chemotherapy in patients with malignant tumors-evidences from a meta-analysis. International journal of clinical and experimental medicine, 8(1), 265.*

ABCB1

ORIGINALITY REPORT

15%	7%	13%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Shweta Lawania, Navneet Singh, Digamber Behera, Siddharth Sharma. "XPC Polymorphism and Risk for Lung Cancer in North Indian Patients Treated with Platinum Based Chemotherapy and Its Association with Clinical Outcomes", Pathology & Oncology Research, 2017 Publication	2%
2	Shweta Lawania, Navneet Singh, Digambar Behera, Siddharth Sharma. "Association of XPA Polymorphisms Towards Lung Cancer Susceptibility and its Predictive Role in Overall Survival of North Indians", Biochemical Genetics, 2018 Publication	1%
3	Sodhi, Kushaldeep Kaur, Charu Bahl, Navneet Singh, Digamber Behera, and Siddharth Sharma. "Functional genetic variants in pre-miR-146a and 196a2 genes are associated with risk of lung cancer in North Indians", Future Oncology, 2015.	1%