1.1 Cancer

Cancer, also known as a malignant tumor or malignant neoplasm originates from a number of molecular changes that alter the normal properties of cells (Peter and Bernard, 2008). That is cancer cell develops from changes that causes normal cells to acquire abnormal functions. Cancer cells do not required signals to induce cell growth and division. As they grow cancer cells develop new characteristics, including changes in cell structure, lack contact inhibition, and production of new enzymes. Loss of contact inhibition accounts for other characteristics of cancer cells including invasiveness of surrounding tissues, and metastasis, or spreading via the lymph system or blood to other tissues and organs (Hanahan and Weinberg, 2000). Such heritable changes allow the cancer cell and its progeny to divide, grow and invade other tissues, even in the presence of normal cells that typically inhibit the growth of nearby cells. The abnormalities in cancer cells usually result from mutations occurring in genes encoding regulatory proteins for cell division. Over time more genes become mutated as the repair genes that encodes for proteins monitoring DNA damages are themselves not functioning normally because they are also mutated (Croce, 2008). Consequently, mutations begin to increase in the cell, causing further abnormalities in that cell and the daughter cells. Some of these mutated cells die, but other alterations may give the abnormal cell a selective advantage that allows it to multiply much more rapidly than the normal cells. This enhanced growth describes most cancer cells, which have gained functions repressed in the normal, healthy cells. As long as the cancer cells remain in their original location, they are considered benign; if they become invasive, they are considered malignant. Cancer cells in malignant tumors can often metastasize, sending cancer cells to distant sites in the body where new tumors may form.

Cancer cells break free from a tumor and travel to and invade other tissues in the body through the process of Metastasis. Cancer cells metastasize to other sites via the lymphatic system and the bloodstream. Cancer cells from the original or primary tumor can travel to other sites such as the lungs, bones, liver, brain, and other areas. These metastatic tumors are "secondary cancers" because they arise from the primary tumor (Chiang and Massague, 2008). Metastatic cancer retains the name of the primary cancer (Figure 1.1). For example, bladder cancer that metastasizes to the liver is not liver cancer. It is called metastatic bladder cancer.

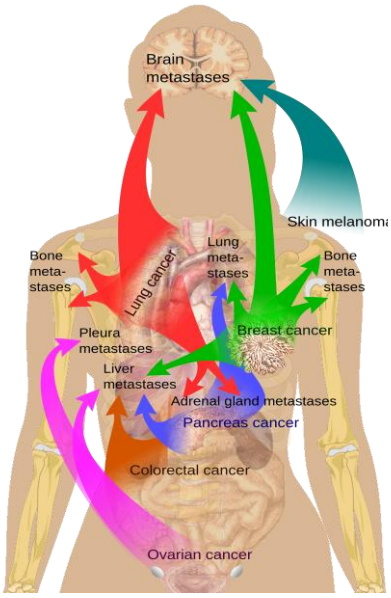

Figure 1.1 Main sites of metastases for some common cancer types. Primary cancers are denoted by "...*cancer*" and their main metastasis sites are denoted by "...*metastases*"

Metastasis is significant because it helps determine the staging and treatment of cancer. Some types of metastatic cancer are curable, but many are not.

1.1.1 Classification of cancer

The type of tumor that forms depends on the type of cell that was initially altered. Accordingly, tumors are group into the following three types:

- **Carcinomas** result from altered epithelial cells, which cover the surface of our skin and internal organs. Nearly, 90% of cancers are carcinomas. Carcinomas can be subdivided into two types: Adenocarcinomas and squamous cell carcinomas. Adenocarcinomas are cancers that develop in an organ or a gland, while squamous cell carcinomas refer to cancers that originate in the skin.

- **Sarcomas** are rare in humans. They are solid tumors of connective tissues, such as muscle, bone, cartilage, and fibrous tissue.

- **Leukemia** arises from blood-forming cells, while **lymphoma** is a cancer of the cells of the immune system. These cancers accounts for approximately 7% of human malignancies. Tumors are further classified according to tissue of origin (e.g., lung or breast carcinomas) and the type of cell involved (e.g., fibrosarcomas from fibroblasts, and erythroid leukemias from precursors of erythrocytes or red blood cells).

1.1.2 Global Cancer Scenario

According to GLOBOCAN 2012 report, there were 14.1 million new cancer cases, 8.2 million cancer deaths and 32.6 million people living with cancer (within 5 years of diagnosis) in 2012 worldwide. 57% (8 million) of new cancer cases, 65% (5.3 million) of the cancer deaths and 48% (15.6 million) of the 5-year prevalent cancer cases occurred in the less developed regions (Table 1.1).

Table 1.1 Cancer Incidence, mortality and prevalence in the world.

Estimated numbers (thousands)	**Men**			**Women**		
	Cases	Deaths	5-year prev.	Cases	Deaths	5-year prev.
World	7427	4653	15362	6663	3548	17182
More developed regions	3244	1592	8616	2832	1287	8297
Less developed regions	4184	3062	6747	3831	2261	8885
WHO Africa region	265	205	468	381	250	895
WHO Americas region	1454	677	3843	1429	618	4115
WHO East Mediterranean region	263	191	461	293	176	733
WHO Europe region	1987	1081	4857	1750	852	4933
WHO South-East Asia region	816	616	1237	908	555	2041
WHO Western Pacific region	2642	1882	4493	1902	1096	4464
IARC membership (24 countries)	3706	1900	9259	3354	1570	9425
United States of America	825	324	2402	779	293	2373
China	1823	1429	2496	1243	776	2549
India	477	357	665	537	326	1126
European Union (EU-28)	1446	716	3759	1211	561	3487

Adapted from GLOBOCAN 2012 (*http://www.globocan.iarc.fr/*)

The overall age standardized cancer incidence rate is almost 25% higher in men than in women, with rates of 205 and 165 per 100,000, respectively. Male incidence rates vary almost five-fold across the different regions of the world, with rates ranging from 79 per 100,000 in Western Africa to 365 per 100,000 in Australia/New Zealand (with high rates of prostate cancer representing a significant driver of the latter). There is less variation in female incidence rates (almost three-fold) with rates ranging from 103 per 100,000 in South-Central Asia to 295 per 100,000 in Northern America. In terms of mortality, there is less regional variability than for incidence, the rates being 15% higher in more developed than in less developed regions in men, and 8% higher in women. In men, the rates are highest in Central and Eastern Europe (173 per 100,000) and lowest in Western Africa (69). In contrast, the highest rates in women are in Melanesia (119) and Eastern Africa (111), and the lowest in Central America (72) and South-Central (65) Asia.

1.2 Nasopharyngeal Cancer

The nasopharynx is made up of the upper part of the pharynx (throat) behind the nose. The pharynx is a hollow tube that begins behind the nose and ends at the top of the trachea (windpipe) and esophagus (tube that goes from the throat to the stomach). The nostrils lead into the nasopharynx, while an opening on each side of the nasopharynx leads into each ear via the Eustachian tubes (Figure 1.2).

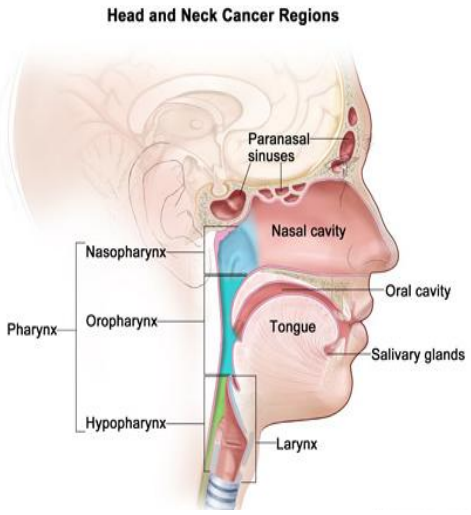

Figure 1.2 Schematic representation of Head and Neck region

Nasopharyngeal cancer (NPC) occurs when malignant cells form in the tissues of the nasopharynx. In most of the cases, NPC begins in the thin, flat cells (squamous cells) that line the oropharynx (the part of the throat behind the mouth) (Wei and Sham, 2005). NPC is one of the most confusing, commonly misdiagnosed and poorly understood disease because of the location of the involved area. The lesion is often situated in a relatively large and inert space where only air and mucus are in transit. It can be silent for a long time causing few primary symptoms. In recent decades, it has attracted worldwide attention because of complex interactions of genetic, viral, environmental and dietary factors, which might be associated with the etiology of this disease (Bei et al., 2012).

1.2.1 Stages of NPC

Based on clinical and radiologic examination, NPC stages can be categorized as: *Stage I* is a small tumor confined to nasopharynx. *Stage II* is a tumor extending in the local area, or that with any evidence of limited neck (nodal) disease. *Stage III* is a large tumor with or without neck disease, or a tumor with bilateral neck disease. *Stage IV* is a large tumor involving intracranial or infra-temporal regions, an extensive neck disease, and/or any distant metastasis. Most patients are generally presented with *Stage III or IV* disease because of non-specific presenting symptoms (cervical nodal enlargement, headache, nasal and aural dysfunction), delay in seeking treatment after the onset of symptoms, and the difficulty of a thorough nasopharyngeal exam (Pan et al., 2015).

1.2.2 Histological subtypes of NPC

NPC is classified into three types based on histology (Wei et al., 2011).

- **Type1:** Characterized by keratinizing squamous-cell carcinomas (SCC). This type of NPC tumor presents as squamous differentiation with the presence of intracellular bridges and/or keratinisation, and the absence of cellular infiltrate.
- **Type2:** Non-keratinizing squamous carcinoma, which is neither anaplastic nor keratinizing. It is a differentiated tumor with a stratified appearance and well defined cell borders. Cytologically, the tumor cells are of moderate size, variable in cell morphology (polygonal and/or spindled) and differentiation.
- **Type3:** Undifferentiated carcinomas that have distinct cytological characteristics. Typically, the cells appear to have prominent nucleoli, indistinct cytoplasm and poorly delineated cell boundaries. The tumor cell morphology of this type is highly variable as it can feature clear cells, spindle cells or anaplastic cells

1.2.3 Racial and geographical distribution

NPC has a remarkable racial and geographical distribution. The annual incidence rate is about 30 per 100,000 in a prevalent region such as Southern China. Ethnic Chinese living in Guangdong province and nearby regions, such as Guangxi and Hong Kong, are prone to the disease. An intermediate incidence (approximately 8 per 100,000 per year) has also been observed in some other population of South East Asian (e.g. Thailand, Philippines, Malaysia and Singapore) Arctic, Arabs in North Africa and Polynesians in New Zealand (Yu and Yuan, 2002).

NPC is common in Northeast region of the India. The National Cancer Registry has reported the prevalence of NPC to be 1.82% among all cancers in this region. Nagaland State has the highest incidence of about 4.3/100,000. High incidences are also observed in Manipur, Mizoram, and Sikkim. However, in Assam, the proportion of NPC among all cancer is only about 0.6%. The district-wise distribution

of the age-adjusted incidence rates (AARs) of NPC in Kohima district in Nagaland states was 19.4/100000 and the Imphal West district in Manipur State followed with a high AAR of 7.4/100 000 (Kataki et al., 2011).

1.2.3.1 Ethic information of northeast India: A high incidence region of NPC

Indians, comprises more that sixth of the world population with more than two thousand ethnic groups. Further, the variation that is found across the population on social parameters such as income and education increases the complexity. Hence, Indian provides a unique resource for dissecting complex disease etiology and pathogenesis. The evolutionary history of India entails migrations from central Asia and South China, resulting in a rich tapestry of socio-cultural, linguistic, and biological diversity. Broadly, Indians belong to the Austro-Asiatic, Tibeto-Burman, Indo-European, and Dravidian language families (Kataki et al., 2011).

Figure 1.3 Geographical map of northeast India States showing high NPC incidence areas viz., Manipur, Mizoram and Nagaland

The Northeastern region of India can be physiographically categorized into the Eastern Himalayas, Northeast Hills (Patkai-Naga Hills and Lushai Hills) and the Brahmaputra and the Barak Valley Plains. It is the eastern-most region of India connected to East India via a narrow corridor squeezed between Nepal and Bangladesh. It comprises the Arunachal Pradesh, Assam, Manipur, Meghalaya, Mizoram, Nagaland, Tripura, and Sikkim (Figure 1.3). Linguistically, the North-Eastern region is

distinguished by a preponderance of the Tibeto-Burman languages, and the population here is thought to comprise migrating peoples from East and Southeast Asia. Culturally, they are different from the rest of the sub-continent. Their unique socio-demographic, life-style and mongoloid ancestry make them more susceptible to NPC.

1.3 Etiology of nasopharyngeal carcinoma

Cancer is a complex disease and involves numerous factors. Similarly, several genetic and environmental factors and their interaction play key role in the manifestation of NPC. In the current study we investigated the association of all these factors.

1.3.1 Environmental factors

Many agents have been found to induce cancer, both in experimental animals and humans which include radiation, chemicals, and viruses. These physical and chemical or biological factors, which induce cancer growth, are called carcinogens.

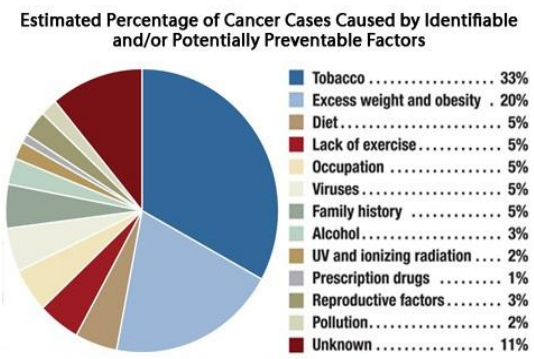

Figure 1.4 Percentages of different factors contributing to cancer

Ninety to ninety-five percentages of cancer can be attributed to environmental factors, including lifestyle, with the remaining 5–10% being due to inherited genetics (Figure 1.4). Among the environmental factors, dietary factors, several lifestyle habits like tobacco and alcohol consumption, as well as viral infections and occupational exposure to environmental pollutants are identified potent etiological factors of cancers including that of the nasopharynx.

1.3.1.1 Dietary factors

Dietary factors have significant contribution on the risk of cancers, and may have either increasing or reducing risk effects. It have been thought that dietary factors account for about 30% of cancers in Western countries (Doll and Peto, 1981), making diet second only to tobacco as a preventable cause of cancer. The contribution of diet to cancer risk in developing countries has been considered to be lower, perhaps around 20% (Miller, 2001). The role of diet in the etiology of NPC remains unresolved. However, there are some common features across the populations demonstrating an elevated risk of NPC. The most consistent results show increased risk associated with the consumption of Cantonese salted fish (Yu et al., 1986). Similar risks effect was observed with intake of other preserved foods in high-risk population of NPC, suggesting a link with NPC pathogenesis (Yuan et al., 2000). Extracts from preserved food including salted fish demonstrate the presence of carcinogenic nitrosamines/precursors, genotoxic and EBV-activating substances (Shao et al., 1988; Poirier et al., 1989). Use of herbal medicine including herbal medicine for clinical treatment, herbal tea and soups containing herbal ingredients have been proposed to be associated with an increase risk of NPC (Zheng et al., 1994). It was suggested that multiple types of herbal medicine extracts can induce EBV activation (Zeng et al., 1983). In addition, adequate fruit and vegetable intake are thought to be associated with a decrease NPC risk which are consistent with those observed for several other cancers (Key et al., 2002). The protective effects of fruits and vegetables may be due to their antioxidant properties.

1.3.1.2 Tobacco use

Tobacco consumption remains the most important avoidable risk factor of cancer. The World Health Organization (WHO) has established tobacco as the second

major cause of death in the world and responsible for the death of one in ten adults worldwide (WHO, 2005). A direct relationship between tobacco use and cancer of the nasopharynx has also been established by many studies (Feng et al., 2009; Fachiroh et al., 2012). It is consumed both in its smoking as well as in smokeless form. According to the International Agency for Research on Cancer (IARC), cigarette smoke contains over 4000 different complex constituents; more than 60 of these compounds are potent carcinogens such as polyaromatic hydrocarbons (PAHs), N-nitrosamines, aromatic amines, heterocyclic amines, aldehydes, volatile hydrocarbons, nitro compounds, miscellaneous organic compounds and metals and other inorganic compounds etc (IARC, 2004). These carcinogenic compounds in cigarette smoke are thought to be responsible for cause of several cancers (Pliarchopoulou et al., 2012; Amtha et al., 2014).

Smokeless tobacco consumption like betel nut or areca nut (with or without tobacco) chewing and other tobacco products like *gutkha*, *paan-masala*, *khaini* are endemic throughout the Indian subcontinent. Tobacco use with areca nut and other tobacco products results in the exposure to tobacco specific nitrosamines (TSNAs) and potent areca or betel nut related carcinogenic alkaloids areoline, arecaidine and guvacine. Two of the six TSNAs identified in smokeless tobacco, N'-nitrosonornicotine (NNN) and 4 (methylnitrosamino)-1 3-pyridyl-1-butanone (NNK), are strong carcinogens in mice, rats, and hamsters, capable of inducing both benign and malignant tumors of the oral and nasal cavity (Brunnemann et al., 1982) as well as of the lung, pancreas among others (Mitacek et al., 1999; Boffetta et al., 2005). These compounds are primarily generated by pyrolysis, but are also produced endogenously from smokeless tobacco. Most of the compounds in tobacco cause oxidative reactions in tissues, and initiate reactions to produce free radicals (Cury et al., 2012). The

presence of reactive oxygen species (ROS) can cause damage to cellular biomolecules, including protein and DNA, consequently resulting in carcinogenesis (Kumar and Muniyandi, 2015).

1.3.1.3 Alcohol consumption

Alcoholic or ethanol beverages are classified as group I carcinogens (carcinogens to human) by the IARC of the WHO. 3.6% of all cancer cases and 3.5% of cancer deaths worldwide are attributable to consumption of alcohol (Willett, 2000). IARC classifies consumption of alcohol as a major cause of several forms of cancer including breast, colon and rectum, liver, cancers of the upper aerodigestive tracts; and as a probable cause of pancreatic cancer (Tuyns, 1979; Blot, 1992). These studies indicate that ethanol present in alcoholic beverages is the crucial compound which causes cancer. However, the exact mechanism(s) is still unclear as ethanol itself is not a carcinogen.

Several mechanisms are involved in alcohol-associated cancer development which includes the effect of acetaldehyde (AA), the first metabolite of ethanol oxidation. Evidence has suggested that acetaldehyde is the predominant carcinogen responsible for the alcohol related cancers, since it binds to DNA and protein, destructs folate and results in secondary hyper-regeneration (Li et al., 2001). In addition chronic alcohol consumption induces cytochrome P-4502E1 (*CYP2E1*) in gastrointestinal mucosa cells and in the liver, leading to the generation of reactive oxygen species (ROS), and enhanced procarcinogen activation of various dietary and environmental carcinogens such as those present in tobacco smoke. It has also been suggested that alcohol acts as a solvent of tobacco and synergizes with tobacco as a risk factor for all upper aerodigestive tract squamous cell carcinomas. Nutritional deficiencies of vitamins (vitamin A) and trace elements (zinc and selenium) that are commonly

observed in chronic alcoholics due to primary and secondary malnutrition may further enhance alcohol associated carcinogenesis (Anand et al., 2008).

1.3.1.4 Occupational exposure

Occupational cancer is caused by exposure to carcinogens in the workplace. It can arise from exposure to many substances; biological carcinogens (e.g. viruses like hepatitis B, HIV viruses etc), chemical carcinogens (asbestos, vinyl chloride, by-products of industrial processes, like polycyclic aromatic hydrocarbons), and physical carcinogens (ionising and non-ionising radiations). Formaldehyde is a well-known carcinogen that can induce carcinoma of nasal cavity in rodents. In 1995, formaldehyde was suggested to be an etiological factor for NPC by IARC (Cogliano et al., 2005). Asbestos was first documented following as an important occupational lung carcinogen in the 1950s (Hughes and Weill, 1991). All different forms of asbestos; chrysotile and amphiboles, including crocidolite, amosite and tremolite are carcinogenic to humans, causing mesothelioma and lung cancer as per the International Programme on Chemical Safety, 1998. Exposure to inorganic arsenic, another known as a carcinogen (IARC, 1987) had increased risk of cancer of the lung, skin, and possibly urinary bladder and liver (Hayes, 1997). Exposure to X- and γ-radiation causes leukaemia and solid tumours in humans (IARC, 2000). An association between coal dust exposure and risk of cancer of the larynx and hypopharynx had been demonstrated (Shangina et al., 2006).

1.3.1.5 Viral pathogenesis

While considering the environmental factors special focus may be given to viruses. Many viruses infect humans but only a few viruses are known to promote human cancer. These include DNA viruses and retroviruses, a type of RNA virus.

Viruses associated with cancer include human papillomavirus (genital carcinomas), hepatitis B (liver carcinoma), Epstein-Barr virus (Burkett's lymphoma and nasopharyngeal carcinoma), human T-cell leukemia virus (T-cell lymphoma); and, probably, a herpes virus called KSHV (Kaposi's sarcoma and some B cell lymphomas).

1.3.1.5.1 Epstein Barr Virus and its role in the pathogenesis of NPC

Epstein Barr Virus (EBV) is a γ-herpes virus identified in 1964 from a patient who had African Burkitt lymphoma (Epstein et al., 1964). It contains a toroid-shaped protein core wrapped with double stranded DNA, a nucleocapsid and an envelope. The outer envelope is coated with multiple external glycoprotein spikes. EBV has a 184-bp long double-stranded DNA genome that encodes more than 85 genes. It present in over 90% of adults worldwide and is associated with polyclonal B lymphoproliferation in immune suppressed patients, Burkitt lymphoma, or Hodgkin's disease (Brady et al., 2007) and nasopharyngeal carcinoma (Raab-Traub, 2002). EBV transmission may occur via saliva or organ and bone marrow transplantation. The mucosal epithelium of the oropharynx is firstly infected, where the EBV replicates its DNA genome and assembles new viral particles (Figure 1.5). They undergo lytic infection within the cells and then release more virions that further infected the B lymphocytes. EBV infects primary resting B lymphocytes to establish a latent infection where the EBV genome exists as a closed, circular, non-integrated (episomal) DNA and transforms them into lymphoblastoid cell lines (LCLs) (Altmann and Hammerschmidt, 2005). In vitro studies demonstrate that complement receptor type 2 (CR2/CD21) is essential for EBV attachment on the surface of B-cells in an interaction mediated by the viral envelope glycoprotein gp350 (Baumforth et al., 1999). Hence, EBV appears to home to the oropharynx, and more specifically, the B cells within the oropharynx. In addition to the tropism towards B cells, the transformation potential of EBV has also been

demonstrated in NPC. In fact, EBV genome can be detected in virtually every NPC cell (Gullo et al., 2008; Tso et al., 2013). However, the mechanisms for EBV entry into nasopharyngeal epithelial cells and its maintenance of latency remain poorly understood. But because the EBV episome is identical in every NPC tumor cell, it supports the theory that NPC originates from clonal proliferation of a single EBV-infected progenitor cell. In these tumor cells, EBV adopts a specific form of latent infection called latency II infection.

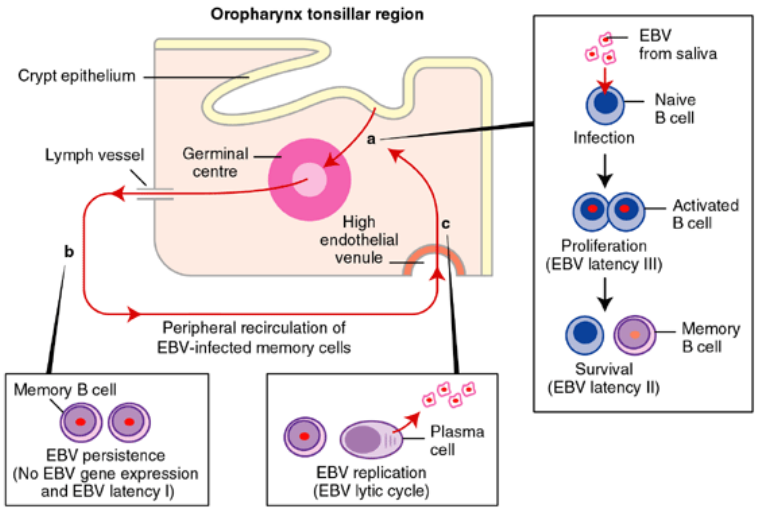

Figure 1.5 A model for Epstein Barr virus (EBV) infection and persistence

Only limited viral genes, including Epstein Barr virus nuclear antigen-1 (*EBNA1*), latent membrane protein-1 (*LMP1*), *LMP2*, and EBV-encoded small RNAs (*EBERs*) are expressed in NPC cells (Baumforth et al., 1999). It also appears that this latent infection occurs after some genetic changes have taken place in the precancerous lesions (especially allelic deletion of chromosomes 3p, 9p and 14q). This is supported by the expression of *EBER* in high grade dysplastic lesion but not in low grade lesions or normal nasopharyngeal epithelium. It have been postulated that the expression of these EBV viral proteins have profound cellular effects on gene expression, and may enhance tumorigenesis and increase the tumor invasive potential (Lo et al., 2004).

1.4 Genetic susceptibility of NPC

Several studies have been conducted to study the genes involved with NPC susceptibility. Since EBV infection has long been suspected to play an etiologic role in NPC development, the associations of specific HLA types with NPC susceptibility were extensively studied, in light of the capability of HLA to present antigen to cytotoxic T cells and, thus, trigger the host immune response against viral infection. Most of the studies focused on HLA class I (A, B and C) and class II (*DRB1, DQA1, DQB1*, and *DPB1*) antigens and used a serological typing approach to determine the HLA serogroup. *HLA-A2* and *HLA-A11* and at the *HLA-B* locus, *HLA-B13* and *HLA-B46* has been reported to be associated with NPC (Liebowitz, 1994; Hildesheim et al., 2002). Metabolic enzymes, which have major function in the metabolism of carcinogens, were examined for the association with NPC. It has been found that polymorphism in *CYP2E1* and *GSTT1* and *GSTM1* genotype are associated with NPC (Jia et al., 2009; Murthy et al., 2013). Polymorphism in DNA repair and damage genes; hoGG1, and *XRCC1* genes have been associated with the onset of NPC (Huang et al., 2011; Laantri et al., 2011). The development and progression of NPC have also included a number of acquired genetic changes (e.g. gene amplification, deletion and mutation) and epigenetic changes (methylation) that contribute to carcinogenesis by altering the expression of genes essential for a wide range of functions, such as proliferation, apoptosis, differentiation, invasiveness and metastasis.

1.4.1 Metabolic activation and detoxification genes

Xenobiotics (*Xenos= strange*) are compounds that are foreign to the body. It includes drugs, food additives, pollutants etc. Most xenobiotics are hydrophobic, a property that allows entry through the lipid bilayers into cells where it interacts with the cells receptors or proteins. This property of hydrophobicity renders it difficult to

eliminate, since in the absence of metabolism, they would accumulate in fat and cellular phospholipid bilayers in cells (Meyer, 1996). The xenobiotic-metabolizing enzymes convert xenobiotics into compounds that are hydrophilic derivatives that are more easily eliminated through excretion into the aqueous compartments of the tissues (Figure 1.6).

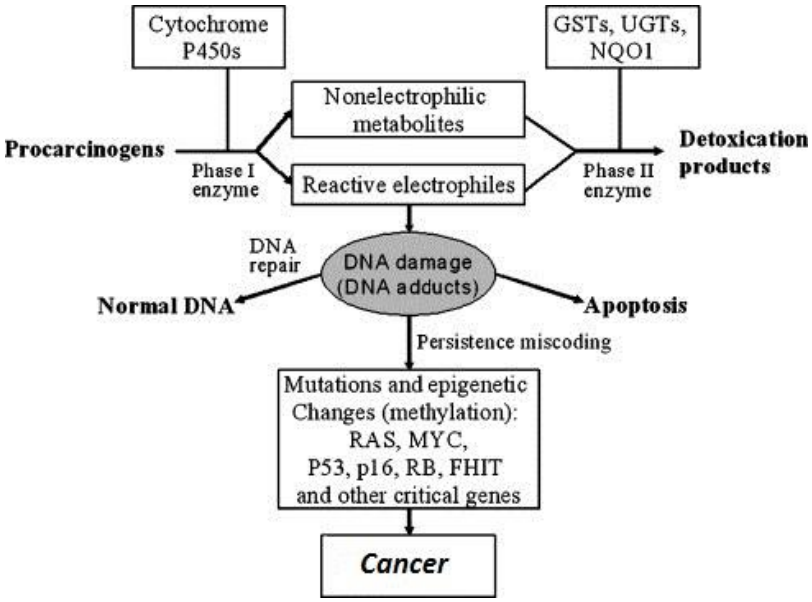

Figure 1.6 Schematic representation of xenobiotic detoxification and processes leading to cancer development

Xenobiotic metabolism is the set of metabolic pathways that modify the chemical structure of xenobiotics to endogenously produce reactive substances (Testa and Kramer, 2006). The liver is one of the most important organs in the body when it comes to detoxifying but a great amount of detoxification occurs in the gastrointestinal tract (Sheweita, 2000). The detoxifying enzymes are highly polymorphic exhibiting wide phenotypic variation. Impaired ability to remove reactive substances from the body may lead to chronic disease conditions. The outcome of biotransformation in most cases is detoxification; nevertheless, metabolism of some xenobiotics produces metabolites that are more reactive than their substrate compound. The biotransformation system involves several enzyme systems that are commonly divided into two phases;

phase I and phase II. The phase I enzymes are responsible for oxidation, reduction or hydrolysis and can be either detoxifying or activating. The phase II enzymes increases the detoxifying potential by conjugation of the phase I product as substrate (Ragin et al., 2010)

1.4.1.1 Phase I metabolic gene

The phase I enzymes lead to the introduction of functional groups, resulting in a modification of the xenobiotics, such that it now carries an –OH, -COOH, -SH, -O- or NH2 group. The addition of functional groups does little to increase the water solubility of the xenobiotics, but can dramatically alter the biological properties of the xenobiotics. Phase 1 metabolism is classified as the functionalization phase of xenobiotic metabolism; reactions carried out by phase 1 enzymes usually lead to the inactivation of active xenobiotics. The phase 1 oxidation reactions are carried out by CYPs, flavin-containing monooxygenases (FMO), and epoxide hydrolases (Yu, Mo et al.). The CYPs and FMOs are composed of superfamilies of enzymes. Each superfamily contains multiple genes (Meyer, 1996).

Cytochrome P450, family 1, subfamily A, polypeptide 1 (CYP1A1) one of the major xenobiotics-metabolizing Phase I enzyme, is included in the cytochrome P450 super family. It is also known as AHH (aryl hydrocarbon hydroxylase) and is located at 15q24.1. Four common polymorphisms have been identified in *CYP1A1*; *m1*- T3801C substitution in 3’ non-coding region, *m2*- A2455G substitution leading to change of isoleucine to valine at codon 462, *m3*- T3205C substitution in 3’ non-coding region and *m4*- C2453A substitution leading to change of threonine to asparagine at codon 461 (Sergentanis and Economopoulos, 2010). The common polymorphisms (*MspI* or CYP1A1*2A polymorphism, rs4646903) located in the 3’ non-coding region containing a T to C nucleotide substitution at 3801, give rise to a *MspI* restriction site, which

results in three genotypes namely *wt/wt*, which lacks the *MspI* restriction site, *wt/m1* and *m1/m1*, which are heterozygous and homozygous respectively for the polymorphic allele with the *MspI* site. The *m1* allele has been associated with higher induction of CYP1A1. The higher enzyme activity would result in increased levels of carcinogenic intermediates, leading to greater risk of cancer development (Mota et al., 2010; Sabitha et al., 2010).

1.4.1.2 Phase II metabolic genes

Phase II enzymes facilitate the elimination of xenobiotic and the inactivation of electrophilic and potentially toxic metabolites produced by oxidation. While many phase reactions result in the biological inactivation of the drug, phase 2 reactions produce a metabolite with improved water solubility and increased molecular weight, which serves to facilitate the elimination of the xenobiotic from the tissue. The phase 2 enzymes include several superfamilies of conjugating enzymes. Among the more important are the glutathione-S-transferases (*GST*), UDP glucuronosyltransferases (*UGT*), sulfotransferases (*SULT*), *N*-acetyltransferases, and methyltransferases (*MT*) (Jancova et al., 2010).

Glutathione S-transferases (GSTs) are another group of xenobiotics-metabolizing enzymes involved in Phase II detoxification of carcinogens. These enzymes play a central role in the detoxification of many endogenous and exogenous substrates through conjugation to glutathione, a tripeptide consisting of glycine, glutamic acid, cysteine to electrophliic compounds, resulting in less reactive and more easily excreatable glutathione conjugates. Among the 3 mammalian *GSTs*, (mitochondrial, cytosolic and microsomal) cytosolic *GSTs* represent the largest family and exhibits significant genetic polymorphism. Cytosolic *GST* isoenzymes can be classified by their substrate specificities, isoelectric points and amino acid sequence

homologes into major classes whch are encoded by a superfamily of genes located at different loci. The different isoenzyrnes of cytosolic GSTs are Mu, Theta, Pi, Sigma, Omega, Alpha and Zeta. The *GSTT1* and *GSTM1* are polymorphic enzymes with inter-individual variations in enzymatic level and activity. These two genes are located at Chr 1p13.3 and 22q11.23 respectively. The homozygous deletion genotypes of *GSTM1* and *GSTT1* result in an absence of enzyme activity which potentially increases an individual susceptibility to cancers (Nair et al., 1999; Schneider et al., 2004).

1.4.2 DNA repair genes

In human, DNA damage due to both environmental factors and normal metabolic processes inside the cell occurs at a rate of 1,000 to 1,000,000 molecular lesions per cell per day. While this constitutes only 0.000165% of the human genome's approximately 6 billion bases (3 billion base pairs), unrepaired lesions in critical genes (such as tumor suppressor genes) can impede a cell's ability to carry out its function and appreciably increase the likelihood of tumor formation (Cooper, 2000). A cell that has accumulated a large amount of DNA damage, or one that no longer effectively repairs damage incurred to its DNA, can enter one of three possible states; an irreversible state of dormancy, known as senescence, cell suicide, also known as apoptosis or unregulated cell division, which can lead to the formation of cancer. Cells cannot function normally if damage in DNA hampers the integrity or accessibility of essential information in the genome. Depending on the type of damage, a variety of repair strategies have evolved to restore lost information:

1) Direct chemical reversal of the damage and

2) Excision repair, in which the damaged base or bases are removed and then replaced with the correct ones in a localized burst of DNA synthesis.

There are three modes of excision repair, each of which employs specialized sets of enzymes:

a) Base Excision Repair (BER) which repairs damage to a single nucleotide caused by oxidation, alkylation, hydrolysis, or deamination. The base is removed with glycosylase and ultimately replaced by repair synthesis with DNA ligase.

b) Nucleotide Excision Repair (NER) which repairs damage affecting longer strands of 2–30 bases. This process recognizes bulky, helix-distorting changes such as thymine dimers as well as single-strand breaks

c) Homologous Recombination Repair (HRR) in which both strands in the double helix are severed, are particularly hazardous to the cell because they can lead to genome rearrangements

1.4.2.1 Polymorphism in repair genes

The repair of damaged DNA is essential to prevent mutations thus prevent mutagenic transformations. Individuals with low DNA repair capacity (DRC) may accumulate genetic alterations and may be at increased risk of developing different kinds of cancer. Polymorphism may result in subtle structural alterations of the repair enzymes and modulation of cancer susceptibility. In this study, two polymorphisms in two DNA repair genes were studied which have strong association with different form of cancers.

XRCC1 (x-ray repair cross-complementation group 1) protein is located at Chr. 19q13.2 is a scaffolding protein directly associated with polymerase and functions in a complex to facilitate the BER and single break-repair processes. Three polymorphisms occurring at conserved sequences in the *XRCC1* gene were reported (Shen et al., 1998). These coding polymorphisms, resulting in amino acid substitutions,

were detected at codons 194 (Arg-Trp), 280 (Arg-His) and 399 (Arg-Gln). Many authors have analysed these polymorphisms in human populations and found a significant association between the Arg194Trp and Arg399Gln variants and increased risk of early-onset colorectal (Abdel-Rahman et al., 2000) and gastric cancer (Shen et al., 2000), in addition to head and neck (Choudhury et al., 2014) and skin cancer (Shen et al., 2000) associated with the Arg194Trp variant, and breast (Romanowicz et al., 2010; Zhu et al., 2015) and lung cancer (Uppal et al., 2014; Kang et al., 2015), amongst others, associated with Arg399Gln polymorphism.

XRCC2 (x-ray repair cross-complementing group 2) gene, located at 7q36.1, is an essential part of the homologous recombination repair pathway and a functional candidate for involvement in cancer progression (Thacker and Zdzienicka, 2003). Common variants within *XRCC2*, including single nucleotide polymorphisms (SNPs) in exon 3 (Arg188His, R188H, rs3218536 or 31479G>A), have been identified as potential cancer susceptibility loci in recent studies, although association results are controversial. The Arg188His polymorphism has been proposed to be a genetic modifier for smoking-related pancreatic cancer (Jiao et al., 2008) and was associated with an increased risk of head and neck cancer risk (Choudhury et al., 2014).

1.4.3 Newly identified susceptible loci (*MECOM*, *TNFRSF19* and *CDKN2B-AS1*)

Environmental factors play a very large role in the development of NPC. It is also been known for quite some time, however, that genetics contribute to NPC susceptibility. Genetic association and linkage studies consistently report that NPC appears to be associated with the HLA-A region (Li et al., 2009; Tse et al., 2009), and two genome-wide association studies (GWAS) recently confirmed the HLA region's role in NPC in southern Chinese and Taiwanese populations (Hsu et al., 2012; Tang et al., 2012). Now a large GWAS has revealed new susceptible loci in the genome that are

associated with risk for NPC. Such findings could eventually help develop models for prediction and screening, which in turn would help with early diagnosis (Bei et al., 2010). These new susceptible loci include; *TNFRSF19* on chromosome 13q12.12 (rs9510787), *MDS1-EVI1* on chromosome 3q26.2 (rs6774494) and *CDKN2A/CDKN2B* (or *CDKN2B-AS1*) gene cluster on chromosome 9q21.3 (1412829) (Table 1.2).

Table 1.2 Risk allele associated with the three hotspot genes

SNP	Gene	Risk allele	Reference allele
rs9510787	*TNFRSF19*	G	A
rs6774494	*MDS1-EVI1*	A	G
rs1412829	*CDKN2B-AS1*	A	G

TNFRSF19 (Tumor Necrosis Factor Receptor Superfamily, Member 19) encoded protein is a member of the TNF-receptor superfamily (Hu et al., 1999). This receptor is highly expressed during embryonic development. It has been shown to interact with tumor necrosis factor receptor associated factors (TRAF) family members. When overexpressed, *TNFRSF19* activates the c-Jun N-terminal kinase (JNK) pathway and induces caspase-independent cell death (Eby et al., 2000). This receptor is capable of inducing apoptosis by a caspase-independent mechanism, and it is thought to play an essential role in embryonic development. Alternatively spliced transcript variants encoding distinct isoforms have been described. Given the epithelial expression of *TNFRSF19* in many embryonic tissues (Morikawa et al., 2008) and the epithelial origin of NPC, the dysregulation of *TNFRSF19* in the epithelium of the nasopharynx may be involved in NPC. Moreover, EBV-encoded latent membrane protein 1 is oncogenic and drives cell transformation through a mechanism similar to the *TNF* receptor family members (Eliopoulos and Young, 2001), lending further biological plausibility to the involvement of *TNFRSF19* in NPC.

MECOM [MDS1 (myelodysplasia 1) and EVI1 (ecotropic viral insertion site 1 fusion proteins) complex locus protein] or *MDS1* and *EVI1* complex locus is a protein-coding gene. The protein encoded by this gene acts as a transcriptional regulator binding to DNA sequences in the promoter region of target genes and regulating positively or negatively their expression. They may acts as an oncogene which plays a role in development, cell proliferation and differentiation. And also play role in apoptosis through regulation of the JNK and TGF-beta signalling. *MDS1-EVI1* encodes three proteins, *EVI1, MDS1* and the fusion protein *MDS1-EVI1*. *EVI1* is a transcription factor involved in leukemic transformation of hematopoietic cells (Metais and Dunbar, 2008). *EVI1* can suppress the effect of transforming growth factor (TGF)-α on growth inhibition, which in turn promotes tumor growth; *EVI1* can also protect cells from stress-induced cell death by inhibiting c-JNK (Kurokawa et al., 2000; Metais and Dunbar, 2008). In contrast, when *EVI1* was fusion with *MDS1*, its capacity to repress TGF-α signalling was significantly impaired (Nitta et al., 2005). Given that TGF-α and JNK signalling pathways are known to be involved in EBV-related tumorigenesis of NPC (Xu et al., 1999) the interruption of the balance between *EVI1* and *MDS1- EVI1* proteins may be important for the pathogenesis of NPC.

CDKN2B-AS1 (CDKN2B antisense RNA 1) gene is located within the *CDKN2B-CDKN2A* tumor suppressor gene cluster at chromosome 9p21 locus, which is the strongest genetic susceptibility locus for several diseases (Zeggini et al., 2007; Burdon et al., 2011; Lee et al., 2014), and also linked to cancers (Iacobucci et al., 2011; Sanson et al., 2011). Many disease-associated single-nucleotide polymorphisms in this locus affect the structure and expression of this gene, suggesting that modulation of this gene expression mediates disease susceptibility. This gene interacts with both polycomb repressive complex-1 (*PRC1*) and -2 (*PRC2*), and may function as a regulator for

epigenetic transcriptional repression (Popov and Gil, 2010; Zhang et al., 2014). Multiple alternatively spliced transcript variants have been generated from this gene, and all of them are long non-coding RNAs. It has been found that some of splice variants are tissue specific, and different splice variants may have distinct roles in cell physiology. All of the genes; *TNFRSF19, MDS1-EVI1* and *CDKN2B-AS1* have previously been shown to be involved in leukemia (Gruss and Dower, 1995; Soderholm et al., 1997; Berndt et al., 2013). This suggests that there might be common disease mechanisms between that disease and NPC. Significantly, leukemia is found at higher than average rates in people with NPC (Li et al., 2013a).

1.4.4 Mitochondria

Mitochondria - from the Greek *mitos* (thread-like) and *khondros* (grain or granule) are small organelles that are present in all nucleated cells. Mitochondria are unique structures that exert essential functions in energy metabolism, free radical production, calcium homeostasis and apoptosis (Wallace, 2010). Mitochondria range from 0.5 to 1.0 μm in diameter and up to 10 μm long. The content of mitochondria is modulated according to cell's metabolic activity and may undergo significant changes under diverse internal or external microenvironments, such as hypoxia and steroid hormone stimulation (Hoppeler et al., 2003; Shadel, 2008).

Human mtDNA is 16569 bp, maternally inherited and it encodes for few products (Figure 1.7); (i) Genes for 13 polypeptides, which form subunits of the oxidative phosphorylation (OXPHOS) complexes, (ii) Genes involved in the protein synthesis machinery, i.e. ribosomal ribonucleic acid (rRNA) and (iii) set of genes for transfer ribonucleic acids (tRNA) (Anderson et al., 1981).

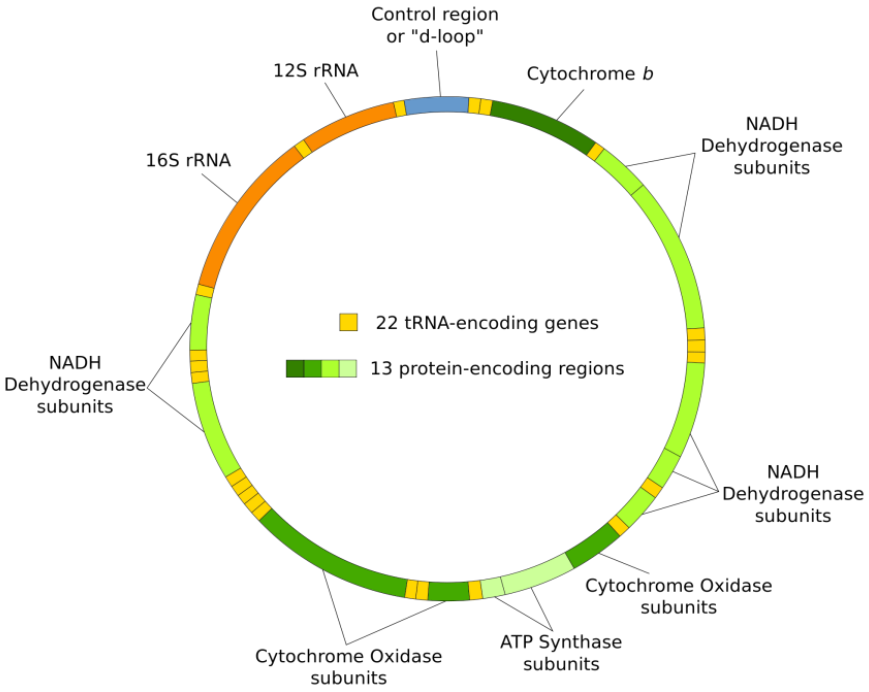

Figure 1.7 Illustration of human mitochondrial genome with its 37 genes and the coding region

Human mtDNA consists of two strands, heavy (H), is completely saturated with genetic information; it encode 2 rRNA genes, 12 proteins and 14 tRNA. The light strand (L), encodes only one mitochondrial protein (*ND6*) and eight genes for tRNA. The main regulatory region in mitochondrial DNA is called the D-loop it is 1.24 Kb, and because of the nascent H strand creates a triple-stranded structure with the displacement (D) of the old strand this region contains the origin of replication of the H strand (O_H) and two promoters, the heavy-strand (*HSP*) and the light-strand (*LSP*) promoters. It can be divided into three domains: the central domain, well conserved in evolution but whose function is still unknown, and two peripheral, left and right domains, very variable, called the conserved sequence box (*CSB*) and extended termination-associated sequences (*ETAS*) domains respectively, on the basis of some conserved boxes contained in them (Clayton, 2000; Clayton, 2003). Mitochondria host a protein pool of approximately 3000 proteins, only 13 of which is coded by the host organelle. Other protein are coded by the nuclear genome and is associated with other

functions like fatty acid synthesis, porphyrin synthesis etc. This sharing of space and function has poise a serious relation of inter dependency of the mitochondria on the nuclear genome.

1.4.4.1 Mitochondria in cancer

Mitochondrial functional defects have long been hypothesized to contribute to the development and probably progression of cancer. It was in 1920s when Otto Warburg hypotheses that because of mitochondrial malfunction, cancer cells had to depend on anaerobic glycolysis even in the presence of oxygen to generate ATP (Warburg et al., 1927). This phenomenon, known as the Warburg effect, has repeatedly observed in cancer cells, along with metastasis, angiogenesis, and endless replication (Kondoh et al., 2007). It is however in contrast to normal cell which produces energy (ATP generation) mainly through an oxygen-dependent process called oxidative phosphorylation (OXPHOS) and is shifted to glycolysis only in deceased level of oxygen (Pasteur effect). The Warburg effect is actually the basis for the widespread application of positron emission tomography in which a glucose analog tracer (2-^{18}fluoro-2-deoxy-D-glucose) is used to differentiate between normal and tumor tissue (Gatenby and Gillies, 2004). Persistent activation of glycolysis in cancer cell provides building blocks for the synthesis of nucleotides (via the pentose phosphate pathway) and amino and fatty acids (from intermediates formed in the glycolytic and tricarboxylic acid cycles) required for tumor proliferation (Xu et al., 2005; Pelicano et al., 2006). In addition, increased lactic acid production may result acidification of the tumor microenvironment and facilitate tumor invasion and metastasis (Gatenby et al., 2006). The enhanced activity of the pentose phosphate shunt may also lead to an elevated production of nicotinamide adenine dinucleotide phosphate (NADPH) and glutathione which would increase the resistance of tumor cells against oxidative insults and some

chemotherapeutic agents. Recent data also suggest that cancer cells may dependence on glycolysis for ATP generation increases as malignant transformation occurs. It was found that accumulation of glucose metabolites resulting from activation of glycolysis favors activation of Hypoxia-inducible factors 1 (*HIF-1*) (Lu et al., 2005) which plays key role in the transcription of several genes that encode proteins favoring cancer development, including those involved in glucose metabolism, apoptosis resistance, invasion, metastasis, and angiogenesis (Lopez-Lazaro, 2006a; Lopez-Lazaro, 2006b).

Apart from this mitochondrial respiration is also associated with the generation of reactive oxygen species. Absence of protective histones with limited DNA repair capacity and lack of introns and its close physical proximity to high levels of endogenous reactive oxygen species (ROS) in the mitochondrial inner membrane, mtDNA is extremely prone to oxidative or other genotoxic damages and thus acquire mutations at a much higher rate (10- to 200-fold) than nDNA (Liu and Demple, 2010).

1.4.4.2 Mitochondrial DNA copy number alterations in human cancer

Mitochondrial DNA (mtDNA) copy number can be defined as to the number of functional mitochondrial genomes present per cell or tissue. The mtDNA copy number or content is measured by quantitative PCR assay targeting total mtDNA and normalized to input template determined by amplifying a nuclear target (Mondal et al., 2013a). The mtDNA copy number per cell is maintained within a constant range to sustain the normal physiological functions. In normal human cell, each mitochondrion contains about 2-15 copies of the mitochondrial genome. It varies significantly among the population from 1000 to 10,000 copies per cell and with the type of cell, where the number is significantly higher in active cells like muscle and brain that requires a large amount of energy. While this copy number enables ease of assaying mitochondrial genome for diagnostic purposes, it is also established that changes in mtDNA copy

number underlie certain pathologic conditions, and are demonstrated in several cancers as well (Hosgood et al., 2010; Thyagarajan et al., 2013).

Mitochondrial DNA copy number is regulated by various complex mechanisms within a cancer cell and may have certain degree of specificity for the tumor sites or the type of the tumor. Alterations in mtDNA copy numbers can be resulted from factors that destruct mtDNA and adversely affect mtDNA propagation. The mtDNA lacks its own repair capacity and its integrity is maintenance by certain nuclear gene. Mutation in the nuclear genes can also causes mtDNA depletion. Furthermore, mutations in the mtDNA D-loop region have been found to be significantly associated with reduced in mtDNA content in different forms of cancers, mostly because D-loop acting as a promoter for mtDNA replication and transcription (Lee et al., 2004; Lee et al., 2005). Mutations in mitochondrial polymerase gamma (*POLG*), transcription factor A (*TFAM*) influence the rate of replication and transcription and lead mtDNA depletion (Singh et al., 2009; Guo et al., 2011). Abnormality of p53 observed in more than 50% of human cancers may lead to a remarkable increase of mtDNA's sensitivity to oxidative stress and disturbs mitochondrial ROS homeostasis, presumably inducing variance enhancement of mtDNA content (Lebedeva et al., 2009). Indeed, the concomitant presence of reduced mtDNA content and *p53* or *POLG* mutations has been discovered in multiple cancer types (Chang et al., 2009; Singh et al., 2009).

1.5 Rationale of the propose study

Although considered a rare disease, nasopharyngeal carcinoma (NPC) is mostly confined to the Cantonese and Chinese populations living in the southeastern region of China and also among their descendants who have migrated to other geographical areas. The annual incidence rate is about 30/100,000 in a prevalent region

such as Southern China. A moderately high incidence of NPC has also been reported in some other population groups also such as Eskimos in the Arctic, Arabs in North Africa, Malays of South East Asia and Polynesians in New Zealand (Jia and Qin, 2012).

In spite of very high incidence of oral cancer in the Indian subcontinent, NPC is uncommon in Indian subcontinent except for North-East region of the country. Over 160 Scheduled Tribes and over 400 other sub-tribal communities and groups are present in the region. The northeastern (NE-India) tribes have a distinct genome pools compare from rest of mainland Indian and other groups; and are more closely related to East-Asians (Cordaux et al., 2003). The region is distinguished by a preponderance of the Tibeto-Burman languages, and the population is thought to comprise migrating peoples from East and Southeast Asia. The Mongoloid race in this region has shown an increase in NPC incidence particularly; *Naga, Manipuri* and *Mizo* groups (Kataki et al., 2011). As per the National Cancer Registry programme (NCRP), 2009-2011 Nagaland State has the highest incidence with age-adjusted incidence rate (AAR) of 21/100,000 among males. Similarly, high incidences among males were also observed in Manipur (4.5/100,000) and Mizoram States (4.9/100,000).

The epidemiology of NPC is complex and evidence from endemic regions has suggested that both genetic and environmental factors (dietary, lifestyle and viral factors) play a vital role in its development (Baumforth et al., 1999; Bei et al., 2012; Jia and Qin, 2012). And interestingly, the unique socio-demographic, dietary, life style factors and their mongoloid ancestry might played a crucial role in making the northeast ethnic community more susceptible to NPC. Despite of the exposure to the same environmental conditions few people are susceptible to NPC suggesting the involvement of genetic factors and the possibility of a strong interaction between them in NPC carcinogenesis. In the northeast India consumption of traditional preserved

foods (smoked meat, fermented fish etc.) and use of herbal medicine for various ailments has always been a custom. Moreover, tobacco related habits like betel quid chewing (with/without tobacco) and smoking cigarette or bidi, and consumption alcohol (mostly locally prepared) are a common practice. Ethanol and acetaldehyde are known carcinogens present in alcoholic drinks (Lachenmeier et al., 2012). Tobacco both smoke and smokeless contains a complex mixture of carcinogenic compounds (Nagaraj et al., 2006; Sabitha et al., 2010; Talukdar et al., 2013); and salted fish, meat and other preserved foods are rich in nitrosamines (Sarkar et al., 1989; Diggs et al., 2011) and generate reactive oxygen species (ROS) which can damage both nuclear and the mitochondrial genome. The phase I and II metabolic enzymes activate and detoxify the carcinogens before eliminating from the body (Xue and Warshawsky, 2005; Shukla et al., 2013). However, inter-individual genetic variation may alter enzymatic activity and subsequently carcinogens, activation or deactivation, thereby increasing susceptibility to cancer risk (Rodriguez-Antona and Ingelman-Sundberg, 2006). On the other hand damage in the genome is monitored by the DNA repair systems. Any defect in the DNA repair pathway will lead to gene mutations and chromosomal instability, known to play major role in tumorigenesis (Berwick and Vineis, 2000; Shi et al., 2004; Metsola et al., 2005).

Studies in endemic areas have investigated the association of polymorphism in the metabolic (*CYP1A1, GSTT1 & GSTM1*) genes (Guo et al., 2008; Jiang et al., 2011) and repair (*XRCC1 & XRCC2*) genes (Cho et al., 2003; Li et al., 2013b) in NPC but have generated equivocal results. Moreover, data are lacking for the interaction of these genes and the environmental factors. It is also well established that the EBV infection is a risk factor for NPC. Experimental evidence in cell and animal models have suggested the presence of possible EBV inducers or activators in tobacco, preserved food samples,

herbal medicine, however limited study involving human subjects were carried out (Bouvier et al., 1991; Hildesheim et al., 1992; Jenson et al., 1999; Fang et al., 2012).

Many reports have also indicated that mitochondria is involved in apoptosis (Wang and Youle, 2009) and probably in tumorigenesis (Park et al., 2009), which has led researchers to examine the potential role of mitochondrial DNA (mtDNA) alterations in the development and maintenance of cancers. The mitochondrial genome is susceptible to ROS and other types of genotoxic damage due to lack of protective histones and its limited mtDNA repair capabilities (Mondal et al., 2013b). The mtDNA copy number per cell is maintained within a constant range to meet the energy requirement of the cell to sustain normal physiological functions. It is likely that the variations in the copy number of mitochondria reflect the net results of gene–environmental interactions between unknown hereditary factors and the levels of oxidative stress (an imbalance between ROS production and the antioxidant capacity), caused by a variety of endogenous and exogenous factors, such as hormones, age, dietary and environmental oxidants/ antioxidants and reaction to oxidative damage, all of which are thought to be the risk factors for various types of cancer.

In addition, a recent GWAS on NPC risk within a Han Chinese population suggested the involvement of variants at the 9p21 (*CDKN2B-AS1*), 13q12 (*TNFRS19*), and 3q26 (*MDS1-EVI1*) in NPC carcinogenesis (Bei et al., 2010). However, it is not established whether these additional susceptibility loci are relevant to other populations. We therefore examine the selected risk factors which include mostly diet and lifestyle, EBV infection, genes polymorphisms, mtDNA copy number variation and their interactions in the development NPC. This is a first kind of study on ethnic groups from Northeast India prone to NPC that has included cases, first-degree relatives and controls to examine the prospects of genetic susceptibility to cancer. Moreover, the study would

be helpful in formulating appropriate prevention and intervention programmes by public health authorities.

1.6 **Objectives of the study:**

- To study the prevalence of Nasopharyngeal Carcinoma and its association with viral and environmental factors in the ethnic population North East India.
- To study the association of novel variation(s) in metabolic-genes (*GSTT1, GSTM1* and *CYP1A1*), DNA repair genes (*XRCC1* and *XRCC2*) other susceptible loci (*MECOM, TNFRSF19* and *CDKN2B-AS1*) and NPC risk.
- To predict for any chances of NPC development in the first degree relatives (FDR) of NPC affected individuals.
- To explore high order interactions among environmental, viral and genetic factors towards NPC progression.
- To study the association of mtDNA copy number and NPC risk.