IDENTIFICATION OF SNPs IN MLH1 GENE AND THE RISK OF LUNG CANCER IN INDIAN POPULATION

Enrollment No. : **101557**

Name of Student : **Nirwan Upmanyu**

Name of supervisor : **Dr. Aklank Jain**

July 2014

Submitted in fulfillment of the Degree of

Bachelor of Technology

DEPARTMENT OF BIOTECHNOLOGY

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY,

WAKNAGHAT (H.P)

CERTIFICATE

This is to certify that the work titled "**Identification of SNPs in MLH1 Gene and the risk of lung cancer in indian population**" submitted by "**Nirwan upmanyu**" in fulfillment for the award of degree of 4 Year Degree Program B.Tech of Jaypee University of Information Technology, Waknaghat has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of this or any other degree or diploma.

Signature of Supervisor:

Name of Supervisor : **Dr. Aklank Jain**

Designation : **Associate Professor (Biotechnology)**

Date :

ACKNOWLEDGEMENT

We take this opportunity to express our profound gratitude and deep regards to our guide **Dr. Aklank Jain** for his exemplary guidance, monitoring and constant encouragement throughout the course of this thesis. The blessing, help and guidance given by him time to time and the support of the parents shall carry us a long way in the journey of life on which we are about to embark. I owe my special debt of gratitude and heart full thanks to Miss. Akanksha Khandelwal and Mr. Sanjay for devoting their precious time and guiding me with their valuable suggestions and for bearing my mistakes and shortcomings. I would also like to mention about Mr. Ankit who at every step has given his unconditional support and helped me time to time.

Signature of the student:

Name of Student : Nirwan Upmanyu

Date :

SUMMARY

Mismatch repair (MMR) plays a very important role in repairing nucleotide errors (mismatches) during DNA replication. Defects in MMR genes are associated with hereditary nonpolyposis colon cancer (HNPCC) and has been widely studies. MLH1 and MSH2 are two of the MMR genes. Reduced expression levels of the MLH1 and MSH2 proteins have been reported in HNPCC and some human cancers, including lung cancer. Several mutations and polymorphism that have been identified in the MLH1 and MSH2 genes. It is possible that some of these variants may have an effect on the DNA repair capacity and therefore may modulate the susceptibility to lung cancer.Hence the main aim of the project is to find polymorphism i.e. SNP the MLH1 and MSH2 gene with further may act as a genetic biomarker in screening for the risk of Lung Cancer in the Indian population

Signature of Student	**Signature of Supervisor**
Name: Nirwan Upmanyu	**Name: Dr. Aklank Jain**
Date	**Date**

Key Words : MLH1; MSH2; SNPs; Carcinomas; Lung Cancer; polymorphism; Indian; RFLP; Genotyping

TALBLE OF CONTENTS

TABLES

FIGURES

CHAPTER 1: INTRODUCTION

INTRODUCTION

Cancer is a multifactorial disease that results from complex interactions between the genetic background and environmental factors. Cancer risks factors are diverse, complex, and partially understood. These may include tobacco use, dietary factors, radiation exposure, environmental pollutants, certain infections, etc. [1]. It acts as a group that account for approximately 13% of all deaths each year. This makes cancer to be the leading cause of death in the developed world and the second leading cause of death in the developing world [2].

Each day our genome gets exposed to various endogenous and exogenous mutagenic as well as carcinogenic damages. Some of these damages are directly related to replication errors that may hamper the genome integrity [3]. To overcome such defects, our genome has evolved certain defense mechanisms. And this protection is ensured from encoding of about known 150 DNA repair genes [4-5]. Among these, most display variations which when interacted with these agents may explain the individual's susceptibility to cancer risk [3]. These variations are also coupled to the maintenance of genome integrity and regulation of DNA repair pathways that have been identified and reported in public databases [6]. And any defect in these pathways or accumulation of unrepaired DNA would either lead to cell apoptosis or predispose the cell to tumor formation ultimately leading to cancer formation.

Among the genetic variations, commonly occurring single nucleotide polymorphisms (SNPs) in DNA repair genes have been shown to incrementally contribute to cancer risk because of their critical role in maintaining genome integrity such as in case of Lung Cancer[7]. Such mutations and single nucleotide polymorphisms (SNPs) in repair pathway genes may contribute to deficient repair capacity and human cancer risk [8-9]. Therefore, an extensive study is required to unfold the factors governing the cancer predisposition. Hence, in this I have summarized the risk factors involving DNA repair pathway genes of the MMR pathway and their role towards Lung cancer epidemiology.

1.1 DNA REPAIR PATHWAYS

The four major pathways that keep a check on genome integrity are: (a) Base-excision repair (BER), (b) Nucleotide-excision repair (NER), (c) Mismatch repair (MMR), and (d) Double-strand break repair (DSBR) [5]. These pathways involve different sets of DNA repair genes and activates according to the DNA damage that is encountered.

1.2 MISMATCH REPAIR PATHWAY

DNA mismatch repair (MMR) is a highly conserved biological pathway that plays a key role in maintaining genomic stability. MMR corrects DNA mismatches generated during DNA replication, thereby preventing mutations from becoming permanent in dividing cells. MMR also suppresses homologous recombination and was recently shown to play a role in DNA damage signaling. Defects in MMR are associated with genome-wide instability, predisposition to certain types of cancer including HNPCC, resistance to certain chemotherapeutic agents, and abnormalities in meiosis and sterility in mammalian systems. The Escherichia coli MMR pathway has been extensively studied and is well characterized. In E. coli, the mismatch-activated MutS-MutL-ATP complex licenses MutH to incise the nearest unmethylated GATC sequence. UvrD and an exonuclease generate a gap. This gap is filled by pol III and DNA ligase. The GATC sites are then methylated by Dam. Several human MMR proteins have been identified based on their homology to E. coli MMR proteins. These include human homologs of MutS and MutL. Although E. coli MutS and MutL proteins are homodimers, human MutS and MutL homologs are heterodimers. The role of hemimethylated dGATC sites as a signal for strand discrimination is not conserved from E. coli to human. Human MMR is presumed to be nick-directed in vivo, and is thought to discriminate daughter and template strands using a strand-specific nick.

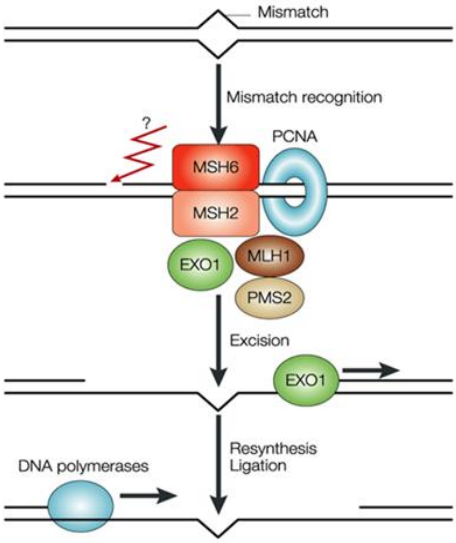

Figure 1: Mismatch repair pathway

Mismatch repair (MMR) genes are involved in numerous cellular functions including:

- Repairing DNA synthesis errors;
- Repairing double-strand DNA breaks;
- Apoptosis;
- Anti-recombination,
- Destabilization of DNA.

1.3 RATIONALE

The *MLH1 genes belong to the* mismatch repair gene family and play an essential role in DNA repair.Variants in *MLH1 are highly heterogeneous and* contribute to genome instability, and for this reason are associated with increased risks of cancer. Germline mutations in four human mismatch repair genes (MSH2, MLH1, PMS1, and PMS2) have been reported to cause hereditary non-polyposis colon cancer syndrome (HNPCC) [10]. *Little studies have been carried out for the MLH1 gene's role in lung cancer.* ***Hence we hypothesize that polymorphism in these MLH1 and MSH2 gene might be also associated with risk of lung cancer.***

1.4 MAIN OBJECTIVES

- To study the genetic polymorphisms in Mismatch Repair Pathway genes (MLH1 and MSH2) in relation to lung cancer
- To compare the incidence rate of the two MMR genes in increasing the susceptibility of an individual to the Lung cancer in the population under study
- To identify the novel polymorphisms of MLH1 and MSH2 genes in addition to the known polymorphisms

CHAPTER 2: LUNG CANCER

LUNG CANCER

2.1 BACKGROUND MATERIAL & LITERATURE REVIEW

Lung cancer is the uncontrolled growth of abnormal cells that start off in one or both lungs; usually in the cells that line the air passages. The abnormal cells do not develop into healthy lung tissue, they divide rapidly and form tumors. As tumors become larger and more numerous, they undermine the lung's ability to provide the bloodstream with oxygen. Tumors that remain in one place and do not appear to spread are known as "**benign tumors**". **Malignant tumors**, the more dangerous ones, spread to other parts of the body either through the bloodstream or the lymphatic system. Metastasis refers to cancer spreading beyond its site of origin to other parts of the body. When cancer spreads it is much harder to treat successfully.

Primary lung cancer originates in the lungs, while **secondary lung cancer** starts somewhere else in the body, metastasizes, and reaches the lungs. They are considered different types of cancers and are not treated in the same way.

Lung cancer can be broadly classified into two main types based on the cancer's appearance under a microscope: non-small cell lung cancer and small cell lung cancer. Non-small cell lung cancer (NSCLC) accounts for 80% of lung cancers, while small cell lung cancer accounts for the remaining 20%. Non-small cell lung cancer (NSCLC) is the most common type of lung cancer. It usually grows and spreads more slowly than small cell lung cancer.

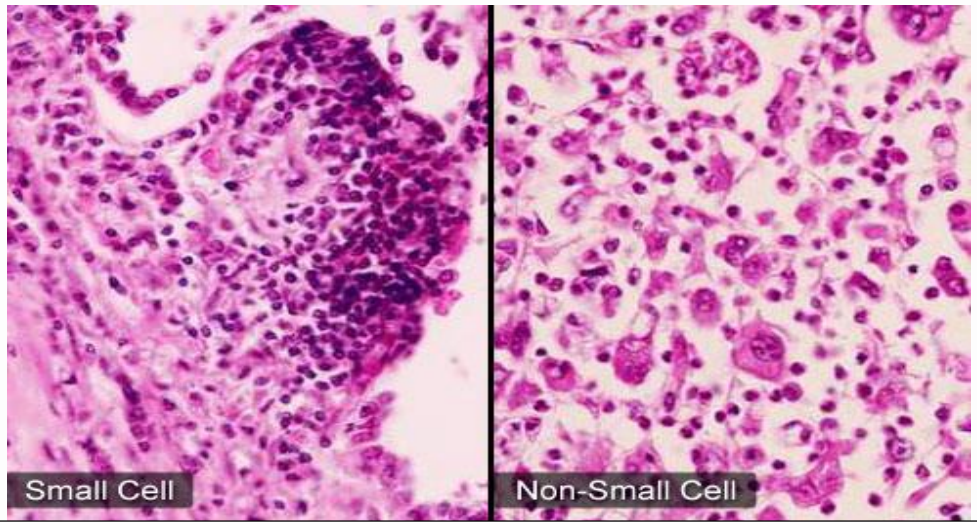

Fig 2 : Difference between small cell lung cancer and non-small cell lung cancer (Nature 2004)

NSCLC can be further divided into four different types, each with different treatment options:

- **Squamous cell carcinoma** or epidermoid carcinoma. As the most common type of NSCLC and the most common type of lung cancer in men, squamous cell carcinoma forms in the lining of the bronchial tubes.
- **Adenocarcinoma.** As the most common type of lung cancer in women and in nonsmokers, adenocarcinoma forms in the mucus-producing glands of the lungs.
- **Bronchioalveolar carcinoma.** This type of lung cancer is a rare type of adenocarcinoma that forms near the lungs' air sacs.
- **Large-cell undifferentiated carcinoma.** A rapidly growing cancer, large-cell undifferentiated carcinomas form near the outer edges or surface of the lungs.

Small cell lung cancer (SCLC) is characterized by small cells that multiply quickly and form large tumors that travel throughout the body. Almost all cases of SCLC are due to smoking.

2.1.1 CAUSES

Smoking causes most cases of lung cancer. The risk depends upon the number of cigarettes smoked every day and for how long someone has smoked. Being around the smoke from others (secondhand smoke) also raises your risk for lung cancer. However, people who do not smoke and have never smoked have become sick with lung cancer. High levels of air pollution and drinking water containing high levels of arsenic can increase your risk for lung cancer. Radiation therapy to the lungs can also increase the risk.

Working with or near the following cancer-causing chemicals or materials can also increase your risk:

- Asbestos
- Products using chloride and formaldehyde
- Certain alloys, paints, pigments, and preservatives

2.1.2 SYMPTOMS

Early lung cancer Symptoms includes:

- Cough that doesn't go away
- Coughing up blood

- Shortness of breath
- Wheezing
- Chest pain
- Loss of appetite
- Losing weight without trying
- Fatigue

Other symptoms that may be due to NSCLC:

- Weakness
- Swallowing difficulty
- Nail problems
- Joint pain
- Hoarseness or changing voice
- Swelling of the face
- Eyelid drooping
- Bone pain or tenderness
- Shoulder pain or weakness

2.2 INCIDENCE RATE

Lung cancer has been the most common cancer in the world for several decades. There are estimated to be 1.8 million new cases in 2012 (12.9% of the total), 58% of which occurred in the less developed regions. The disease remains as the most common cancer in men worldwide (1.2 million, 16.7% of the total) with the highest estimated age-standardised incidence rates in Central and Eastern Europe (53.5 per 100,000) and Eastern Asia (50.4 per 100,000). Notably low incidence rates are observed in Middle and Western Africa (2.0 and 1.7 per 100,000 respectively). In women, the incidence rates are generally lower and the geographical pattern is a little different, mainly reflecting different historical exposure to tobacco smoking. Thus the highest estimated rates are in Northern America (33.8) and Northern Europe (23.7) with a relatively high rate in Eastern Asia (19.2) and the lowest rates again in Western and Middle Africa (1.1 and 0.8 respectively).

Lung cancer is the most common cause of death from cancer worldwide, estimated to be responsible for nearly one in five (1.59 million deaths, 19.4% of the total). Because of its high fatality (the overall ratio of mortality to incidence is 0.87) and the relative lack of variability in survival in different world regions, the geographical patterns in mortality closely follow those in incidence.

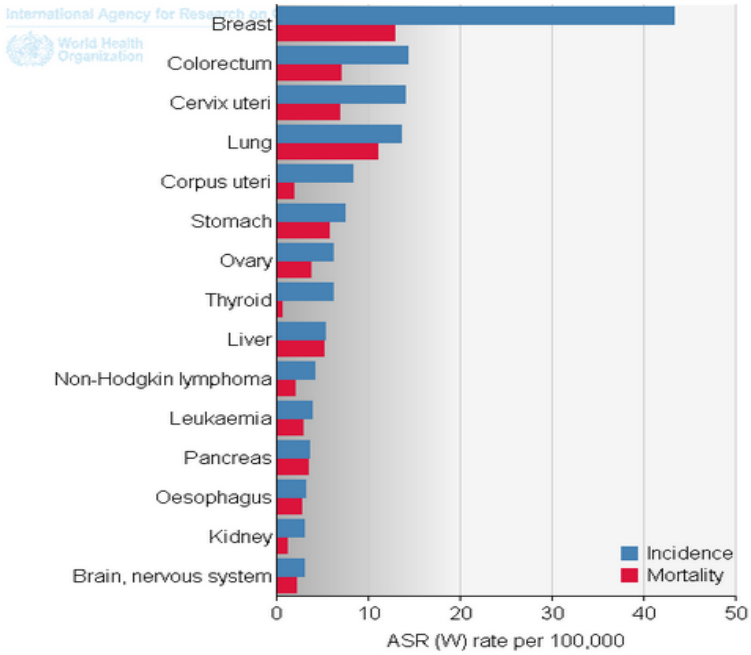

Fig 3. Estimated Age standardized incidence and mortality rates : men (source GLOBOCAN)

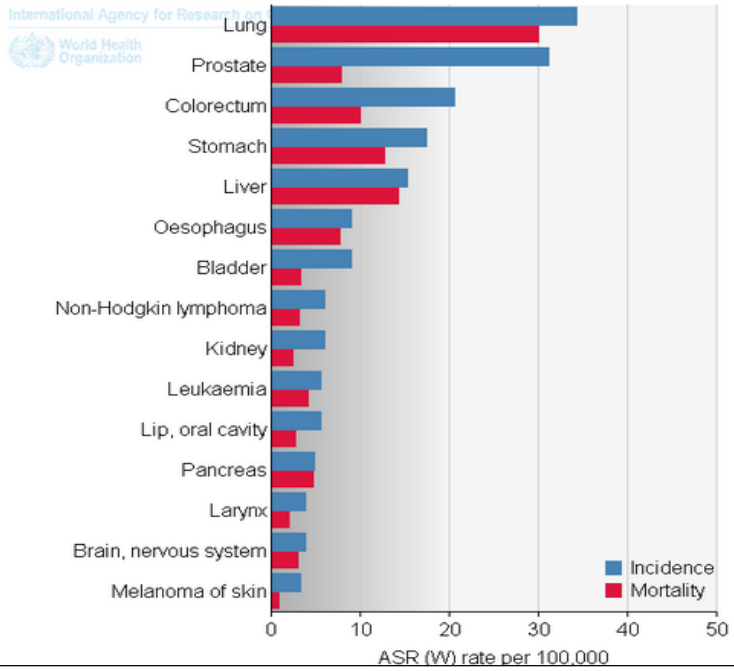

Fig 4. Estimated Age standardized incidence and mortality rates : women source)GLOBOCAN

2.3 THE INDIAN CONTEXT

One million of the current 5 million deaths in world, and 2.41 million in developing countries is contributed by India2,10 and, in 2020, this figure is projected at 1.5 million!

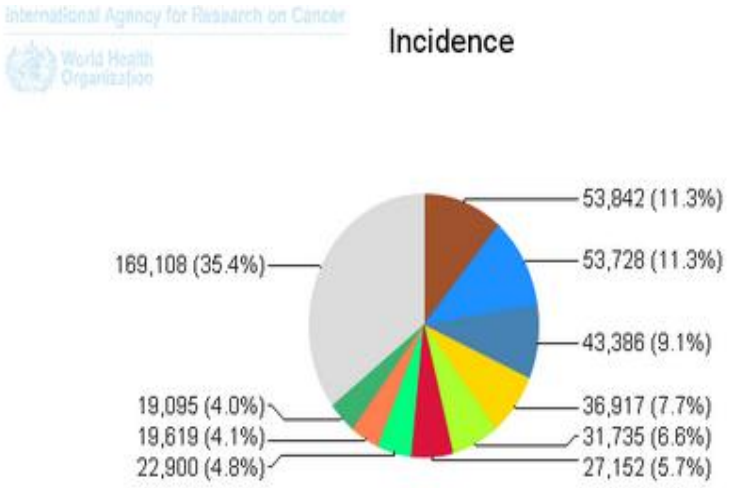

Fig 3. Estimated Age standardized incidence and mortality rates: men source GLOBOCAN

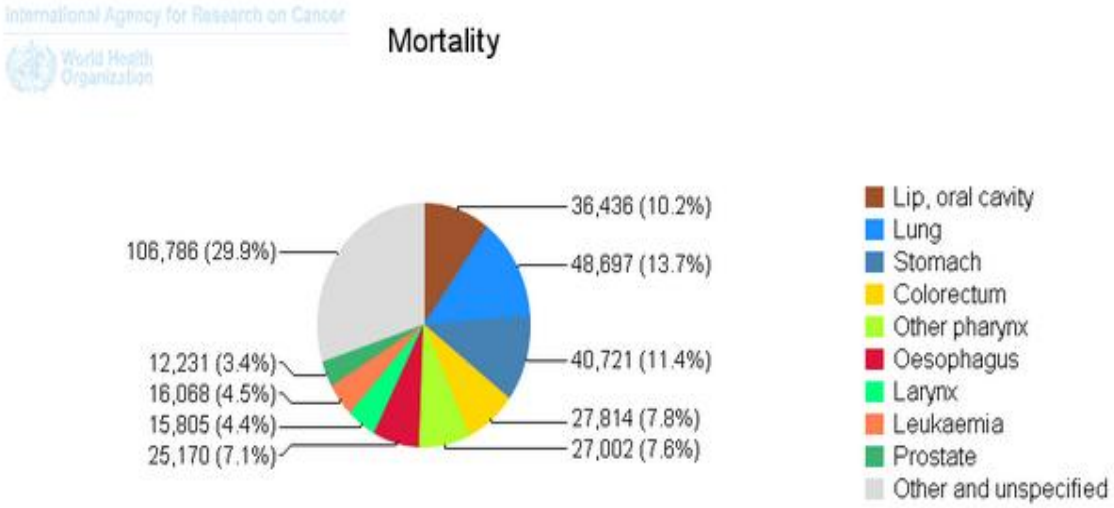

Fig 5. Incidence and Mortality Rates: Indian scenario (Source GLOBOCAN)

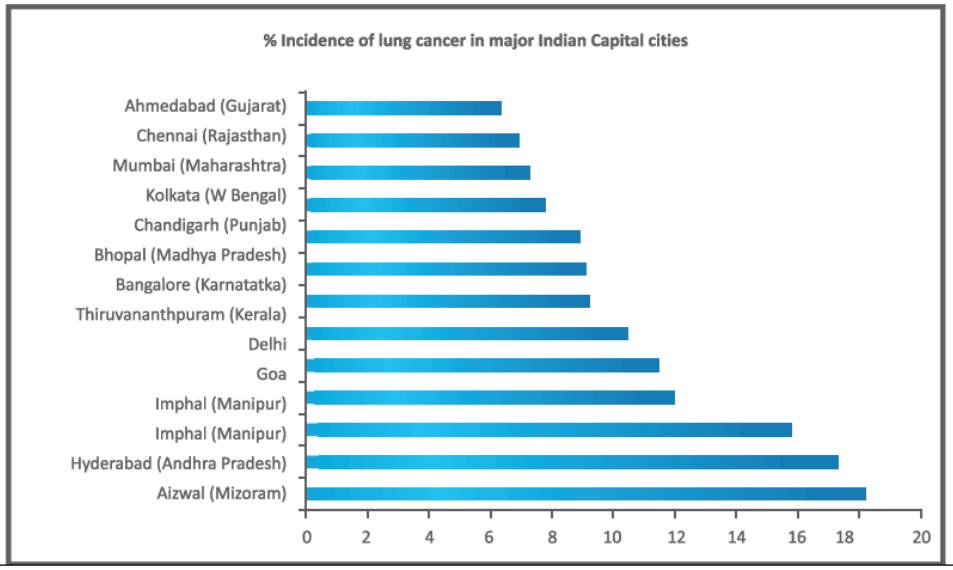

Fig 6. Incidence of Lung Cancer in males in some indian cities (source Siro)

The Indian council of medical research (ICMR), after studying lung cancer data of 24 years (1982-2005), has found that while new cases of lung cancer per one lakh male population has increased by around 160% in Chennai, 100% in Bangalore and 40% in Delhi during this period, such cases have fallen by 60% in Mumbai.

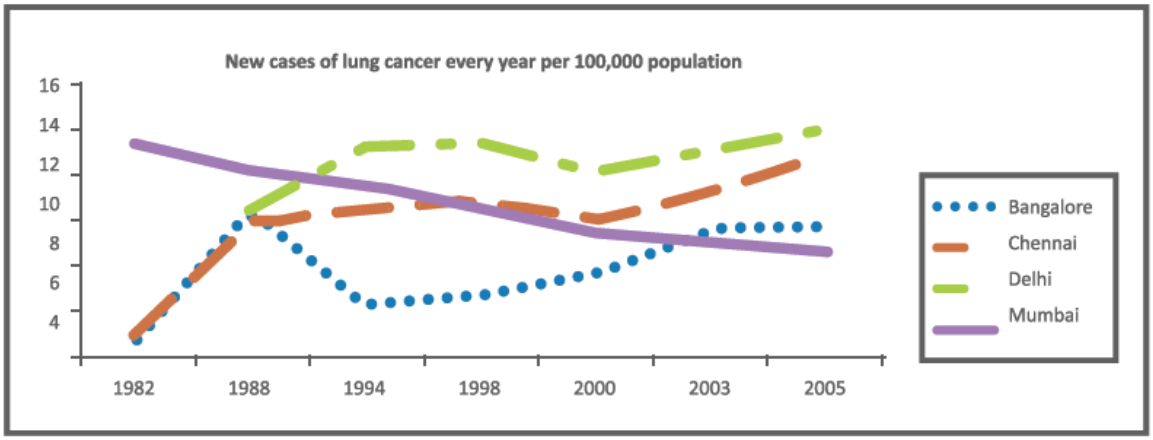

Fig 7.Lung Cancer trends in 4 Metro cities over 24 years

2.4 STAGING OF LUNG CANCER

With the assistance of the diagnostic procedures elaborated the NSCLC can be staged.

- TX : Primary tumor cannot be assessed, or tumor proven by the presence of malignant cells in sputum or bronchial washings but not visualized by imaging or bronchoscopy.
- T0 : No evidence of primary tumor.
- T1 :Tumor 3 cm or less in greatest dimension, surrounded by lung or visceral pleura without bronchoscopic evidence of invasion more proximal than lobar bronchus (i.e., not in the main bronchus).
- T2 : Tumor with any of the following features of size or extent. More than 3 cm in greatest dimension. Involving main bronchus, 2 cm or more distal to the carina. Invading the visceral pleura. Associated with atelectasis or obstructive pneumonitis that extends to the hilar region but does not involve the whole lung.
- T3 : Tumor of any size that directly invades the chest wall (including superior sulcus tumors), diaphragm, mediastinal pleura, phrenic nerve or parietal pericardium. Tumors in the main bronchus less than 2 cm distal to the carina but without involvement of carina. Associated atelectasis or obstructive pneumonitis of the entire lung.
- T4 : Tumor of any size, which invades the mediastinum, heart, great vessels, trachea, esophagus, vertebral body or carina. Tumor with a malignant pleural effusion. Secondary nodules in the same lobe of the lung.

2.5 DIAGNOSIS

Chest X-ray

Patients with lung cancer often have obstructive features (37%) and pleural effusions (22%). Lung cancer patients rarely present with a normal chest X-ray (only 2% in one study).Chest radiography is a simple, cost-effective measure and it imparts very little radiation to the patient. Naturally, it is routinely carried out at every institute.

CT / MRI scanning

Results from computed tomography (CT) scanning are subject to variation caused by different scanning techniques, but suggest that CT scanning of the chest has a high sensitivity (89 to 100%) but a relatively low specificity (56 to 63%) and a poor negative predictive value

(60 to 100%).CT has now become the mainstay of staging chest malignancies and is routinely performed at all major centers in India. Superiority of magnetic resonance imaging (MRI) over CT scan for the detection of bronchial and chest wall invasion or for nodal staging is unestablished. Also as the CT is less expensive and widely available in India it is preferred and routinely advised.

PET scanning

Positron emission tomography (PET) scanning has a diagnostic sensitivity of 96% and a specificity of 78% but there is considerable variation within the studies included. The diagnostic studies indicate negative predictive values as low as 47%.

The considerable cost of the instrument imaging agents as well as the short half-life of positron emitting isotopes (which require a nearby cyclotron for generation), has prevented widespread acceptance and hence these units are available only at a few specialized centers.

Bronchoscopy

The value of bronchoscopy depends on the location of the primary tumor. Peripheral tumors in subsegmental bronchi may not be visible.Flexible bronchoscopy has good diagnostic sensitivity (83% to 88%) for central lesions.

Sputum cytology

There is a wide variation (10% to 97%) in the sensitivity of sputum cytology in the diagnosis of lung cancer. High sensitivity is only achieved by the use of specific and carefully controlled protocols for sample collection

2.6 TREATMENT

1. Radiotherapy

Radiotherapy has an established role in management of lung cancer, both on its own and in combination with chemotherapy. Radiotherapy has a well-documented effect in palliating thoracic symptoms and, in selected cases with NSCLC, it may be curative.

Radiotherapy can also be useful in treating locally symptomatic metastases.

2. Chemotherapy

NSCLC

• Chemotherapy with a platinum-based combination double regimen should be considered

in all patients who are not suitable for curative resection or radical radiotherapy and are fit enough to receive it.

• Second-line chemotherapy with docetaxel should be considered for stage IIIB/IV patients with good performance status.

SCLC

• A regimen containing a platinum agent with etoposide is recommended for first line treatment.

• Second-line in SCLC cases must be considered depending upon the duration of response to first line chemotherapy and on patient's performance status and desire.

3. Surgery

Surgical removal of the tumor is generally performed for limited stages (stage 1 or stage 2). This is the choice of treatment for cancer that has not spread beyond the lungs

2.7 REVEIW

Mismatch repair (MMR) plays an important role in repairing nucleotide mismatches during DNA replication. Defects in MMR genes are associated with some sporadic tumors. MLH1 and MSH2 are two of the MMR genes. Mismatch repair (MMR) is one of the major DNA repair pathways, which is responsible for the repair of single-base mismatches and insertion–deletion loops that result from slippage during replication of repetitive sequences or during recombination . Identification of germline MMR gene mutations in hereditary nonpolyposis colon cancer (HNPCC) confirmed that hereditary cancer could result from MMR gene defects. The MLH1 and MSH2 genes are two of the key components in the MMR pathway located on chromosomes 3p22.2 and 2p21, respectively. The MSH2 gene is responsible for recognizing nucleotide mismatches occurring during DNA replication, and MLH1 is involved with mismatch strand excision and subsequent repair. Reduced expression levels of the MLH1 and MSH2 proteins have been reported in HNPCC and some human cancers, including lung cancer. In addition to germline mutations that have been identified in the MLH1 and MSH2 genes, several polymorphisms have also been identified while their functional contributions, however, are largely known. It is possible that some of these variants may have an effect on the DNA repair capacity and therefore may modulate the susceptibility to lung cancer.

The MLH1 −93A>G polymorphism was significantly associated with the risk of lung cancer among never smokers. Never smokers who develop lung cancer represented a unique subset of all lung cancer patients[13]. Global statistics estimate that 15% of cases of lung cancer in males and 53% of such cases in females are not attributable to smoking , and overall, approximately 10–15% of lung cancers worldwide occur in lifetime never smokers. This was the first epidemiological study thet emphasized the important role of MLH1 in the development of lung adenocarcinoma in never smokers, who represented a particularly susceptible subset of the general population i.e Taiwan population. They found that the MLH1 −93A>G polymorphism was significantly associated with the risk of lung adenocarcinoma among never smokers, which is modulated by low-level exposure to carcinogens, such as ETS. The MLH1 −93A>G (rs1800734) polymorphism has been associated with several cancers, including lung and breast cancers. This polymorphism is located in the MLH1 CpG island, at −93 nucleotides from the transcription start site, in potential transcription factor binding sites. Recent studies have shown that methylation of the MSH2 16638C>T polymorphism showed some indications with risk of lung cancer in the group of ever-smokers. However, the paper did not find any association of the MSH2 polymorphisms with the risk of lung cancer. But the sample size of the patients were small.

In conclusion, they found that the MLH1 −93A>G polymorphism is associated with an increased risk of lung adenocarcinoma among never smokers and that this association may be further modified by ETS exposure. The study also supports the concept in the literature that MLH1 polymorphisms and ETS exposure have a role in tumorigenesis of lung adenocarcinoma among never smokers[13]

CHAPTER 3: EXPERIMENTAL WORK OUTLINE

EXPERIMENTAL WORK OUTLINE

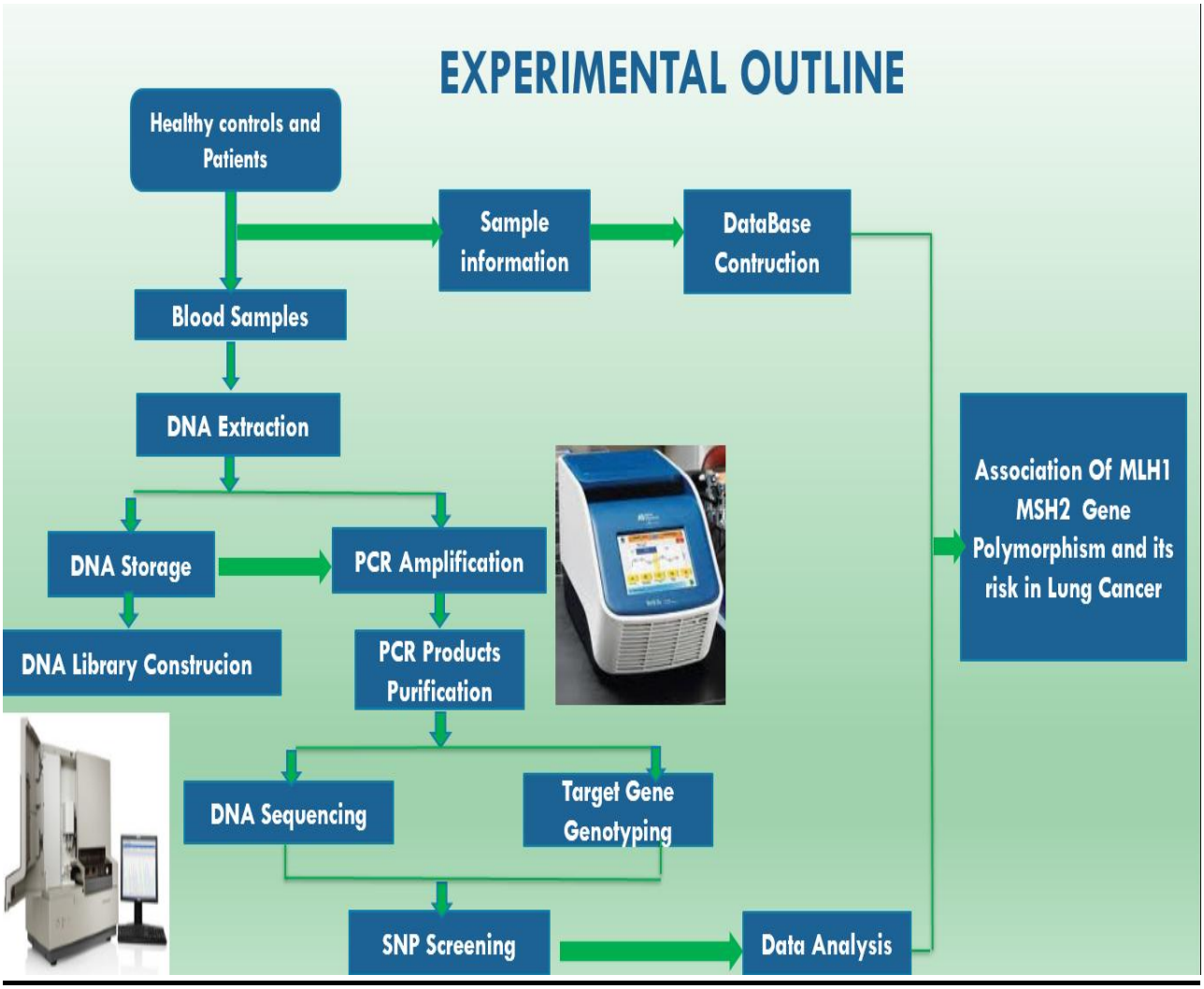

Fig 8. Experimental outline for the SNP

The experimental work outline will be based on the comparative analysis from the healthy control blood samples and the lung cancer patient blood sample taken from IGMC,Shimla with the consent of the patients. Once the blood samples are received then DNA will be extracted which will be run for the PCR with the designed primer for the MLH1 and MSH2 gene .The amplified PCR gene product is then sequenced and further analysis is done using various BI tools such as CUSTALW (multiple alignment tool), and hence fourth will be further aligned to check for any polymorphism at the particular position of the amplified gene and its risk in lung cancer

CHAPTER 4: METHODOLOGY

METHODOLOGY

4.1 ISOLATION OF GENOMIC DNA FROM BLOOD SAMPLES

4.1.1 Materials required

1.Tris(hydroxymethyl)aminomethane-chloride (Tris-cl)

2.Ammonium chloride

3.Di-sodium ethylene diamine tetra acetate

4.Red blood cell(RBC) lysis buffer

5.Tris-EDTA (TE) buffer(pH-7.3,8.0)

6.Sosium dodecyl sulphate(SDS)

7.Ammonium acetate

8.Chilled dehydrated ethyl alcohol

9.Ethanol (70% and 100%)

4.1.2Method:

1. To 300µl of blood sample,added RBC lysis buffer(three times the volume of blodd sample taken) and kept for incubation on a rocker,to permit perpectual shaking at room temperature(RT) until the RBC's completely lysed.

2.Centrifuged the solution at 13,000 rpm foe 1min to obtain a creamish white WBC pellet

3.The supernatant was discarded ant the WBC pellet was thoroughly suspended in 300µl TE buffer(pH-8.0),using a vortexing machine.Thereafter,20µl of 10% SDS solution was added to the above solution and the mixture wsa incubated at 56ºc for 30 minutes in dry bath.

4.Subsequently added 150 µl of 7.5 M ammonium acetate and mixed vigorously for about 1minute per sample,on a vortexer.centifuged the mixture at 13,000rpm at RT for 15 minutes,thereby resulting in separation of the precipitated proteins as pellet.

5.Supernatant was transfred to a fresh vial.to this added chilled absolute ethyl alcohol.The tube was rocked a couple of times to allow precipitation of genomic DNA.

6.The genomic DNA precipitates were centrifuged at 13,00 rpm for 10 min.The latter was subsequently washed in 70% ethanol and air dried at RT foe about 10-15 minutes.

7.100µl of TE buffer was used to dissolve the dried DNA pellet by incubating at 65ºc for 10 minutes.The dissolved DNA was finally stored at 20ºc till PCR is done.

8.The isolated genomic DNA was subjected to quantitative analysis by standard agraose gel electrophoresis

9. PCR amplification of the extracted genomic DNA was followed.

4.2 MICROVOLUME NUCLEIC ACID QUANTIFICATION USING A NANODROP SPECTROPHOTOMETER

1. To begin, clean the upper and lower optical surfaces of the microvolume spectrophotometer sample retention system by pipetting 2 to 3 µL of clean deionized water onto the lower optical surface.

2. Close the lever arm, ensuring that the upper pedestal comes in contact with the deionized water. Lift the lever arm and wipe off both optical surfaces with a clean, dry, lint-free lab wipe.

3. Open the NanoDrop software and select the Nucleic Acid application. Use a small-volume, calibrated pipettor to perform a blank measurement by dispensing 1 µL of buffer onto the lower optical surface. Lower the lever arm and select "Blank" in the Nucleic Acid application.

4. Once the blank measurement is complete, clean both optical surfaces with a clean, dry, lint-free lab wipe.

5. Choose the appropriate constant for the sample that is to be measured.

6. Dispense 1 µL of nucleic acid sample onto the lower optical pedestal and close the lever arm. Because the measurement is volume independent, the sample only needs to bridge the gap between the two optical surfaces for a measurement to be made.

7. Select "Measure" in the application software. The software will automatically calculate the nucleic acid concentration and purity ratios. Following sample measurement, review the spectral output.

8. The software will automatically calculate the nucleic acid concentration and purity ratios. Following sample measurement, review the spectral image to assess sample quality.

9. A typical nucleic acid sample will have a very characteristic profile.

10. To accurately assess sample quality, 260/280 or 260/230 ratios should be analyzed in combination with overall spectral quality. Pure nucleic acids typically yield a 260/280 ratio of ~1.8 and a 260/280 ratio of ~2.0 for DNA and RNA, respectively. This ratio is dependent on the pH and ionic strength of the buffer used to make the blank and sample measurements. Acidic solutions will under-represent the ratio by 0.2-0.3, while a basic solution will over-represent the ratio by 0.2-0.3. Significantly different purity ratios may indicate the presence of protein, phenol or other contaminants that absorb strongly at or near 280 nm. The 260/230 purity ratio is a second measure of DNA purity with values for a "pure" nucleic acid commonly in the range of 1.8-2.2. Purity ratios that are significantly lower than the expected values may indicate the isolation technique used may require further optimization

4.3 AGAROSE GEL ELECTROPHORESIS

1. Rinse and dry the gel casting tray (with 95% ethanol if available).

2. Tape the ends of the casting tray as demonstrated. Set the casting tray on a level surface.

3. Adjust the level of the comb so it rests evenly with a few mm of space between teeth and the tray; this will allow wells to form in the agarose.

4. The 1% agarose has already been prepared (1 g agarose powder per 100 mL 1x TBE), and melted; it is stored in 68oC water bath next door. Gently swirl, make sure there are no particles. Using a 50 mL beaker, measure out about 30 mL. Return agarose stock to water bath.

5. Pour the 25-30 mL agarose into the casting tray.

6. Allow the gel to solidify (about 10 minutes); then remove the comb and tape.

7. Add loading dye to each sample. Plug in the power supply.

8. Turn the power on and adjust the voltage control knob to 120 volts. You should not need totouch any other knob.

9. Electrophorese (run) until the bromophenol blue has migrated to within ¾ of the positive electrode end of the gel.

10. Shut off the power supply, unplug the leads, and unplug the power supply.

11. Lift the gel casting tray from the chamber.

12. EtBr-stained DNA bands will light up under UV illumination.,observe and take a gel pic for future refrences.

4.4 POLYMERASE CHAIN REACTION

1. Usually 20 to 50 µl total in volume and will include the following:
    a. X µl, 0.1 to 1 µg of genomic DNA or cDNA, ~0.1µg should be sufficient for genomic DNA (5)
    b. 10X PCR buffer to give a final concentration of 1X
    c. 4 mM dNTP mix (dCTP, dATP, dGTP, dTTP) to give a final concentration of 0.2 mM

2. Both the forward and reverse primer added at a final concentration of 0.1 µM to 1 µM of each primer
    a. 1 unit/µl *Taq* polymerase
    b. H_2O (DNA and DNase free) to bring volume to 20 µl to 50 µl

3. An example 20 µl reaction
    a. 1 µl of dsDNA template (~0.1 µg)
    b. 2 µl of 10X buffer
    c. 1 µl of 4 mM dNTP mix
    d. 1 µl of 10 µM forward primer to a final concentration of 0.5 µM
    e. 1 µl of 10 µM reverse primer to a final concentration of 0.5 µM
    f. 1 µl of 1 unit/µl *Taq* polymerase
    g. 13 µl of water

4. Combine the reagents in the 0.5-ml tube or the 0.2-ml PCR tube. Be sure to keep the reagents on ice. Tap tube gently to mix and spin briefly in microcentrifuge to get all contents to bottom, then place on ice until ready to load in thermocycler. If thermocycler does not have a heated lid, layer thin film of mineral oil over mixture to prevent evaporation during cycling.

5. Upon completion of PCR, hold samples at 4°C. Prepare the DNA for loading by addition of 1/10 volume stop-loading buffer (contains EDTA, glycerol, and bromphenol blue). Analyze by gel electrophoresis and be sure to include size markers in at least one well on the same gel.

4.5 RESTRICTION DIGESTION OF GENOMIC DNA:

1. Combine the following in a microfuge tube in order (this is for a 20 µl digest; this can be scaled up or down):

a. 2 µl 10x Buffer

b. 10 µl-12 µl of nuclease free water (to 20 µl final volume)

c. 0.2 µl 100X BSA (optional**)

d. 6 µl DNA (0.5-3 µg depending on application)

e. 2 µl Enzyme*

*In general, aim for about 10-fold overdigestion in a 2 hour digest. By definition, 1 unit of enzyme cuts 1 µg of DNA per hour. So if you are cutting 1 µg of DNA for two hours, This means a 10-fold excess would be 5 units of enzyme. Most digests use about 0.5µl of enzyme per reaction, which is usually 5 units.

** In nearly all cases, enzyme performance is enhanced by the presence of acetylated bovine serum albumin (BSA) in the reaction. Sometimes the effect is minor, sometimes major. Rather than try to remember which enzymes require BSA, I always add it!

2. Gently flick the tubes (don't vortex, this denatures proteins) and spin down in microfuge for a few seconds.

3. Place in heat block, water bath, or air incubator or a thermocycler. If you are digesting at a high temperature (50-60C), be aware that the liquid will evaporate, leading to a dried sample at the bottom of the tube and water on the lid. Either spin down the liquid every 15 min, or cover with sterile mineral oil.

CHAPTER 5: REPORTED PRIMERS

REPORTED PRIMERS

5.1 FOR MLH1 GENE

SET I

Forward primers

5'- TAATCCGCTCGAGCCGAGACTCCTAAAAACGAACC-3' and

Reverse primers

5'-GCATCCCAAGCTTCTTTGATAGCATTAGCTGGC-3’

SET II.

Forward Primers

5'- AATATGTACATTAGAGTAGTT-3' and

Reverse Primer

5’-CAGAGAAAGGTCCTGACTC -3’

CHAPTER 6 : RESULTS

RESULTS

6.1 Genomic DNA Isolation (Gel Electrophoresis)

6.1.1: Set I

Table1:Concentration and purity of isolated DNA (Nanodrop readings):

Sr. no	Sample ID	Conc. (ng/µl)	A260/260	A260/230
1.	Control 1	74.7	1.85	1.4
2.	Control 2	157.8	1.79	1.88
3.	Control 3	169.6	1.55	1.4
4	Lung cancer 1	69.7	1.85	2.02
5	Lung cancer 2	195.8	1.84	2.48
6	Lung cancer 3	519.7	1.8	1.64
7	Lung cancer 4	931.5	1.76	1.46
8	Lung cancer 5	457.8	1.85	2.22

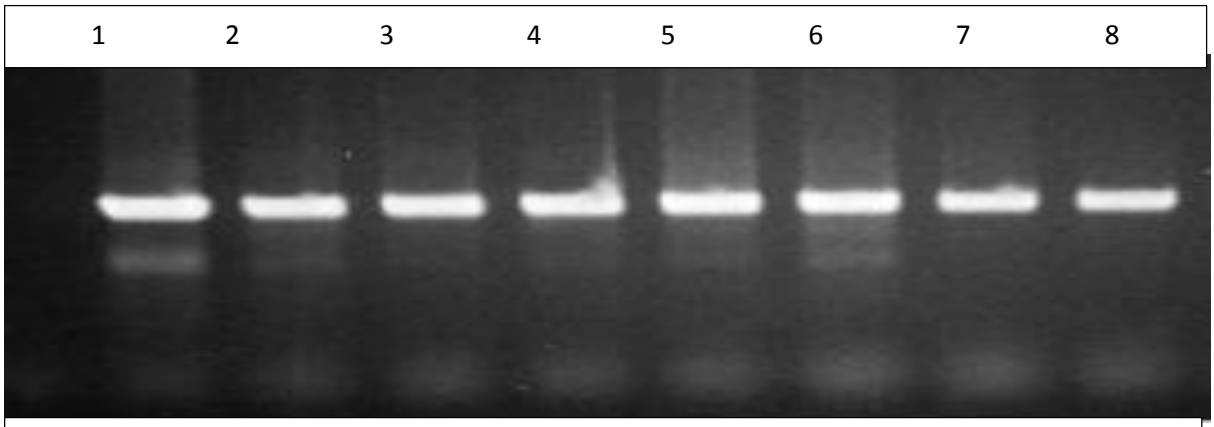

Figure 9: Gel Electrophoresis pattern of genomic DNA isolated from control and lung Cancer patient samples

Lanes	Sample ID
1	Control 1
2	Control 2
3	Control 3
4	Lung cancer 1
5	Lung cancer 2
6	Lung cancer 3
7	Lung cancer 4
8	Lung cancer 5

Table 2:Concentration and purity of isolated DNA (Nanodrop readings):

6.1.2: Set II

Sr.no	Sample ID	Conc. (ng/µl)	A260/280	A260/280
1	Control 4	14.3	1.86	1.13
2.	Control 5	32.3	1.76	2.16
3.	Control 6	115.4	1.90	2.04
4	Lung cancer 6	648.2	1.81	2.03
5	Lung cancer 7	585.1	1.83	2.01
6	Lung cancer 8	1044.6	1.56	1.24
7	Lung cancer 9	39.1	1.56	.86
8	Lung cancer 10	357.0	1.78	1.89

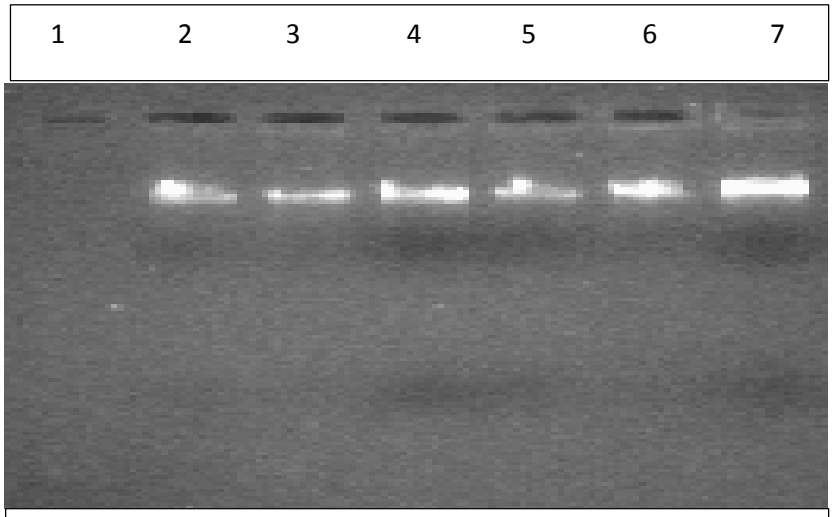

Figure 10: Gel Electrophoresis pattern of genomic DNA isolated from control and Lung Cancer patient samples

Lanes	Sample ID
1	Negative Control
2	Control 5
3	Lung cancer 6
4	Lung cancer 7
5	Lung cancer 8
6	Lung cancer 9
7	Lung cancer 10

6.2.1 MLH1 GENE AMPLIFICATION OF GENOMIC DNA USING SET I OF PRIMER:

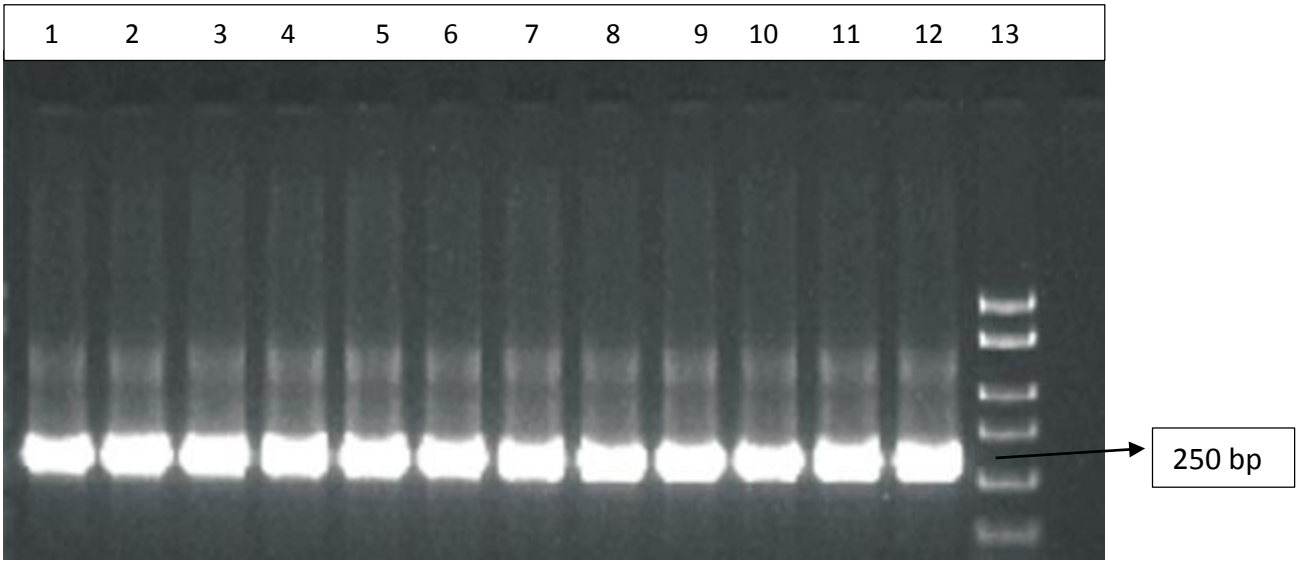

Figure 12: Gel Electrophoresis pattern of Polymerase Chain Reaction (PCR) for detection of MLH1 gene 250 bp amplification.

Lanes	Sample ID
1	Control 1
2	Control 2
3	Control 3
4	Lung cancer 1
5	Lung cancer 2
6	Lung cancer 3
7	Lung cancer 5
8	Lung cancer 4
9	Lung cancer 5
10	Lung cancer 8
11	Lung cancer 9
12	Lung cancer 10
13	Ladder (100) bp

6.2.2 MLH1 GENE AMPLIFICATION USING SET II OF PRIMER:

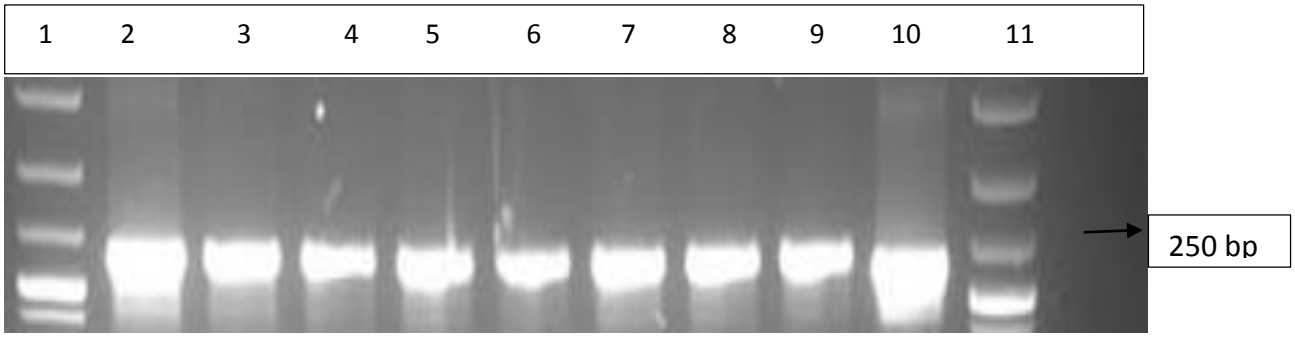

Figure 13: Gel Electrophoresis pattern of Polymerase Chain Reaction (PCR) for detection of XPA gene using set II of primer showing 250 bp amplication.

Lanes	Sample ID
1	Ladder 100 bp
2	Lung cancer 1
3	Lung cancer 2
4	Lung cancer 3
5	Lung cancer 4
6	Lung cancer 5
7	Lung cancer 6
8	Lung cancer 7
9	Lung cancer 8
10	Lung cancer 9
11	Ladder 100 bp

6.3: Restriction Digestion of Control and patient samples with HindIII, using I set of primer

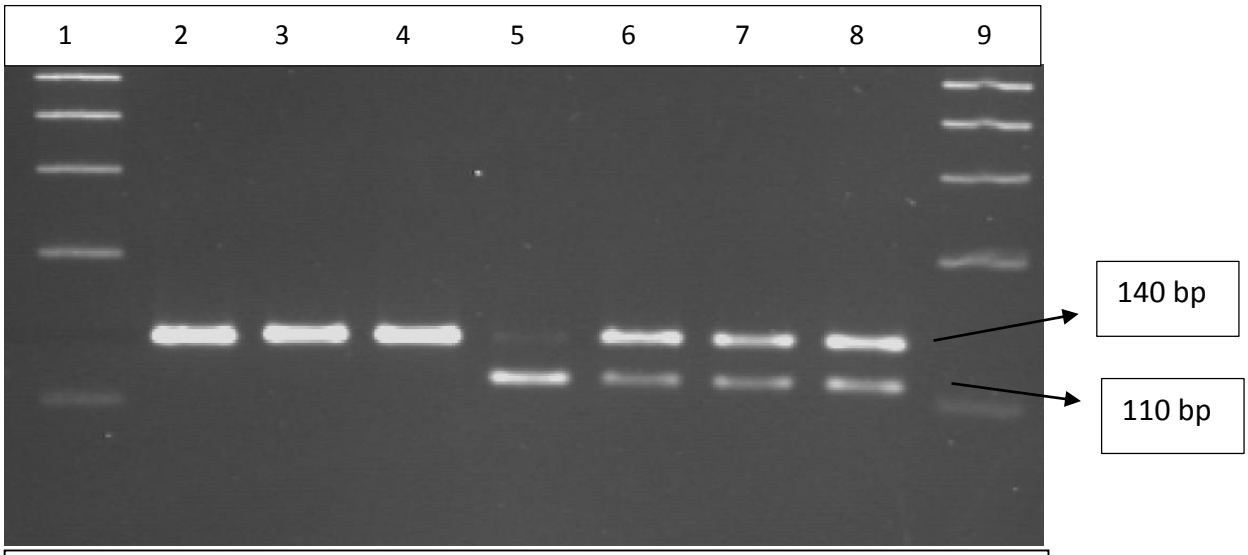

Figure 14: Gel Electrophoresis pattern of restriction digestion from three restriction enzymes used (HindIII,).

Lanes	Sample
1	Molecular ladder (100 bp)
2	Negative control
3	Lung cancer 1 (PCR product)
4	Lung cancer 2(PCR product)
5	Lung cancer 3(PCR Product)
6	Lung cancer 4 (PCR product)
7	Lung cancer 5 (PCr product)
8	Lung cancer 7 (PCR product)
9	Molecular ladder 8 (100 bp)

6.4 SNP DETECTION IN MLH1 GENE

Alignment of five Lung Cancer Samples vs. two Controls of Forward Primer(SetII)

Alignment of five Lung Cancer Samples vs. two Controls of Forward Primer

```
6_MLH1-2F CGCCTTTCGAA--TTAAGGCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 48
7_MLH1-2F CCCACTCCGGG---TA-GGCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 46
8_MLH1-2F CGGGTTTCCAA--TTAGGCCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 48
9_MLH1-2F CGCCTTTC-AA--TAAGGCCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 47
10_MLH1-2FCGCGTTTCCAAAGTTGGGC-GGGCAGCCAGCCGGGCCTGGCGCAGCATCA 49
Control 1 CCCGTTCCGAA---TA-GGCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 46
Control 2 CGCGCTCCCGA---CAAGGCGGGCAGCCAGCCGGGCCTGGCGCAGCATCA 47
           *    * *         *  ****************************

6_MLH1-2F GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 98
7_MLH1-2F GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 96
8_MLH1-2F GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 98
9_MLH1-2F GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 97
10_MLH1-2F GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 99
Control 1 GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 96
Control 2 GTGCCCGCTGCCGCTTCCGCTCGATACTCGCCCGCACCGAGGCAGGCAGC 97
           ************************************************

6_MLH1-2F TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 148
7_MLH1-2F TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 146
8_MLH1-2F TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 148
9_MLH1-2F TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 147
10_MLH1-2FTCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 149
Control 1 TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 146
Control 2 TCCGCGGGTTGCTCTAAAGCCGCCGCCTCCGGCAAAGCCCCGTCGGCCGC 147
           ************************************************

6_MLH1-2F CGCCATCTCCGGACCACTCCT-AGAACCTAGAA-- 180
7_MLH1-2F CGCCATCTCCGGACCACTCCG-AGGACCTAGA--- 177
8_MLH1-2F CGCCATCTCCGGACCACTCCT-AGAACCTAGA--- 179
9_MLH1-2F CGCCATCTCCGGACCACTCCT-AGAACCTAGAATA 181
10_MLH1-2FCGCCATCTCCGGACCACTCCT-AGAACCTAGAA-- 181
Control 1 CGCCATCTCCGGACCACTCCC-AGGACCTAGA--- 177
Control 2 CGCCATCTCCGGACCACTCCCTAGGACCTAGA--- 179
           ********************  ** *******
```

Alignment of five Lung Cancer Samples vs. two Controls of Reverse Primer

```
6_MLH1-2R ---GGGG--GCGGGGGCCGAGGGGGGC-TTTGCCGGAGGCGGCGGCTTTA 44
7_MLH1-2R --TGAAGGAAGCCGGC--GAAGGGGGC-TT-GCCGGAGGCGGCGGCTTTA 44
8_MLH1-2R GGCGAAG--AGCCGGCCGGAAGGGGGC-TTTGCCGGAGGCGGCGGCTTTA 47
9_MLH1-2R -GCGGCG--ACGCCGCCCAAAGGGGGCCTTTGCCGGAGGCGGCGGCTTTA 47
10_MLH1-2R-CCGGAA--CGGCGGCCGAACGGG----CTTGCCGGAGGCGGCGGCTTTA 43
Control 1 --CGGGA--AAGCGGCCGGAAGGGGGC-TTTGCCGGAGGCGGCGGCTTTA 45
Control 2 --CGGGA--AGGCCGGCCGACGGGGGC-TTTGCCGGAGGCGGCGGCTTTA 45
               *          *    * ***     * ********************

6_MLH1-2R *AGCAACCCGCGGAGCTGCCTGCCTCGGT-GCGGGCGAGTATCGAGCGGA 93
7_MLH1-2R GAGCAACCCGCGGAGCTGCCTGCCTCGGTGCGGG-CGAGTATCGAGCGGA 93
8_MLH1-2R GAGCAACCCGCGGAGCTGCCTGCCTCGGTAAGGGGCGAGTATCGAGCGGA 97
9_MLH1-2R GAGCAACCCGCGGAGCTGCCTGCCTCGGTAGCGGGCGAGTATCGAGCGGA 97
10_MLH1-2RGAGCAACCCGCGGAGCTGCCTGCCTCGGATGCGGGCGAGTATCGAGCGGA 93
Control 1 GAGCAACCCGCGGAGCTGCCTGCCTCGGTAGCGGGCGAGTATCGAGCGGA 95
Control 2 GAGCAACCCGCGGAGCTGCCTGCCTCGGT-GCGGGCGAGTATCGAGCGGA 94
           ****************************    ** ***************

6_MLH1-2R AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 143
7_MLH1-2R AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 143
8_MLH1-2R AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 147
9_MLH1-2R AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 147
10_MLH1-2 AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 143
Control 1 AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 145
Control 2 AGCGGCAGCGGGCACTGATGCTGCGCCAGGCCCGGCTGGCTGCCCGGCCC 144
           ************************************************

6_MLH1-2R TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 182
7_MLH1-2R TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 182
8_MLH1-2R TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 186
9_MLH1-2R TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 186
10_MLH1-2RTACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 182
Control 1 TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGC- 184
Control 2 TACTCGGCGACGGCGGCTGCGGCTACTGGAGGTTTGGGCC 184
           ************************************
```

The highlighted nucleotides are the ones showing the difference in samples compared to the healthy controls.
This is a preliminary analysis data of the obtained sequences which will be further analyzed with more samples and with the one reported.

CONCLUSION

To the best of our knowledge, this is the first epidemiological Study in India emphasizing the important role of MLH1 in the development of lung adenocarcinoma. It is a well-researched fact that the MLH1 −93A>G (rs1800734) polymorphism has been associated with several cancers, including lung and breast cancers. This polymorphism is located in the MLH1 CpG Island, at −93 nucleotides from the transcription start site. It is noteworthy that the SNP (rs1800734) is associated with MLH1 methylation in endometrial cancers. The Polymorphisms in the MSH2 gene may have an effect on the DNA repair capacity. However, the functional relevance of the MLH1 −93A>G polymorphism is unknown. The aim of the project deats with finding the possible association of the MLH1 gene with the lung cancer so that it may act as a potential screening biomarker for the susceptibility of the lung cancer. However there were many polymorphism in the gene site but to validate the study, large case controls and lung cancer sample patients must be sequenced and their association must be studied. Furthermore, other genes included in the MMR pathway should also be investigated, which may reveal additional etiological roles of the MMR pathway in lung cancer.

CHAPTER 8: REFERENCES

REFERENCES

1. Anand P, Kunnumakkara AB, Sundaram C, Harikumar KB, Tharakan ST, Lai OS, Sung B, Aggarwal BB. Cancer is a preventable disease that requires major lifestyle changes. *Pharm Res*. 2008 Sep;25(9):2097-116.

2. Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011 Mar-Apr;61(2):69-90.

3. Jalal S, Earley JN, Turchi JJ. DNA repair: from genome maintenance to biomarker and therapeutic target. *Clin Cancer Res*. 2011 Nov 15;17(22):6973-84.

4. Wood RD, Mitchell M, Sgouros J, Lindahl T. Human DNA repair genes. *Science*. 2001 Feb 16;291(5507):1284-9.

5. Wood RD, Mitchell M, Lindahl T. Human DNA repair genes, 2005. *Mutat Res*. 2005 Sep 4;577(1-2):275-83.

6. Packer BR, Yeager M, Burdett L, Welch R, Beerman M, Qi L, Sicotte H, Staats B, Acharya M, Crenshaw A, Eckert A, Puri V, Gerhard DS, Chanock SJ. SNP500Cancer: a public resource for sequence validation, assay development, and frequency analysis for genetic variation in candidate genes. *Nucleic Acids Res*. 2006 Jan 1;34(Database issue):D617-21.

7. Wang M, Chu H, Zhang Z, Wei Q. Molecular epidemiology of DNA repair gene polymorphisms and head and neck cancer. *J Biomed Res*. 2013 May;27(3):179-92.

8. Pastorelli R, Cerri A, Mezzetti M, Consonni E, Airoldi L. Effect of DNA repair gene polymorphisms on BPDE-DNA adducts in human lymphocytes. *Int J Cancer*. 2002 Jul 1;100(1):9-13.

9. Hou SM, Fält S, Angelini S, Yang K, Nyberg F, Lambert B, Hemminki K. The XPD variant alleles are associated with increased aromatic DNA adduct level and lung cancer risk. *Carcinogenesis*. 2002 Apr;23(4):599-603.

10. Päivi Peltomäki et ol.:Deficient DNA mismatch repair: a common etiologic factor for colon cancer. *Hum. Mol. Genet. (2001) 10 (7):735-740.doi: 10.1093/hmg/10.7.735*

11. Lung cancer Focus: India : a report by SIRO 2012

12. Yen-Li Loa, Chin-Fu Hsiaoa, Yuh-Shan Joub, Gee-Chen Chang :Polymorphisms of MLH1 and MSH2 genes and the risk of lung cancer among never smokers:: *Lung Cancer 72 (2011) 280–286*

NOTES