Identification of role of regulatory SNPs rs3213150 and rs3213183 of *E2F1* gene in susceptibility of common cancer in Himachal Pradesh population

Submitted in partial fulfillment of requirements for the award of the degree of

BACHELOR OF TECHNOLOGY
IN
BIOTECHNOLOGY AND BIOINFORMATICS

Under the supervision of

Dr. Harish Changotra
(Associate Professor)
By
Vrishbhanu Handa (141845)
Abhideep Kaur Sihra (141840)
To

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
WAKNAGHAT, SOLAN- 173234
HIMACHAL PRADESH, INDIA

Contents

CERTIFICATE

This is to certify that the work which is being presented in the project title **"Role of regulatory SNPs rs3213150 and rs3213183 of E2F1 gene in susceptibility of common cancer in Himachal Pradesh population"** in partial fulfillment of the requirements for the reward of the degree of Bachelor of technology and submitted in Biotechnolgy and Bioinformatics Department, Jaypee university of Information Technology, Waknaghat is an authentic record of work carried out by **Abhideep Kaur Sihra** (141840) and **Vrishbhanu Handa** (141845) during a period from July 2017 to May 2018 under the supervision of **Dr. Harish Changotra**, Associate Professor, Department of Biotechnology and Bioinformatics, Jaypee University of Information Technology ,Waknaghat.

The above statement made is correct to the best of our knowledge.

Date: /5/2018

Signature of supervisor

Dr. Harish Changotra
Associate Professor
Department of Biotechnology and bioinformatics
Jaypee University of Information Technology
Solan, Himachal Pradesh -173234

ACKNOWLEDGEMENT

We have taken efforts in this project. However, it would not have been possible without the kind support and help of many individuals. We would like to extend our sincere thanks to all of them.

We are highly indebted to Dr.Harish Changotra for their guidance and constant supervision as well as for providing necessary information regarding the project & also for their support in performing the project.

We would like to express our gratitude towards Sir Sanjay, Ma'am Ambika and Ma'am Avni for their kind co-operation and encouragement, which help us in performing this project.

We would like to express our special gratitude and thanks to lab people for giving us such attention and time.

Our thanks and appreciations also go to our college in developing such labs to perform experiments and allowing us to work on such projects and people who have willingly helped us out with their abilities.

We, greatly acknowledge, Dr.Sudhir Syal, Head of the Department for his valuable assistance.

In the end we would like to thank our parents and friends for their support.

Vrishbhanu (141845)

Abhideep (141840)

Date: /05/2018

Place: JUIT, Waknaghat

SUMMARY

Cancer is one of the leading causes of mortality and morbidity worldwide. Cancer results in disturbing the normal life cycle of the cells. This leads to survival of abnormal, old or damaged cells, that should die, and new cells are formed when they are not needed. The damaged cells can multiply without stopping and may results in formation of growths called tumors. It takes multiple mutations to occur over a lifetime to cause cancer. Most of the genes that are mutated in human cancer are directly involved in cell cycle regulation. Early gene factor (E2F) family plays a crucial role in controlling cell cycle and regulates the expression of many genes, which are involved in cell growth and differentiation.

E2F1 is also considered to be a critical modulator of cellular senescence in human cancer. E2F1 genetic 3'untranslated region (3'UTR) miRNA binding site variant are significantly associated with cancer risk. So, we have done a population based study in which we check the susceptibility of common cancers, which are lung cancer, head and neck cancer and cervical cancer in Himachal Pradesh population.

Two variants of E2F1 gene are investigated in this study i.e. rs3213150 and rs3213183. PCR-RFLP is the technique, which is used to check the genotype of all the patients, and then analyzed the data using MedCalc, which revealed the association of variant rs3213150 with lung cancer risk and association of variant rs3213183 with cervical cancer risk.

If these studies are performed on larger number of population and association is found, then these SNPs can be used as the biomarker to check the susceptibility of these cancers.

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ABBREVIATION	
bp	Base pair
ºC	Degree Celsius
CI	Confidence interval
DNA	Deoxyribonucleic acid
SNP	Single Nucleotide Polymorphism
EDTA	Ethylene diamine tetra acetic acid
EtBr	Ethidium Bromide
DHI	di-hydroxy indol
NFW	Nuclease Free Water
µl	Microlitre
mins	Minutes
mM	Millimolar
M	Molar
OR	Odds ratio
PCR	Polymerase chain reaction
RFLP	Restriction fragment length polymorphism
ROS	Reactive oxygen species
secs	Seconds
SNP	Single nucleotide polymorphism
TE buffer	Tris – EDTA buffer
UVR	Ultraviolet rays

CHAPTER 1: INTRODUCTION

Cancer refers to a condition that occurs due to the collection of related diseases. It involves uncontrolled growth of abnormal cells in any part of the body. In all cancer types, abnormal cells start dividing rapidly without stopping and spread to other parts of the body. Cancer can occur anywhere in the human body. Normal human cells grow and divide to form new cells according to the need of the body. When cells grow older or are damaged, they die and new cells are formed.

Cancer results in disturbing the normal life cycle of the cells. This leads to survival of abnormal, old or damaged cells, that should die, and new cells are formed when they are not needed. The damaged cells can multiply without stopping and may results in formation of growths called tumors. [1]

Cancer is one of the leading causes of mortality and morbidity worldwide. Cancer-caused life loss kept increasing since 2000.Cancer of the oral cavity (Oral cancer) is the 11th most common malignancy in the world. The estimated incidence, mortality and 5-year survival due to lip, oral cavity cancer in world are 3, 00, 373(2.1%), 1, 45, 328(1.8%) and 7, 02, 149(2.2%) respectively according to data of GLOBOCAN 2012. A changing trend in incidence and prevalence of oral cancer has been observed with more women and youngsters being affected by oral cancer. Lung cancer has been the most common cancer in the world. There are 1.8 million new cases estimated in 2012 i.e.12.9% of the total, 58% of these cases occurred in the less developed regions around the world. Till date lung cancer remains as the most prevalent cancer among men worldwide (1.2 million, 16.7% of the total), with the highest estimated age standardized incidence rates in Central and Eastern Europe (53.5 per 100,000), and Eastern Asia (50.4 per 100,000).[2]

Mutations happen often, and the human body is normally able to correct most of them. Depending on where the change occurs in the gene, a mutation may be beneficial, harmful, or make no difference at all. So, single mutation alone is unlikely to result to cancer. Usually, it takes multiple mutations to occur over a lifetime to cause cancer. Most of the genes that are mutated in human cancer are directly involved in cell cycle regulation, because these genes are mostly linked to the mechanism that controls cell proliferation.

Many of the genes that contribute to the development of cancer fall into broad categories:

- Tumor suppressor genes are called protective genes. The main function of these genes is to limit the cell growth by monitoring how quickly cells divide to new cells, mismatched DNA repair and controls when a cell dies. Cell growth becomes uncontrolled when a tumor suppressor gene is mutated, and that results in the formation of the mass called tumor. Some of the examples of tumor suppressor genes are **BRCA1, BRCA2**, and **p53**. The germ-line mutations in BRCA1 or BRCA2 genes increase the risk for developing hereditary breast or ovarian cancers in women. p53 is the most commonly mutated gene in cancer patients. In fact, more than 50% of all cancers involve a damaged or missing p53 gene. Most of the p53 mutations are acquired mutations and Germ-line mutations are rare. [3]

- Oncogenes are those genes, which turn a healthy cell into a cancerous cell. These genes do not have inherited mutations.

- DNA repair genes fix mistakes made when DNA is copied. But if a person has an error in a DNA repair gene, these mistakes are not corrected. And then they become mutations, which may eventually lead to cancer.

Cancer is basically a disease of uncontrolled cell division. Its development and progression are usually linked to a series of changes in the activity of cell cycle regulators. For example, inhibitors of the cell cycle keep cells from dividing when conditions aren't right, so too little activity of these inhibitors can promote cancer. Similarly, positive regulators of cell division can lead to cancer if they are too active. In most cases, these changes in activity are due to mutations in the genes that encode cell cycle regulator proteins. [4]

Advances in genetics and molecular biology have improved our knowledge of the inner workings of cells, the basic building blocks of the body. The cells can change during a person's life to become cancer and certain types of changes can build on inherited gene changes to speed up the development of cancer, and this information can help us better prevent and treat cancer.

Early gene factor (E2F) family plays a crucial role in controlling cell cycle and function of the tumors suppressor proteins. E2F regulates the expression of many genes, which are involved in cell growth and differentiation. It mainly regulates the G1/S transition, the DNA replication and gives downstream signaling during following mechanisms:

DNA damage checkpoint response, DNA repairs gene activation, Cell cycle check point, apoptosis induction, differentiation and development. [5]

E2f is the 8-member family (E2F1-E2F8), which are downstream effectors of the retinoblastoma tumor suppressor (Rb), which contains one or more conserved DNA binding domains, which binds to target promoters and regulate their expression.

E2F1 belongs to the E2f family of transcription factors, which coordinates the expression of key genes involved in regulation of cell cycle and proliferation. Due to its pivotal and multifunctional role in cell cycle control, E2F1 is expected to be a significant player in carcinogenesis. E2F1 gene is located on chromosome 20q, approx 10.71Kb and contains 7 exons. Excess of E2F1 may

promote proliferation, but at the same time it may also enhance apoptosis. Over expression or lack of E2F1 has both positive and negative effect on tumorogenesis. It is well known that genetic variants in microRNA (miRNA) binding regions can result in altered gene functions. miRNA are capable of regulating the E2F activity and miRNA dysregulation has been implicated in malignancy. E2F1 is also considered to be a critical modulator of cellular senescence in human cancer. E2F proteins and their target genes comprise a genetic pathway that is possibly the most altered pathway in human cancer [6]

E2F1 genetic 3'untranslated region (3'UTR) miRNA binding site variant are significantly associated with cancer risk.

Worldwide work has been reported on E2F1 gene SNPs such as following:

- rs35301225 polymorphism in miR-34a promotes development of human colon cancer by deregulation of 3′UTR in E2F1 in Chinese population.[7]
- The *TERT* promoter SNP rs2853669 decreases E2F1 transcription factor binding and increases mortality and recurrence risks in liver cancer. [8]

Similarly, rs3213183 and rs3213150 SNPs have high MAFs i.e. 0.2997 and 0.2971 respectively. Work on rs3213183 has been reported in squamous cell carcinoma of the Head and Neck (Meixia Lu et.al, 2012). To date, the function of these single nucleotide polymorphisms (SNPs) located in the promoter region has not been investigated in Indian population. Our project involves the population study to check the association of these mutant alleles with head and neck cancer, lung cancer and cervical cancer.

Significance of the study

There are different association studies available regarding cell cycle regulatory genes like p53, p73, Mdm-2, BRCA-1 and BRCA-2 which reveals the association of genetic variants of these genes with cancer risk. But there are limited studies available regarding *E2F1* regulatory SNPs. Therefore; in this

study we analyzed the association of E2F1 regulatory SNPs with cancer risk. If mutant alleles have higher frequency in cases than in controls then it may have role in carcinogenesis and if high risk in these SNPS is reported then they could be used as a biomarker for cancer patients.

The main objective of this study is to identify the role of regulatory SNPs rs3213150 and rs3213183 of *E2F1* gene in susceptibility of common cancers in Himachal Pradesh population.

CHAPTER 2: REVIEW OF LITERATURE

2.1. WHAT IS CANCER?

Cancer is the collection of diseases involving the uncontrolled growth of abnormal cells. These cells continue to grow, divide and re-divide; without dying and forming new abnormal cells. This results in formation of subset of neoplasms or tumors. A tumor is any abnormal growth of cells and which often results in formation of a mass or lump, but may be diffusely distributed. A tumor can of two types i.e., benign tumor or malignant tumor.

A benign tumor, remains confined to its original location, it do not invade to surrounding normal tissues and other sites of the body. For example, Skin warts.

A malignant tumor, has the ability to invade to other surrounding normal tissues and spreads to other parts of the body through the circulatory or lymphatic systems, which is the main characteristic of cancer cells i.e., metastasis. These malignant tumors are referred as cancers, and this property of cancerous cells to invade and metastasize makes cancer so dangerous. Usually benign tumors can be removed surgically, but malignant tumors are resistant to such localized treatment due to their ability to spread to other parts. [9]

Normally, cells are able to detect any change and repair DNA damage and if the damage is severe and cell is not able to repair itself, then it undergoes programmed cell death or apoptosis. Cancer is caused when damaged cells continue to grow, divide, and spread abnormally instead of dying or undergoing self-destruction as normally happens. [10]

Differences between Cancer Cells and Normal Cells

Cancer cells grow out of control and are invasive in nature; this property of cancer cells makes them different from normal cells. Normal cells are more

specialized than the cancer cells. Matured normal cells have distinct cell types with specific functions, whereas cancer cells do not possess these properties. Due to this reason cancer cells continue to divide without stopping, unlike normal cells.

In addition, cancer cells do not have the ability to detect the signals that normally tell cells to stop dividing or to undergo programmed cell death, or apoptosis, this process is normally used by body to get rid of unwanted cells.

Cancer cells have the ability to alter the surrounding normal cells, and blood vessels, which forms an area called microenvironment to feed growing tumor. Cancer cells induce surrounding normal cells to form blood vessels, which supply oxygen and nutrients to tumor for its growth and also remove waste products from tumors.

Cancer cells mostly escape the immune system, which helps our body to protect from any infection or disease. Normally the immune system removes damaged or abnormal cells from the body but sometimes cancer cells are able to protect themselves from the immune system. [11]

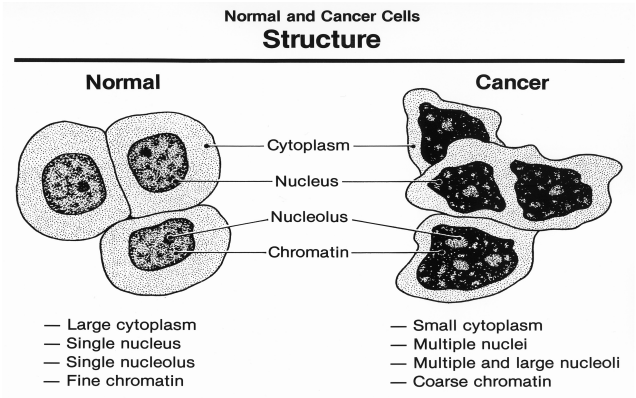

Fig 1: Difference between Normal cells and cancer cells. (National Cancer Institute, an agency part of the National Institutes of Health, with the ID 2512)

CAUSES OF CANCER

Cancer is mainly caused by mutations in various genes such as in oncogenes, tumor-suppressor genes, and microRNA genes. These mutations are usually somatic mutations, however germ line mutations can make a person susceptible to heritable cancer. A single mutation is usually not enough for the

development of a malignant tumor. Various studies suggest that cancer development is a multistep process, which involves sequential alterations in several genes such as oncogenes, tumor-suppressor genes, and microRNA genes in cancer cells. [12]

CANCER CAUSING AGENTS

Cancer causing agents are as following:

Chemical carcinogens

Various chemicals and environmental toxins are responsible for causing alterations in normal cell DNA. These mutation-causing substances are called Mutagens and those cause cancers are called carcinogens.

Various studies have shown association of different substances with specific types of cancer. For example, Smoking Tobacco causes 90% of lung cancer and also it is associated with many other cancer types.

Tobacco is also one of the main causes for different cancers such as lung, head and neck, stomach, bladder, kidney and pancreas as it contains many other known carcinogens, such as nitrosamines and polycyclic aromatic hydrocarbons.

Ionizing radiations

Ionizing radiations are one of the main causes for various cancers. Long exposure to the radiations from radon gas and to ultraviolet radiation can cause various skin malignancies. Radiation therapy can also result in causing other cancer type in later stages while treating some other cancer type. For example, patients those are undergoing chest radiation for lymphomas have the chances to develop breast cancer in later stages. [13]

Viral and Bacterial infections

Various viral and bacterial infections cause cancer. Liver cancers are mainly caused due to Hepatitic B and C infections; cervical cancer is caused due to Human Papilloma virus (HPV) infection; Burkitt's lymphoma is caused by

Epstein Barr virus and gastric or stomach cancer is caused by Helicobacter pylori infection. [14]

Genetic or inherited cancers

Genetic or inherited cancers are mainly caused by inheritable change in the genes. For example, BRAC1 and BRAC2 genes are responsible for inherited breast cancer and ovarian cancer. Any defect in defect in the p53 gene leads to various cancers such as bone cancers, breast cancers, soft tissue sarcomas, brain cancers etc. Individuals with Down's syndrome are more prone to develop cancers such as leukemia and testicular cancer. [15]

Hormonal changes

Hormonal changes can also result in development of cancer. A common example is the change in female hormone levels such, as females with excess of estrogen levels are more prone to develop uterine cancer.

Immune system dysfunction

If our immune system is not strong enough to protect our body from various pathogens, this can lead to several cancers. For example, Kaposi's sarcoma, and non-Hodgkin's lymphoma can occur due to HIV- infection. [16]

2.2 EPIDEMIOLOGY OF LUNG CANCER:

The epidemiology of cancer has been depicted as the study of malignancy distribution by age, sex, financial status, etc. and of those factors, which decides its pervasiveness. [17]

INCIDENCES

Lung cancer is the most widely recognized malignancy in men worldwide with an age-standardized rate (ASR) of 33.8 per 100,000, and it is the fourth most recurrent malignancy in women (13.5 per 100,000). Incidence and mortality credited to lung cancer has increased steadily since the 1930s, primarily due to the prevalence of cigarette smoking.

In men, the highest incidence rates are seen in North America, East Asia, Central Eastern and Southern Europe (48.5 to 56.5 per 100,000). In under developed countries, the highest rates are observed in West Asia, South Africa, and the Caribbean (25.7 to 32.2 per 100,000). In case of men, various nations have now passed the peak of the tobacco related outbreak, and now a decrease in incidence and mortality is observed. [18]

In women, the lung cancer incidence is lower worldwide; the highest rates are observed in North America and in Northern Europe (35.8 to 37 per 100,00). However, the incidence rates has increased over recent time period but has leveled in the past decade, most recently calculated to be 51 per 100,000 in 2009 compared with 39 per 100,000 in 1984.In the past 100 years, lung cancer has transformed from a rare disease into a global problem. Efforts to enhance results have not only prompt to a greater understanding of the etiology of lung cancer, but also the histologic and molecular characteristics of individual lung malignancies. The long-term trends in the age balanced lung cancer incidence among men and women are constant with the historic trend of tobacco use. Incidence rates of lung cancer also vary by ethnicity. [18]

In India, lung malignancy comprise 6.9 per cent of all recent cancer cases and 9.3 per cent of all cancer related mortality in both sexes, it is the most common cancer and reason for cancer related deaths in men, with the highest described incidences from Mizoram in both males and females (Age balanced rate 28.3 and 28.7 per 100,000 population in males and females, respectively).

The time patterns of lung cancer show a significant increase in Delhi, Chennai and Bengaluru in both sexes. The incidence and patterns of lung cancer vary as per geographic area and ethnicity and largely reflect the prevalence and pattern of smoking. [19]

MORTALITY

Worldwide, lung cancer is the commonest cause of death from cancer, with 1.38 million deaths reported in 2008 (18.2% of the total) of cancer deaths. [18]

Global lung malignancy mortality does not vary significantly by region, with 43% of deaths taking place in more developed countries and 57% occurring in under developed countries. Gender differences in lung cancer mortality patterns depicts historical differences between men and women in the increase and decrease in cigarette smoking over the past 50 years. [18]

RISK FACTORS

Tobacco

Smoking is one of the most predominant risk factor for the development of lung cancer. It is assessed that ~90% of deaths in men and 75-80% of deaths in women occur due to lung cancer. There are no less than two ways that smoking is associated with lung malignancy. First, polycyclic aromatic hydrocarbons and carcinogenic compounds present in tobacco smoke that induce mutations in the *p53* gene, which are crucial for cell cycle deregulation and carcinogenesis.

The second way is, the N-nitroso compounds, that are another important group of chemicals found in tobacco smoke, many of these are potential animal carcinogens. These compounds are excreted in the urine of smokers. [20]

Radon

The U.S. Environmental Protection Agency has decided radon to be the second driving reason of lung malignancy after cigarette smoking. The increased risk is credited to radon is from domestic exposure, due to dissemination of radon from the soil. High radon concentrations have been associated to the increased risk of lung cancer in underground miners. Many recent epidemiological studies of domestic radon exposure also recognized it as a risk factor for lung cancer. Inhalation of radon can carcinogenically affect the lungs, due to its discharge of α particles upon decay. [21]

Other Predisposing Risk Factors

The risk of developing a lung cancer again in patients who survived the lung tumor resectioning is ~1 to 2% per patient per year for non-small cell lung cancer (NSCLC), and 6% for small cell lung cancer (SCLC). 10 years after starting treatment of small cell lung cancer, cancer risk rises from ~2% to >10% per patient per year. This risk of developing a second primary lung cancer can convert into an important combined risk and is a common reason for death in lung cancer survivors. [22]

2.3 EPIDEMIOLOGY OF HEAD AND NECK CANCER

Head and neck cancer (HNC) includes a heterogeneous group of upper aero-digestive tract cancers that together contain the seventh commonest cancer worldwide. The yearly incidence of head and neck cancers worldwide is more than 550,000 cases with around 300,000 deaths each year. Male to female proportion ranges from 2:1 to 4:1. Around, 90% of all head and neck cancers are squamous cell carcinomas (HNSCC). HNSCC is the sixth driving malignancy by incidence worldwide. Most HNSCCs emerge in the epithelial lining of the oral cavity, oropharynx, larynx and hypo pharynx. [23,24]

Head and neck cancers are rising in India as crucial public health problems, which are dependent on lifestyle. It also has a long latent period and need fully committed infrastructure and human resource for its treatment. Overall, 57.5% of worldwide head and neck cancers (excluding esophageal malignancies) occur in Asia mainly in India, for both sexes. Head and neck cancers in India represented 30% of all cancers in males. In females they consist of 11 to 16% of all locations of cancers. Over 200,000 cases of head and neck cancers occur every year in India. Almost 80,000 oral cancers are diagnosed each year in our country. [25]

Two major causes of HNC are: (1) Tobacco and alcohol usage, and (2) human papillomavirus (HPV) infection. Historically, most head and neck cancers have been caused by tobacco and alcohol, but over the past decade, Human papillomavirus has been recognized as another major cause of HNC. [26]

RISK FACTORS FOR HEAD AND NECK CANCER

The HNC risk factors incorporate some important risk factors linked with both types of HPV - positive and HPV-negative HNC, and in addition to other causative factors. [26]

Non-infectious Risk Factors for Head and Neck Cancer

Tobacco

Tobacco is one of the major causes of HNC, and use of alcohol collectively increases the risk of HNC already caused by tobacco use. Historically, around 90% of HNC patients have a history of tobacco use. The carcinogenic effects of tobacco are dependent on the dose and the risk of HNC is closely related to the frequency, duration, and intensity of cigarette smoking. Smokeless tobacco is also a significant HNC risk factor, as it is responsible for the malignancies of oral cavity. Examples of smokeless tobacco are, snuff or chewing tobacco. [27]

Alcohol

Alcohol consumption solely increases the risk of HNC and about 1% to 4% of cases are caused due to alcohol use and it increases the risk to two times in odds of HNC for drinkers who never use tobacco. However, the large effect is seen when alcohol is used with tobacco. The combination of alcohol and tobacco use increases the risk of cancer multiplicative times. [28]

Gender and age

Men in various countries are two to five times for susceptible to risk of HNC than women. This is due to greater use of elements such as tobacco and alcohol use, among men than in women. The HNC risk increases with age. [29]

Immune suppression

Immune suppression that is done for solid organ transplantation or in case of HIV infection is associated with an increased HNC risk. After any organ transplantation, chances of lip cancer increase by 10 times or more and also there is two to five times increase of risk at other head and neck locations.

Similarly, HIV- infected patients are at a two to five times increased risk for HNC, which has risen in the post-HAART (highly active anti-retroviral therapy) era because they survive longer and are more susceptible to develop non-AIDS defining malignancies, such as HNC. [30]

Risk Factors for Human Papillomavirus-Positive Head and Neck Cancer

Human papillomavirus

HPV infection is a sexually transmitted, that causes HPV-positive HNC (HPV-HNC), different subtypes of HNCs that takes place majorly in the oropharynx and initiate from the lymphoid tissues of the palatine and lingual tonsils. Oral HPV infection is the pivotal precursor to HPV-HNC. Despite this fact that, there are more than 100 types of HPV but only few of these are considered as high risk factors or are carcinogenic in nature. HPV16 is responsible for various types (more than 90%) of HPV-HNCs. [31]

2.4 EPIDEMIOLOGY OF CERVICAL CANCER

Cervical cancer is the most common cancer cause of mortality among women in developing countries. Mortality that occur due to cervical cancer indicates about the health inequities, about 86% of all deaths, occur due to cervical cancer in developing, low and middle economy countries.

Each year, 122,844 women are diagnosed with cervical cancer and 67,477 die due to cervical cancer in India. In India, 432.2 million women aged 15 years and older, who are at greater risk of developing malignancy. It is the second most common cancer among the women aged between 15–44 years. [32]

RISK FACTORS

Cervical cancer risk is mainly dependent on the following risk factors:

Measures of sexual activity: cervical cancer risk is associated with how sexually active a woman is, depend on number of sexual partners, age at first sexual intercourse and sexual behavior of the opposite partners.

Tobacco: Smoking is one of the major risk factor for cervical cancer. Smoking and cervical cancer risk is directly associated as direct carcinogenic effect of cigarette smoking on the cervix is proved by the presence of nicotine metabolites in the cervical mucus of women who smoke. [33]

Use of Oral contraceptives: Long-term use (12 years or more) of oral contraceptives is associated with the increased risk of cervical cancer. This association is somehow stronger for adenocarcinomas than squamous cell carcinomas.

Diet: The various studies on effect of diet on risk of cervical cancer has shown that a high intake of foods containing beta carotene and vitamin C and to some extend, vitamin A may help in reducing the risk of cervical cancer.[34]

HPV INFECTION AND CERVICAL CANCER

Every year, HPV is responsible for 50,000 new cases of cervical cancer and 250,000 cervical cancer associated mortality worldwide.

HPV infections are asymptomatic in nature, and due to which individuals are generally not aware that whether they are infected or not and thus facilitates the easy spread of the infection.

About 50% of men and women have chances of acquiring HPV infection during their lifetime. Almost every sexually active woman is infected with HPV at least once during her lifetime, and the highest pervasiveness is seen, soon after the onset of sexual activities. [35]

HPV transmission particularly occurs due to direct skin-to-skin contact with an infected partner. Sexual intercourse is not necessary for transmission; the virus can be transmitted through sexual foreplay. HPV can only multiply in stratisfied squamous epithelium. HPV infection is among one of the most common sexually transmitted diseases (STDs). The major risk factor for HPV infection is how sexually active woman is i.e, sexual behavior, involving early age of sexual activity onset, different sexual partners, and co-infection with HIV. Besides the fact that the deciding factors of risk for persistent infection and progression to malignancies are not fully understood, persistence appears to depend on HPV type and concurrent infection with different virus types. [36]

HPV is considered as necessary but is not sufficient as to cause for cervical cancer. Out of more than 100 HPV types only 18 types are considered as high-risk types, while the remaining types are low risk types for cervical cancer.

Various molecular studies have shown that HPV-16 and 18 are the two major highly oncogenic types of HPV found in cervical malignancies, and out of these two HPV-16 is more commonly found in cervical cancer cases. Other

high risk types are not very prevalent. Beside this, genital warts and benign cervical lesions are commonly associated with 11 low risk HPV types. Out of these low risk HPV types, 6 and 11 are responsible for causing 90% of the genital warts. HPV has been detected in women with benign cervical cytology and it is also found in healthy women. [37]

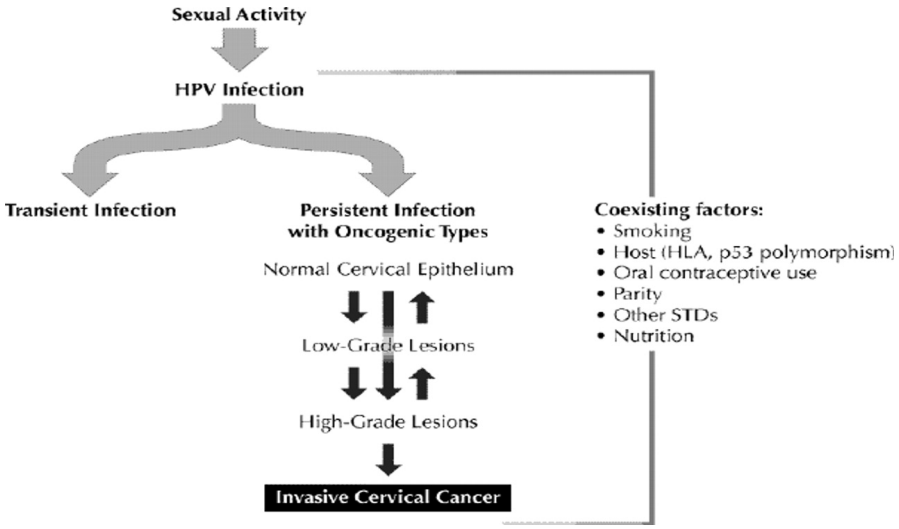

Fig. 2: Etiological model of cervical cancer and human papillomavirus (HPV) infection, By: Lianne Friesen(Edaurdo L. Franco et. al., 2001)

2.5 CELL CYCLE AND CANCER

The process of dividing a cell and replicating DNA can be explained as a series of coordinated events that comprises a ‘‘cell division cycle.’’

There are two types of cell cycle control mechanisms which have been recognized till date, it is a cascade of protein phosphorylations which subsequently pass a cell from one stage to the other and there s a group of checkpoints that monitor the critical events for the completion of cell cycle and it delays the progression for the next stage if it’s necessary.

The control in the first type involves a significantly regulated kinase family[38]. Activation of kinase enzyme requires association with the second subunit which is expressed at the specific period of the cell cycle; the ‘cyclin’ subunit associates with its specific partner ‘cyclin dependent kinase’ (CDK) and later with unique substrate specificity, it creates an active complex.

The second type of regulation in cell cycle checkpoint control is more administrative. The checkpoints, present in cell cycle checks the flaws in critical events such as chromosome segregation and DNA replication. When these checkpoints are activated by damaged DNA, then signals are passed to the cell cycle progression machinery, and further these signals delay in progression of cell cycle, unless the mutation has been arrived and changes the function of cell cycle genes.

There is a obvious connection between the cell cycle and cancer as the cell cycle machinery is responsible for cell proliferation and cancer is a disease of uncontrolled cell proliferation. As all cancers shows the existence of large number of cells. However, the cell number excess is connected with complex cycle with the reduced sensitivity with signals, which normally guides a cell to multiply, adheres, or dies. [39]

The first genetic variation, which has been contributed for the development of cancer, was the result of function mutations in the genes. These mutations are defined as a set of ‘oncogenes’, which are the mutant versions of normal

cellular ‘proto-oncogenes’. However, the transformation can be redundant by the individual oncogenes and the mutation of one of many genes will further lead to transformation or it can be cell specific which means mutations will help in transforming some cells but simultaneously have no effect on the others. Hence, it suggests that the multiple different pathways of genetic changes will lead to cancer but at the same time all pathways have their same role in each cell type. Recently, the loss of the function mutations in producing or causing cancer has become highly apparent.[40]

Mutations in the tumor suppressor genes were initially recognized to have a very important role in susceptibility of the cancer. The inactivation of both the copies of a tumor suppressor gene results in the loss of function. However, one single heterozygous for mutations in tumor supression is more likely to develop cancer as only one mutational event is required to stop the synthesis of any gene product that is functional.

Loss of function of the tumor suppressor gene product e.g pRb is predicted to liberate E2F transcriptional repressors and activators without needing the phosphorylation and finally bypass a normal negative regulation and then controlling entry into the cycle. [41]

The products of cell cycle regulatory genes are very important determinants of progression of all the malignancies.

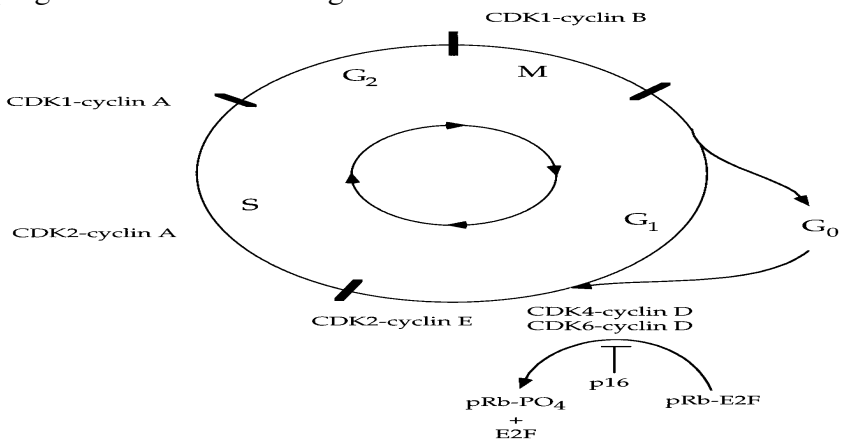

Fig 3: Representation of the mammalian cell cycle. (KATHLEEN COLLINS et. al., 1996)

2.6 E2F FAMILY

The family of transcription factors (E2F) is downstream effectors of (RB) Retinoblastoma protein pathway and also plays a crucial role in controlling the cell division. As it is a discovery, E2F is viewed as a positive and the main regulator of genes, which are required for DNA synthesis.

E2F have many more diverse activities, as it has been found to play significant roles in various different contexts. E2F acts as both an activator and repressor for transcription of genes. It mainly acts as cell proliferation and differentiation factor, as cell death regulator, and in tumor suppression and oncogenesis. [42]

MEMBERS OF E2F FAMILY

In mammals, activity of E2F is produced by a huge number of several interconnected complexes, that are, 7 E2F genes (E2F1– E2F7), 2 DP genes (DP1 and DP2) and 3 genes which encodes RB (retinoblastoma) related proteins (p107 and p130) RB. E2F1, E2F2 and E2F3 mainly E2F3a are major transcriptional activators that interact mainly with pRB and periodically expressed during the cell cycle regulation. E2F4 and E2F5, both are expressed in all of the cell types mostly. They function as repressors as they recruit pocket proteins to E2F regulated promoters, but they are poor transcriptional activators. E2F4 can significantly interact with all the pocket proteins. [43]

Generally, the E2Fs activator promotes progression of the cell cycle, whereas E2Fs, which plays its role as a repressor, are required for cell cycle differentiation, progression and exit. Hence, E2Fs function in opposite ways as a repressor and activator.

E2F3a is an activator of E2F family with regulated cell cycle expression patterns; E2f3b has been expressed as similar to both, E2F4 and E2F5.

E2F/RB is the predominant complex of E2f3b in some of the quiescent cells, present in Q stage of cell cycle and it is believed to act as a transcriptional repressor.

E2F6 and E2F7 have such features, which makes them different from other members of the E2F family. E2F6 and E2F7 lack the Trans activation and binding of pocket protein domains and the repression of transcription are done in protein independent manner.

E2fs, as an activator is required for cell proliferation, whereas E2fs, as a repressor is the main function in cell cycle and differentiation and exit. [44]

2.7 E2F1 GENE

The E2F family plays a significant role in controlling the cell cycle and function of tumor suppressor proteins. E2F1 is a part of (E2F) family of transcription factors, which regulates the expression of important genes, which are involved in regulation of cell cycle and progression as well.

E2F1 is the main regulator of proliferation and apoptosis of the cell, and the deregulated expression of *E2F1* is found in various malignant cancers. Earlier studies identified that 3' untranslated region, (3'UTR) microRNA binding site variants of E2F1 gene are significantly associated with the cancer risk.

The location of *E2F1* gene is chromosome 20 q, which is approximately 10.71 kb and it contains 7 exons .As E2F1 has its important and significant role in cell cycle control, it is expected to be the main player in carcinogenesis (causing cancer).

Large number of E2F1 may enhance proliferation, but at the same time it also proliferates apoptosis, and there have been various examples, which show the over-expression or lack of E2F1gene which has both positive and negative effects on the tumor formation. The main thing is to balance between growth and death of the cell, which depends on the level of E2F1 deregulation, but also on the background of the cell. [45]

E2F1 has its role in oncogenic activities as well as tumor suppressive activities. However, in mouse models there are several examples of both positive and negative effects of deletion and over expression of E2F1 on tumor genesis. Deregulated activity of E2F family is observed in the huge majority of human tumors and it occurs through various different mechanisms. It includes functional loss of retinoblastoma, amplification of *CCND1* that encodes cyclin D1 and also enhances the phosphorylation of retinoblastoma, loss of INK4A, a CDK inhibitor which inhibits the phosphorylation of RB, and expression of the (HPV) Human Papilloma

Virus (E7-onco protein), which disturbs the Rb–E2f complexes."E2F1 and p53 stabilizes in response to different stresses, particularly DNA damage, and these both genes are phosphorylated by ataxia telangiectasia mutated (ATM). Although, CHK1 and CHK2 were shown to enhance E2F1 stabilization and activity of this gene after genotoxic stresses and hence contribute to the up-regulation of p73 and apoptosis induced by E2F1 gene.

E2F1 increases the susceptibility for tumor genesis in various tissues as E2F1 can induce apoptosis that is regulated by p53-dependent and independent pathways .It is known that variants in the promoter regions, miRNA binding regions and coding regions can result in mutated gene functions. However, studies have shown different results of *E2F1* and *E2F2* variants in association with the risk of cancer. [46]

2.8 SINGLE NUCLEOTIDE POLYMORPHISM (SNPs)

SNPs are the variants present in the genome and occur naturally in the human population. SNPs are also pronounced as “snips”. Every individual inherits an allele copy from each parent, which makes genotype of individual at an SNP site AA, BB, or AB. In the Human Genome Project, about 15 million common DNA variants, maximum SNPs are identified.

SNP is a genomic locus in which two or more alternative bases are present with appreciable frequency (>1%). SNPs occurs once every hundred base pair sequences throughout the genome and are the most frequent type of variation in human genome. SNP can occur in a regulatory region where it results in the change in gene expression, or it can also occur between the genes.

Single nucleotide polymorphisms (SNPs) within micro RNA binding sites consists a novel genre of biomarkers of cancer. As regulation of miRNA is dependent on complementarity of the sequences between the mRNA transcript and miRNA, the single nucleotide changes can have significant effects. Many examples from past of these functional miRNA binding site of SNPs have been identified as biomarkers for various cancers.

While many of the research till date focuses on cancer risk association, many more studies are linking these SNPs to cancer prognosis and response to treatment. [47]

SNPs under study:

rs3213150 and rs3213183

Location of gene: 20q11.22

Exons count: 7

1.rs3213150

Table 1: rs3213150 summary

Alleles	C/T
MAF (Minor allele frequency)	A=0.2971/1488 (1000 Genomes) A=0.2790/35038 (TOPMED)
Allele position	20:33684395
Location	Intronic

2. rs3213183

Table 2: rs3213183 summary

Alleles	C/G/T
MAF (Minor Allele frequency)	A=0.2997/1501 (1000 Genomes) A=0.2909/36526(TOPMED)
Chromosomal position	20: 33675156
Location	Promoter

2.9 SNP GENOTYPING: PCR- RFLP

(RFLP) Restriction Fragment Length Polymorphism is homologous DNA sequences of different sizes that can be analyzed after digestion of the DNA samples with specific restriction enzymes. The fragments of different sizes are formed after digestion, which allows the analysis for the genotype.

RFLP is specific to a single clone/restriction enzyme combination, as a molecular marker.

PCR-RFLP is a popular technique for genetic analysis, which is also known as cleaved amplified polymorphic sequence (CAPS). It has been used for the detection of both interspecies and intraspecies variations.

In PCR-RFLP analysis, the first step is to amplify the fragment containing the variation. Then specific restriction enzyme is treated on the amplified DNA fragment. The restriction fragments of different sizes are formed which depends on the restriction enzyme recognition sites. Allelic identification can be done by resolving the fragments in gel electrophoresis. The advantages of PCR-RFLP technique is that it is comparatively less expensiveness and no advanced instruments are required for this.

PCR-RFLP are divided into different steps including primers designing, identification of specific restriction enzyme, amplification, treatment of amplified products with restriction enzyme and gel electrophoresis to resolve the fragments of DNA.

After digesting the amplicons with the selected restriction enzyme, the fragments are resolved by gel electrophoresis. Generally, it is done by slab gel electrophoresis with polyacrylamide gel or agarose gel, which is used as a molecular sieving matrix PCR-RFLP analyses are mostly conducted with unlabelled primers. Visualization of resolved fragments, in that case is done by complexation of DNA fragments with (EtBr) ethidium bromide or any other fluorescent dye during the gel electrophoresis. [48]

CHAPTER 3: OBJECTIVES

1. To optimize the PCR-RFLP conditions for rs3213150 and rs3213183, promoter SNPs of E2F1 gene.

2. To Analyze the association of genotype of rs3213150 and rs3213183 SNPs in cancer patients.

The main objective of this study is to identify the role of regulatory SNPs rs3213150 and rs3213183 of *E2F1* gene in susceptibility of common cancers in Himachal Pradesh population.

If these mutant alleles have higher frequency in cases than in controls then it may have role in carcinogenesis and if high risk in these SNPS is reported then they could be used as a biomarker for cancer patients.

CHAPTER 4: MATERIALS AND METHODS

➢ EXPERIMENTAL PROCEDURES

1. Population study

The study group comprised of 280 cancer patients and 230 healthy individuals that were taken as controls. These controls matched with the patients with respect to age,sex and geographical distribution. All participants signed a written consent before drawing the blood sample and are provided with the knowledge about the project.

2. Sampling

2ml of venus blood was collected in NA_2EDTA coated tubes, from patients and healthy controls. Blood samples were preserved at -20^0C.

3.Isolation of genomic DNA from blood samples

- 300µl blood sample was mixed 1000µl RBC lysis buffer in a 2ml autoclaved micro centrifuge tube. Kept for incubation on rocker at RT for 30-60 minutes until it turned shiny.
- The tube was centrifuged at 13,000 RPM for 1 min to obtain a white pellet.
- The supernatant was discarded and the pellet was thoroughly suspended in 300µl of TE Buffer (pH 8.0) using a vortexing machine. 20µl of 10% SDS solution was added, and the mixture was incubated at 56^0C for 30 mins in water bath.
- 160µl of 7.5M ammonium acetate was added and mixed vigorously by vortexing for 1 min. Centrifuged at 13,000 RPM for 15 mins which resulted in separation of the proteins as a pellet.

- The supernatant was transferred to fresh sterile in 1.5ml micro centrifuge tube. Chilled ethyl alcohol twice the volume of the supernatant was added and the tube was rocked gently for 2-3 times to allow precipitation of gDNA.

- The tube was centrifuged at 13,000 RPM for 10 min to pellet down the gDNA. The pellet was washed in 150µl of 70% ethyl alcohol and air dried at RT for 10-15 min."

- Dried DNA pellet was dissolved in 60µl of TE Buffer (pH 7.3) by incubating at 65^0C for 10 mins. Dissolved DNA was stored at -20^0C for further use.

- The isolated genomic DNA was checked by Agarose gel (0.8%) and quantified using Nano Drop plus Spectrophotometer (GE Healthcare, US). The concentration of DNA was read by measuring the absorbance of a sample at 260 on a spectrophotometer.

4. Primer Designing

PROTOCOL:

Primer designing for rs3213150

Checked for natural restriction site for rs3213150 by using NEB cutter.

TCTTAATGAGTGGGTGGGAGGGGGTTGTACYAGGGGTGGGC
(natural restriction site was not present)

To create restriction site for ACC651 one substitution is made i.e. TGTACY to GGTACT

Primers for rs3213150 are:

Forward Primer:
CCCCCCAAAAAATTAATGAGTGGGTGGGAGGGGGT**G**GTAC

Reverse Primer: GACCCTCAGGTTGTTAGATTTCCGCAG

Tm(^{0}C) : 66^{0}C

Amplicon size (bp) : 189bp

Primer designing for rs3213183

Checked for natural restriction site for rs3213183 by using NEB cutter.

AATAAGTAGGAGTTAACCAGAGGAAGAAAAYACCAAGCAGA

(natural restriction site was not present)

To create the restriction site for Rsa I we have done two substitutions in sequence AGGAAGAAAAYAC to form AGGAAGAAAGTAC.

Primers for rs3213183 are:

Forward primer:

CCCCCCCCAAAAAAAAGGAGTTAACCAGAGGAAGAA**A**A**G**

Reverse primer: TTTTAAGATGCGAAAGAGGAGG

Tm (^{0}C) : 54^{0}C

Amplicon size (bp) : 201bp

Reconstitution of primers : For 100µM, following amount of nuclease free water is added.

Table 3: Amount of distilled water required for primer reconstitution

Primers	NFW (µl)
rs3213150F	171
rs3213150R	263
rs3213183F	183
rs3213183R	170

PCR OPTIMIZATION

Gradient PCR

Gradient PCR was done to find the best annealing temperature for both sets of the primers for the successive amplification of the specific product.

PCR (Polymerase chain reaction) :

PCR amplification of the desired gene product was performed using specific primers. PCR was performed using following protocol:

The reaction of 12 µl was prepared according to the composition given in the following table 4:

Table 4: PCR components used for DNA amplification

Reaction components	Per reaction volume (µl)	
	For rs3213150	For rs3213183
Master mix	5.4µl	5.4µl
DNA Template	1µl	1µl
Forward Primer	0.2µl	0.2µl
Reverse Primer	0.2µl	0.2µl
Nuclease free water	5.2µl	5.2µl
Total	12µl	12µl

Table 5: PCR cycling conditions : 5 (a) For rs3213150

Steps	Temperature	Time	Cycle/s
Initial Denaturation	94^0C	3 min	1
Denaturation	94^0C	35 seconds	35
Annealing	66^0C	35 seconds	
Extension	72^0C	1 min	
Final Extension	72^0C	5 min	1
Hold	4^0C	∞	-

5(b):For rs3213183

Steps	Temperature	Time	Cycle/s
Initial Denaturation	94^0C	3 min	1
Denaturation	94^0C	30seconds	30
Annealing	54^0C	30seconds	
Extension	72^0C	1 min	
Final Extension	72^0C	5 min	1
Hold	4^0C	∞	-

GENOTYPING PCR-RFLP :

Allelic specific product was identified using RFLP. Restriction enzyme was added to amplified product to find the genotype present in the individuals. PCR product for rs3213150 was digested with restriction enzyme ACC651 and PCR product for rs3213183 was digested with restriction enzyme RSAI.

Table 6 : Reaction conditions for RFLP :

Reaction components	Reaction volume	
	rs3213150	rs3213183
Restriction enzyme	ACC65I	Rsa I
Restriction site of enzyme	5'…GGTACC…3'	5'…GTAC…3'
	3'…CCATGG…5'	3'…CATG…5'
Enzyme (1 unit)	0.1µl	0.05µl
Buffer	1.5	1.5
Water	3.4	3.4
PCR Product	10	10
TOTAL	15	14.95

Analysis of Digested PCR Product:

- Agarose gel electrophoresis was performed to visualize the banding pattern.
- These digested RFLP products were analyzed on agarose gel containing EtBr.
- A 100bp Marker was used as a ladder.
- 14µl of amplified and restriction enzyme treated product was loaded into the wells.
- Then the gel was run at 100volts in 1X TAE buffer for 40-50 mins.

- Bands were visualized using Gel Doc.

Table 7: Digested band pattern of rs3213150 and rs3213183

rs3213150		**rs3213183**	
Genotype	**Band size (bp)**	**Genotype**	**Band Size (bp)**
CC (WW)	40+149	CC (WW)	201
CT (WM)	40 + 149 + 189	CT (WM)	55 + 146 + 201
TT (MM)	189	TT (MM)	55 + 146

Statistical Analysis

Statistical tools were applied to assess the association of E2F1 gene polymorphism with cancer susceptibility.

A chi-square test was used to find out the control genotype frequency using HWE calculator.

If p-value < 0.05, then frequency is considered to be in dis equilibrium. To assess the risk association with variant alleles, the ODD ratio and 95% CI (confidence Intervals) were calculated.

Odd Ratio was calculated by using MedCalc Software and the significance of the Odd ratio is determined by Z test.

CHAPTER 5: OBSERVATION AND RESULTS

- **GRADIENT PCR GEL RESULTS**

A) rs3213150

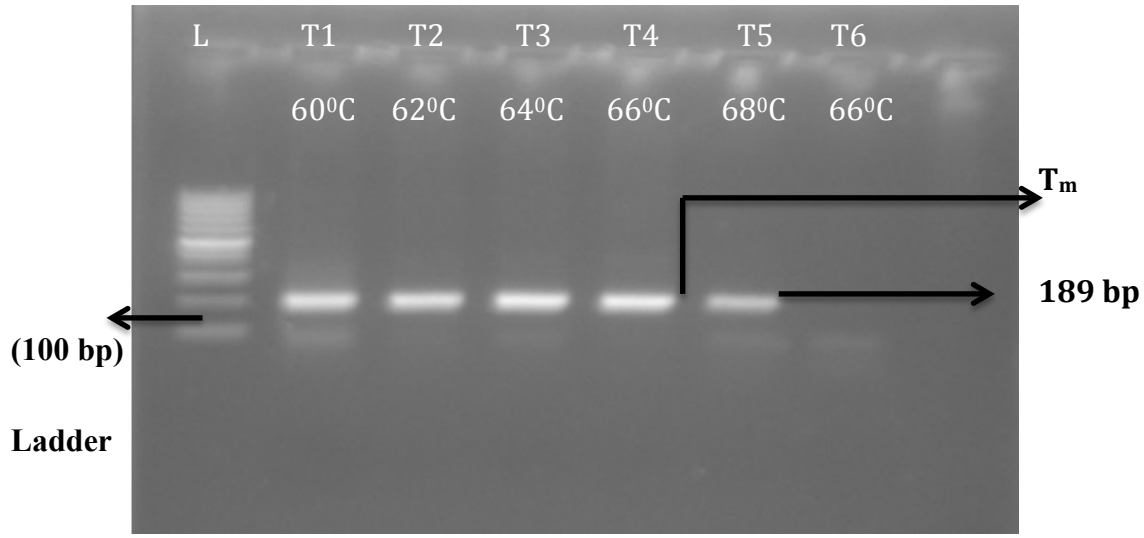

Fig 4: a) Agarose gel electrophoresis (1.5% w/v) showing the amplification of the DNA sample at different annealing temperatures i.e. 60^0C- 68^0C. (Expected band size is 189 bp)

Lane 1: Ladder (100 bp); **Lane 2**: Sample at 60^0C; **Lane 3**: Sample at 62^0C; **Lane 4**: Sample at 64^0C; **Lane 5**: Sample at 66^0C; **Lane 6:** Sample at 68^0C; **Lane 7**: Negative control at 66^0C.

B. **rs3213183**

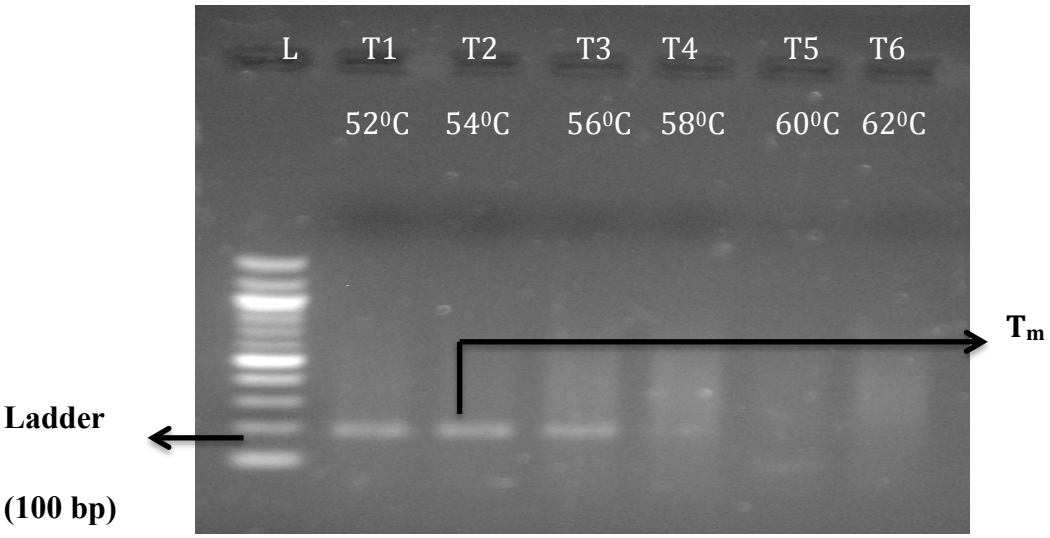

Fig4 : b) Agarose gel electrophoresis (1.5% w/v) showing the amplification of the DNA sample at different annealing temperatures i.e. 52^0C- 62^0C. (Expected band size is 201bp)

Lane 1: Ladder (100 bp); **Lane 2**: Sample at 52^0C; **Lane 3**: Sample at 54^0C; **Lane 4**: Sample at 56^0C; **Lane 5**: Sample at 58^0C; **Lane 6:** Sample at 60^0C; **Lane 7**: Negative control at 62^0C.

➢ **RESTRICTION DIGESTION GEL RESULTS**

A) rs3213150

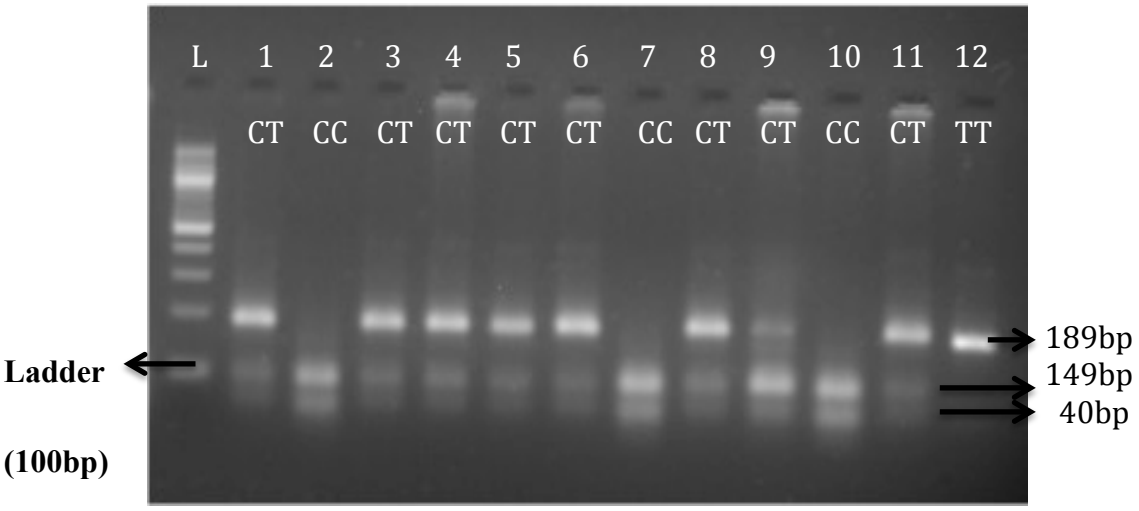

Fig 5: a) Agarose gel electrophoresis (3% w/v) showing digestion of 12 Samples which was amplified at annealing temperature 66^0 C. (Expected genotype = CC: 2 bands (149bp +40 bp); CT : 3 bands (189bp + 149bp +40 bp); TT : 1 band (189bp)).

Lane 1: Ladder (100 bp); **Lane 2**: Sample 1; **Lane 3**: Sample 2; **Lane 4**: Sample 3; **Lane 5**: Sample 4; **Lane 6:** Sample 5; **Lane 7**: Sample 6; **Lane 8** : Sample 7; **Lane 9**: Sample 8; **Lane 10**: Sample 9; **Lane 11**: Sample 10; **Lane 12** : Sample 11; **Lane 13:** Sample 12.

B) rs3213183

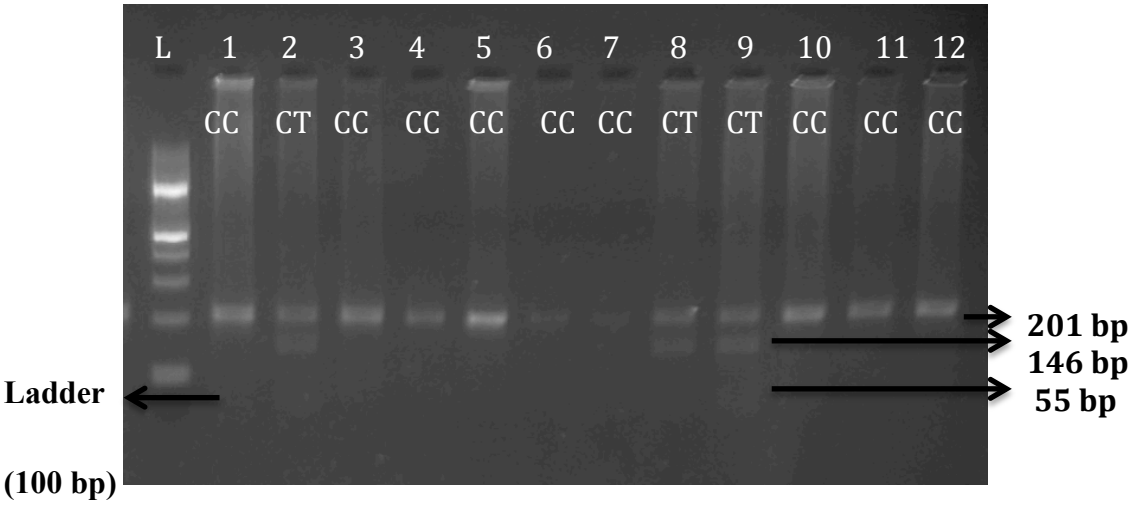

Fig 5: b) Agarose gel electrophoresis (3% w/v) showing digestion of 12 Samples which was amplified at annealing temperature 54^0 C. (Expected genotype = CC: 1 bands (201 bp); CT : 3 bands (201bp + 146bp +55 bp); TT : 2 band (146 bp + 55 bp)).

Lane 1: Ladder (100 bp); **Lane 2**: Sample 1; **Lane 3**: Sample 2; **Lane 4**: Sample 3; **Lane 5**: Sample 4; **Lane 6:** Sample 5; **Lane 7**: Sample 6; **Lane 8** : Sample 7; **Lane 9**: Sample 8; **Lane 10**: Sample 9; **Lane 11**: Sample 10; **Lane 12** : Sample 11; **Lane 13:** Sample 12.

RESULT ANALYSIS

rs3213150

Statistical analysis of Intronic variant rs3213150 (C/T) revealed the association of this variant with the lung cancer risk in different genetic models. This genetic variant of E2F1 gene was significantly associated with lung cancer in heterozygous model,(OR= 5.0510, 95% CI= 2.784-9.162, P-value = <0.0001); in Dominant model ,(OR= 4.8971, 95% CI= 2.704-8.868, P-value = <0.0001); in Allelic Model,(OR= 2.0028, 95% CI= 1.414-2.836, P-value = 0.0001).

Protective association of variant rs3213150 (C/T) with Head and Neck cancer was found in different genetic models that are, heterozygous model, (OR= 0.5637, 95% CI=0.341-0.931, P-value = 0.0253); in Dominant model,(OR= 0.5400, 95% CI= 0.327- 0.891, P-value = 0.0159); in Allelic Model,(OR= 0.6083, 95% CI= 1.414-2.836, P-value = 0.0235).

No significant association was observed in any genetic model i.e., heterozygous model, (OR= 0.6973, 95% CI=0.386-1.258, P-value = 0.2310); in Dominant model,(OR= 0.6861, 95% CI= 0.382- 1.232, P-value = 0.2072); in Allelic Model,(OR= 0.7609, 95% CI= 0.474-1.219, P-value = 0.2564). Thus, this variant is not significantly associated with cervical cancer.

Table 8: Genotypic analysis of rs3213150 for Lung Cancer

Variant	Genotype	LC Cases (100)	Control (280)	Odd Ratio (95% CI)	P value
rs3213150					
	CC	16	111	Ref	
	CT	83	114	5.051(2.784-9.162)	<0.0001
	TT	1	5	1.387(0.152-12.64)	0.7715
	CT+TT	84	119	4.897(2.704-8.868)	<0.0001
	C	115	336	Ref	
	T	85	124	2.002(1.414-2.836)	0.0001

Table 9: Genotypic analysis of rs3213150 for Head and Neck cancer (HNC)

Variant	Genotype	HNC Cases (90)	Control (280)	Odd Ratio (95% CI)	P- Value
rs3213150					
	CC	57	111	Ref	
	CT	33	114	0.563(0.341-0.931)	0.0253
	TT	0	5	0.1763(0.009-3.244)	0.2428
	CT+TT	33	119	0.540(0.327-0.891)	0.0159
	C	147	336	Ref	
	T	33	124	0.608(0.395-0.935)	0.0235

Table 10: Genotypic analysis of rs3213150 for Cervical cancer (CX)

Variant	Genotype	CX Cases (90)	Control (280)	Odd Ratio (95% CI)	P- Value
rs3213150					
	CC	49	41	Ref	
	CT	40	48	0.697(0.386-1.258)	0.2310
	TT	1	2	0.4184(0.036-4.781)	0.4832
	CT+TT	41	50	0.686(0.382-1.232)	0.2072
	C	138	130	Ref	
	T	42	52	0.7609(0.474-1.219)	0.2564

rs3213183

Statistical analysis of the promoter variant rs3213183 (C/T) revealed no significant association of this variant with the lung cancer risk in different genetic models i.e., heterozygous model,(OR= 1.769, 95% CI= 0.964-3.247, P-value = 0.0653); in Dominant model ,(OR= 1.509, 95% CI= 0.843-2.699, P-value = 0.1653); in Allelic Model,(OR= 1.257, 95% CI= 0.743-2.127, P-value = 0.3930).

Also, no association of this variant rs3213183 (C/T) with Head and Neck cancer was found in different genetic models that are, heterozygous model,

(OR= 1.339, 95% CI=0.692-2.591, P-value = 0.3860); in Dominant model, (OR= 1.092, 95% CI= 0.574- 2.077, P-value = 0.7874); in Allelic Model,(OR= 0.8997, 95% CI= 0.494-1.636, P-value = 0.7293).

A significant association of variant rs3213183 was observed with cervical cancer in different genetic models. This genetic variant of E2F1 gene was significantly associated with cervical cancer in heterozygous model, (OR= 4.183, 95% CI=1.683-10.39, P-value = 0.0021); in Dominant model,(OR= 3.850, 95% CI= 1.644-9.014, P-value = 0.0019); in Allelic Model,(OR= 3.095, 95% CI= 1.468-6.524, P-value = 0.0030).

Table 11: Genotypic analysis of rs3213183 for Lung cancer (LC)

Variant	Genotype	LC Cases (100)	Control (230)	Odd Ratio (95% CI)	P- Value
rs3213183					
	CC	77	192	Ref	
	CT	22	31	1.769(0.964-3.247)	0.0653
	TT	1	7	0.356(0.043-2.943)	0.3381
	CT +TT	23	38	1.509(0.843-2.699)	0.1653
	C	176	415	Ref	
	T	24	45	1.257(0.743-2.127)	0.3930

Table 12: Genotypic analysis of rs3213183 for Head and Neck cancer (HNC)

Variant	Genotype	HNC Cases (90)	Control (230)	Odd Ratio (95% CI)	P-Value
rs3213183					
	CC	74	192	Ref	
	CT	16	31	1.339(0.692-2.591)	0.3860
	TT	0	7	0.172(0.009-3.054)	0.2306
	CT +TT	16	38	1.092(0.574-2.077)	0.7874
	C	164	415	Ref	
	T	16	45	0.899(0.494-1.636)	0.7293

Table 13: Genotypic analysis of rs3213183 for Cervical cancer (CX)

Variant	Genotype	CX Cases (52)	Control (91)	Odd Ratio (95% CI)	P-Value
rs3213183					
	CC	34	80	Ref	
	CT	16	9	4.183(1.683-10.39)	0.0021
	TT	2	2	2.352(0.318-17.39)	0.4019
	CT +TT	18	11	3.850(1.644-9.014)	0.0019
	C	84	169	Ref	
	T	20	13	3.095(1.468-6.524)	0.0030

CHAPTER 6: DISCUSSION

Cellular proliferations are controlled by cell cycle regulatory genes and the mutations in these genes are associated with cancer. The E2F family is the transcription factors that are involved in cell growth and differentiation; they regulate the expression of many genes. E2F1 activates the DNA repair pathway. E2F1 up regulates many genes involved in cell cycle and is involved in various other activities like DNA synthesis and replication, checkpoint control, DNA damage and DNA repair, apoptosis, self-renewal, differentiation and development.

E2F1 is found to be de-regulated in many types of cancers like hepatocellular carcinoma, non-small-cell lung carcinoma, cervical cancer, glioblastoma, pancreatic cancer, renal, breast, and ovarian cancer. In our study we have focused on three main common cancers in Himachal Pradesh population that are lung cancer, cervical cancer and head and neck cancer.

E2F1 is epigenetically regulated through methylation, phosphorylation and acetylation. E2F1 has been associated with oncogenic functions in melanoma, and it is more complex that it can promote or inhibit biological functions linked with metastatic phenotype. The important role of E2F1 is in regulating disease progression and drug resistance as it is related to malignant melanoma. With the known interactions between E2F1 and various signaling pathways, it is seen that the ability of E2F1 to control a multitude of biological processes makes its role in cancer.

Two variants of E2F1 gene that are rs3213150 andrs3213183 are investigated to check the susceptibility of these variants with common cancers in Himachal Pradesh population. PCR-RFLP technique was used for genotyping the cases and the controls. Statistical methods were applied to check the association of these SNPs with the susceptibility of the cancer.

In case of SNP rs3213150, statistical analyses revealed the association of this variant with lung cancer risk as in heterozygous model, it increases the risk of 5 folds in this population; in Dominant model it increases the risk to 4 folds; in Allelic Model the risk is increased to 2 folds.

Although, in case of SNP rs3213183, statistical analysis revealed the association of this variant with cervical cancer risk as in heterozygous model; it increases the risk to 4 folds; in Dominant model and allelic model, the risk is increased to 3 folds.

CHAPTER 7 : CONCLUSION AND FUTURE PROSPECTS

SNPs rs3213150 and rs3213183 of E2F1 gene were investigated for showing any association with the cancer risk in three common cancers that are lung cancer, head and neck cancer and cervical cancer in Himachal Pradesh population. The genotypes of both the SNPs were analyzed, which results in concluded the association of variant rs3213150 with the lung cancer risk and association of variant rs3213183 with cervical cancer risk in Himachal Pradesh population. There is no association of variant rs3213150 with cervical cancer risk but protective association of this variant is observed with head and neck cancer risk. "Although, no association is found in variant rs3213183 with lung cancer and head and neck cancer risk in the same population. This data could be further utilized for studying the etiological factors of the cancer by genotyping more number of cancer samples in different population sets.

APPENDIX

➢ GLASSWARES AND INSTRUMENTS

Glassware's

- Beaker – 1000ml,500ml,100ml
- Eppendorfs (autoclaved)-2ml.1.5ml,0.5ml,0.2ml
- Autoclaved micro tips (100-1000ul,20-200ul,0.1-10ul)
- PCR Tube stand
- Capped Bottles
- Measuring cylinder – 500ml,100ml,10ml
- Eppendorfs stand

INSTRUMENTS

- PCR Tube stand
- Micro pipette
- Thermo-cycler
- Spinner
- Weighing balance
- Autoclave
- Incubator
- Laminar Air Flow
- Hot air oven
- pH meter
- Rocker
- Refrigerator (-80ºC,-20ºC,-4ºC)
- Vortex
- Centrifuge
- Water bath
- Agarose Gel Electrophoresis chamber

- UV trans illuminator

➢ REAGENTS

❖ **Di-sodium ethylene diamine tetra acetate (0.5M,Ph 8.0)**

- Take 186.1g of Na_2EDTA and add to it 800ml of milliQ water.
- Stir vigorously on magnetic stirrer.
- Adjust ph to 8.0 with 10M NaOH.
- Make up the volume to 1000ml.

❖ **Tris (hydroxymethyl) aminomethane-chloride, Tris-Cl (1M,pH8.0)**

- Take 121.2g Tris Base in 800ml of distilled water.
- pH 8 was adjusted by adding 1N HCl.
- Make up the volume to 1L by adding MQ water.

❖ **Tris-Cl (1M:pH 7.3)**

- In 800ml of distilled water add 121.2g of Tris base.
- ph 7.3 was adjusted by adding 1N HCl.
- Make up the volume to 1L."

❖ **Ammonium Chloride, NH_4Cl (1M)**

- In 800 MQ water dissolve 53.5g of ammonium chloride .
- Make up the volume to 1L.

❖ **10% SDS**

- In 70ml of distilled water dissolve 10g of SDS.
- Heat to 68*C to mix the solution.
- Make up the volume to 1L.

❖ **Red Blood Cell Lysis Buffer**

- **Composition :**Tris 10mM , ph – 8.0;EDTA 1mM ; NH_4Cl 125mM, ph 8.0
  i. EDTA (0.5M) 2ml

ii. Tris (1M,ph-8.0) 10ml

iii. NH_4Cl(1M) 125ml

Mix the above reagent in MQ water to obtain final volume of 1L.

❖ **Tris – EDTA (TE) buffer (pH 8.0)**

- **Composition:** Tris 10Mm;EDTA 1mM, pH 8.0

i. EDTA (0.5M) 2ml

ii. Tris (1M , pH8.0) 10ML

Mix the above reagent in distilled water to obtain absolute volume of 1L

❖ **Tris – EDTA (TE) buffer (pH7.3)**

- **Composition: Tris 10Mm;EDTA 1Mm,pH7.3**

i. EDTA(0.5M) 2ml

ii. Tris (1M, pH7.3) 10ml

Mix the above reagents in distilled water to obtain absolute volume 1L.

❖ **Ammonium Acetate (7.5M)**

- In 20ml of MQ water dissolve 28.9g of ammonium acetate salt.
- Final volume was adjusted to 50ml.

REFERENCES

1. G.M. Cooper, "*The Cell: A Molecular Approach*". 2nd edition, Boston University, Sunderland (MA):Sinauer Associates;(2000).ISBN-10:0-87893-106-6.
2. McDaniel JT, Nuhu K, Ruiz J, Alorbi G; Global health association, "*Social determinants of cancer incidence and mortality around the world: an ecological study*",(2017), doi: 10.1177/1757975916686913.
3. Hudson K, Javitt G, Burke W, et al. ASHG statement on direct-to-consumer genetic testing in the United States. *American Journal of Human Genetics* 2007; 81(3):635–637.
4. Reece, J. B., Urry, L. A., Cain, M. L., Wasserman, S. A., Minorsky, P. V., and Jackson, R. B. "*The cell cycle*", In *Campbell biology* (10th ed.),(2011), San Francisco, CA: Pearson, 247
5. Gupta N, Gupta R, Acharya AK, Patthi B, Goud V, Reddy S, Garg A, Singla A, "*Nepal J. Epidemio. Changing Trends in oral cancer - a global scenario"*, (2016); 6(4): 613-619.
6. Meixia Lu, Zhensheng Liu, Hongping Yu, Li-E Wang, Guojun Li, Erich M. Sturgis, David G. Johnson, Qingyi Wei, " *Combined effects of E2F1 and E2F2 polymorphisms on risk and early onset of squamous cell carcinoma of the head and neck*",Mol Carcinog;(2012); 51(Suppl 1): E132–E141.
7. Yunzhao Zhao, Lin Tang, Weiwei Nie, Zexing Wang, Xiaoxiang Guan, "*Functional variants at the miRNA binding sites of the E2F1 gene and its mRNA expression*",Oncol Lett., (2013); 5(1): 398–402.
8. Uthman OA, Jadidi E, Moradi T. "*Socioeconomic position and incidence of gastric cancer: a systematic review and meta-analysis*", J Epidemiol Community Health; (2013); 67(10): 854–860.
9. G.M. Cooper, "*The Cell: A Molecular Approach*". 2nd edition, Boston University, Sunderland (MA):Sinauer Associates;(2000).ISBN-10:0-87893-106-6.

10. Bouwman P, Jonkers J. "*The effects of deregulated DNA damage signalling on cancer chemotherapy response and resistance*", Nat Rev Cancer, (2012); 12(9):587–598.

11. Carlo M. Croce, M.D., "*Oncogenes and Cancer*", N Engl J Med, (2008); 358:502-511
12. Clapp RW1, Jacobs MM, Loechler EL; "*Environmental and occupational causes of cancer: new evidence 2005-2007";* Rev Environ Health. (2008) ;23(1):1-37.
13. Clarke Brian Blackadar; "Historical review of the causes of cancer"; World J Clin Oncol*,* (2016); 7(1): 54-86.
14. Sunny Y. Auyang,"*Cancer causes and cancer research on many levels of complexity*",(2006).
15. Bruce N. Ames, Lois Swirsky Gold t, and Walter C. Wilet, "*The causes and prevention of cancer*", (1995),Vol. 92, pp. 5258-5265.
16. A. J. Phillips., "*The Epidemiology of Cancer*", (1969);15(12):4447.
17. Ferlay J, Shin H R, Bray F et al. GLOBOCAN 2008, "*Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 Lyon*", France: International Agency for Research on Cancer; (2010).
18. Carole A. Ridge, Aoife M. McErlean, and Michelle S. Ginsberg, "*Epidemiology of Lung Cancer*", MD1, (2013); 30(2): 93–98.
19. Prabhat Singh Malik and Vinod Raina, "*Lung cancer: Prevalent trends & emerging concepts*", (2015); 141(1): 5–7.
20. Pfeifer G P, Denissenko M F, Olivier M, Tretyakova N, Hecht S S, Hainaut P; "*Tobacco smoke carcinogens, DNA damage and p53 mutations in smoking associated cancer*". (2002);21(48):7435–7451.
21. Pershagen G, Akerblom G, Axelson O, Clavensjö B, Damber L, Desai G, Enflo A, Lagarde F, Mellander H, Svartengren MN*, " Residential radon exposure and lung cancer in Sweden"*. Engl J Med, (1994); 330(3):159-64.
22. Johnson B E., "*Second lung cancers in patients after treatment for an initial lung cancer*". J Natl Cancer Inst., (1998); 90(18):1335–1345.

23. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. "Cancer statistics",*CA Cancer J Clin*.,(2017);57(1):43-66.

24. Boyle P and Bernard Levin. "*World Cancer Report 2008"*. International Agency for Research on Cancer.

25. Manik Rao Kulkarni , "*Head and Neck Cancer Burden in India*", International Journal of Head and Neck Surgery, (2013);4(1):29-35
26. Eleni M. Rettig, MD , Gypsyamber D'Souza, PhD, MPH, "*Epidemiology of Head and Neck Cancer*", Surg Onco Clin, (2015); 24(3): 379-96.
27. Blot WJ, McLaughlin JK, Winn DM, "*Smoking and drinking in relation to oral and pharyngeal cancer*", Cancer Res, (1988);48(11):3282–7.
28. Hashibe M, Brennan P, Chuang SC, "*Interaction between tobacco and alcohol use and the risk of head and neck cancer: pooled analysis in the Inter- national Head and Neck Cancer Epidemiology Consortium*", Cancer Epidemiol Biomarkers Prev, (2009); 18(2): 541–50.
29. Grulich AE, van Leeuwen MT, Falster MO, "*Incidence of cancers in people with HIV/AIDS compared with immunosuppressed transplant recipients: a meta- analysis*", Lancet (2007); 370(9581): 59–67.
30. Long JL, Engels EA, Moore RD, "*Incidence and outcomes of malignancy in the HAART era in an urban cohort of HIV-infected individuals*", AIDS (2008); 22(4): 489–96.
31. Gillison ML, Alemany L, Snijders PJ, "*Human papillomavirus and diseases of the upper airway: head and neck cancer and respiratory papillomatosis*", Vaccine (2012); 30(Suppl 5): F34–54
32. ICO Information Centre on HPV and cancer. "*Human Papillomavirus and Related Diseases in India"* (Summary Report 2014-08-22); (2014).

33. Schiffman MH, Brinton LA, Devesa SS, Fraumeni JF Jr. Cervical cancer. In: Schottenfeld D, Fraumeni JF Jr, editors. "*Cancer epidemiology and prevention*". New York: Oxford University Press; (1996); p. 1090-116.
34. Moscicki AB., "*Impact of HPV infection in adolescent populations*", *J Adolesc Health*, (2005); 37(6 Suppl): S3–S9.
35. Brown DR, Shew ML, Qadari B, "*A longitudinal study of genital pappilomavirus infection in a cohort of closely followed adolescent women*", J Infect Dis, (2006); 191:182–192.
36. Sowjanya AP, Jain M, Poli UR, "*Prevalence and distribution of high-risk human papilloma virus (HPV) types in invasive squamous cell carcinoma of the cervix and in normal women in Andhra Pradesh, India*". BMC Infect Dis. (2005); 5:116.
37. Disease Specific Documents for XII plan. "*Human Papilloma Virus ICMR: High power Committee to Evaluate Performance of ICMR, 2012–2013"*. New Delhi, India: ICMR; (2014).
38. Morgan, D. O., "*Principles of CDK regulation"***,** Nature (London), (1995), **374,** 131–134
39. Aaronson, S. A. "*Expression of multi-cytokine resistance and multi-growth factor independence in advanced stage metastatic cancer. Malignant melanoma as a paradigm*", Science (1991) 254**,** 1146–1153.
40. Weinberg, R. A. , "*Tumor suppressor genes*", (1991) *Science* 254**,** 1138–1146.
41. Jacks, T. & Weinberg, R. A., "*Cell-cycle control and its watchman*", Nature (London), (1996), 381, 643– 644.
42. Desssislava K Dimova and Nicholas J Dyson , "*The E2F transcriptional network: old acquaintances with new faces*", Oncogene, (2005), 24, 2810–2826.
43. He S, Cook BL, Deverman BE, Weihe U, Zhang F, Prachand V, Zheng J and Weintraub SJ, "*E2F is required to prevent inappropriate S-phase entry of mammalian cells*." Mol. Cell. Biol.,(2000), 20, 363–371.

44. Leone G, Nuckolls F, Ishida S, Adams M, Sears R, Jakoi L, Miron A and Nevins JR. "*Identification of a novel E2F3 product suggests a mechanism for determining specificity of repression by Rb protein*", Mol. Cell. Biol.,(2000), 20, 3626–3632.
45. Yunzhao zhao , Lin Tang, Weiwei Nie, Zexing Wang and Xiaoxiang Guan, "*Functional variants at the miRNA binding sites of the E2F1 gene and its mRNA expression*", oncology letters ,(2013),5: 398-402.
46. Meixia Lu, Zhensheng Liu, Hongping Yu, Li-E Wang, Guojun Li, Erich M. Sturgis, David G. Johnson, Qingyi Wei, " *Combined effects of E2F1 and E2F2 polymorphisms on risk and early onset of squamous cell carcinoma of the head and neck*",Mol Carcinog;(2012); 51(Suppl 1): E132–E141.
47. Carina Preskill and Joanne B. Weidhaas, "*SNPs in microRNA Binding Sites as Prognostic and Predictive Cancer Biomarkers*", Crit Rev Oncog. (2013); 18(4): 327–340.
48. Henrik Berg Rasmussen, " *Restriction Fragment Length Polymorphism Analysis of PCR-Amplified Fragments (PCR-RFLP) and Gel Electrophoresis – Valuable Tool for Genotyping and Genetic Fingerprinting*", Denmark,(2012).