CHAPTER 1

General Introduction
1.1 Cell proliferation, differentiation and cancer

Cancer is a differentiation disorder in which cells deregulate the balance between proliferation and differentiation (Swarting et al., 2015; DeGraff et al., 2013; Frank and Miranti, 2013), Figure 1.1. While cell proliferation and differentiation are tightly regulated during normal development, cancer cells exhibit both uncontrolled proliferation and blocked differentiation (Matushansky et al., 2003). This transition occurs mostly in stem cells, which, due to loss of their regulatory controls give rise to cancer (Sell, 2010; Wu, 2008). Such subtle albeit important coordination is brought about by reorganization of multifarious cellular processes including the cell cycle (Matushansky et al., 2003), differentiation (Frank and Miranti, 2013), transcription regulation (Kulis et al., 2013) and post-transcriptional regulation through microRNAs (Wang et al., 2013). A better understanding of these mechanisms would shed light on how cancers promote proliferation in favour of differentiation and hence enable designing of novel strategies to combat this global malignancy.

**Figure 1.1 : Cell proliferation, differentiation and cancer**

Balance between cell proliferation and differentiation between normal cells, if destabilized, lead to evolution of cancer cells.
1.2 Hallmarks of Cancer

Carcinogenesis, or the initiation of cancer is a complex and multistep process in which alterations (genetic/epigenetic) in normal cells perturb normally tightly regulated cellular proliferation and differentiation. This occurs through disturbance of different cellular pathways, broadly due to deregulation of three elemental cellular processes, viz. a) Cell fate, b) Cell survival and c) Genome maintenance (Vogelstein et al., 2013) vide Figure 1.2.

a) **Cell fate**: The fate of a cell, governed by two opposing parameters—cell division and differentiation, is negated by genetic alterations in cancer cells, promoting division over differentiation, thus offering a selective growth potential through modulation of self-renewal pathways (Perrimon et al. 2012; Hoffman 2012) chromatin modifications (ENCODE Project Consortium, 2012; Rosetto et al., 2012).

b) **Cell survival**: In contrast to normal cells, cancer cells gain selective advantage through perturbation of the cell cycle (Boughton, 2015; Chow, 2010; Sherr, 1996), apoptosis (Correia et al., 2015; Brown and Attardi, 2005) or receptor mediated signalling (Hynes and Lane, 2005; Turner and Grose, 2010).

c) **Genome maintenance**: Spontaneous mutations produced in the body (Hoeijmakers, 2009) or genotoxins produced in the microenvironment (Fiaschi and Chiarugi, 2012) provide survival advantages through mutations in checkpoint surveillance (Derheimer and Kastan, 2010), or DNA repair pathways such as mismatch repair, double strand break repair, etc. (Farrel et al., 2012, Woo et al., 2014; Lieber, 1998).
Alterations of normal cells lead to perturbation of three elemental processes, vide cell survival, cell fate and genome maintenance, leading to selective growth advantage and cancer. Figure adapted from Vogelstein et al., 2013.

As a result, there is a deregulation of several cellular phenomena, leading to the molecular evolution of different cancers. These common and basic mechanisms, widely known as the “Hallmarks of Cancer” were suggested by Hanahan and Weinberg, 2011 and include the following (Figure 1.3):

a) Sustaining proliferative signalling- Masters of their own destinies.
b) Resisting cell death- Circumventing the apoptotic trigger.
c) Inducing angiogenesis- The angiogenic switch.
d) Evading Immune Destruction- Incognito to immune surveillance:
e) Enabling replicative immortality- Rewinding the telomere clock.
f) Tumour promoting inflammation- The enemy within.
g) Evading growth suppressors- Inhibition of tumour suppressors.
h) Activating invasion and metastasis- Epithelial to mesenchymal transition.
i) Reprogramming energy metabolism- Aerobic glycolysis.
j) Genomic instability and mutability- Enhancing auto generated alterations.
Figure 1.3: Hallmarks of cancer:

Deregulation of the ten different basic cellular processes in cancer, widely known as the “Hallmarks of cancer”. Image adapted from Hanahan and Weinberg, 2011.
1.3 Cancer associated genes

1.3.1 Classification of genes associated with cancers

The genes associated with the development and progression of cancer can be classified according to their mode of action as i) Oncogenes, ii) Suppressors and iii) Chameleons (Stepanenko et al., 2013), Figure 1.4.

a) **Oncogenes**: Oncogenes are altered counterparts of proto-oncogenes, genes that encode proteins associated with stimulation of cell division, inhibition of cell differentiation and halting of cell death. The alterations in these genes are dominant, i.e., alteration in a single allele is sufficient for a phenotype to develop (Chial, 2008). Eg. PI3K, RAS (Iurlaro et al., 2014).

b) **Suppressor genes**: Tumour suppressors (TSGs) function in preventing inappropriate cell proliferation, stimulating cell death and DNA repair (Chial, 2008). Alterations in TSGs are recessive and follow the two hit hypothesis, i.e., for a cancer to develop, both alleles of a TSG should undergo alteration (Knudson, 1971). TSGs are further classified into landscapers, gatekeepers and caretakers, based on their mode of action (Srivastava and Grizzle, 2010). Landscaper genes facilitate neoplastic growth by creating a microenvironment for unregulated cellular proliferation, e.g., PTEN (Stratakis, 2003). Gatekeepers are genes that regulate the growth of tissues by inhibiting proliferation, facilitating apoptosis and promote cell differentiation, e.g., RB (Macleod, 2010). Caretaker genes maintain stability of the genome and are hence necessary for normal cell growth, e.g., BRCA1, BRCA2 (Yu, 2000).

c) **Gene chameleons**: Chameleons on the other hand comprise of a paradoxical group which, although showing suppressor properties might behave in an opposite manner under different experimental settings and present antagonistic duality (Schneider-Stock, 2014), e.g., Death associated protein kinase 1 (DAPK 1) (Schneider-Stock, 2014).
On the basis of their functions, genes can be classified as oncogenes, tumour suppressor genes and chameleon genes.

1.3.2 Alterations of genes associated with cancer

Cancer cells accumulate somatic mutations at a rate significantly higher than their normal counterparts through a property called “Mutator Phenotype”, leading to acquisition of the above hallmarks of cancer (Bielas et al., 2006). These alterations in genes associated with cancers can be categorized broadly into i) Genetic, and ii) Epigenetic alterations.

1.3.2.1 Genetic alterations:

Genetic alterations, those directly affecting the DNA of the cancer genome have been identified through exhaustive whole genome/transcriptome/proteome analysis and sequencing. These alterations may either be “drivers”, those that grant a selective growth advantage to cells or “passengers” or pre-neoplastic mutations having no effect on the neoplastic process (Vogelstein et al. 2013). The common aberrations encountered in cells are either a) Chromosome copy number variations or b) Chromosomal alterations (Translocation, Deletion, Amplification, Point mutation, Frameshift)
variations, Viral insertion, etc.) with these being mostly irreversible (Herceg and Hainaut, 2007).

a) **Chromosome copy number variations/ aneuploidy**: Includes addition or deletion of entire chromosome(s) or its arms (Bürger et al., 2013; Schnaiter and Stilgenbauer, 2013; Elbert, 2009).

b) **Chromosomal alterations**: Includes alterations affecting parts of the chromosomes, Figure 1.5.

1. **Translocation** - Results from inappropriate re-ligation of two DNA double-strand breaks (DSBs) in heterologous or same chromosome (Byrne et al., 2014; Zheng, 2013).

2. **Deletion** - Loss of either one allele (loss of heterozygosity) or both alleles (homozygous deletion) (Didžiapetrienė et al., 2011; Mizoguchi et al., 2011).

3. **Amplification** - Increase in copy number of the genes (Chang et al., 2014; Matsui et al., 2013).

4. **Point mutation** - Mutation in a single base, leading to its alteration (Wang et al., 2013; Tan et al., 2015).

5. **Frameshift variant** - Insertion or deletion, leading to alteration of the open reading frame of the protein (Laghi et al., 2002; Korff et al., 2008).

6. **Inversion** - Breakage and joining of chromosomal segment in the reverse orientation (Schmid and Roth, 1983).

7. **Insertion** - Addition of chromosome segment into another (Lafay- Cousin et al., 2004).

8. **Viral insertion** - Integration of viral genome in host, resulting in activation of oncogenic processes, e.g., Hepatitis B virus, Human papilloma virus (Hai et al., 2014; Chandrani et al., 2015)
Figure 1.5: Chromosomal alterations

Different types of chromosomal alterations observed in genes associated with cancers: a- Deletion; b- Amplification; c- Inversion; d- Insertion; e- Translocation; f- Point mutation; g- Viral integration.
1.3.2.2 Epigenetic alterations:

Epigenetics comprise of heritable gene activity and expression without alterations in the DNA sequence (Goldberg et al., 2007; Bird 2007). Epigenetics are important in regulating a multitude of cellular processes, either turning genes “on” of “off”, leading to different functionalities among tissues (Jaenish and Bird, 2003; Lau et al., 2004). Cells have evolved four systems to promote gene silencing, viz. a) DNA methylation, b) Histone modifications, c) RNA mediated gene silencing and d) Alternate splicing (Egger et al., 2004).

a) DNA methylation: DNA methylation is a process whereby a methyl group is added to a cytosine nucleotide immediately adjacent to a guanine residue in the eukaryotic DNA by one of three enzymes called DNA methyltransferases (DNMTs). This site, known as CpG (Egger et al., 2004; Jones and Baylin, 2002; Robertson, 2002), is present sparsely in prokaryotes (Wojciechowski et al., 2013). In mammals, methylation, although sparse is a global phenomenon, distributed at specific CpG sequences throughout the genome, with the exception of CpG islands (CpGIs), almost 45,000 of which present in the haploid genome in humans (Antequera and Bird, 1993). Almost 72% of the promoters are associated with high CpGs (Saxonov et al., 2005), 70% of which are mostly methylated (Strichman-Almashanu et al., 2002). Insertion of the methyl group alters the appearance and structure of DNA, thereby participating in stand recognition during DNA replication, altering the binding of transcription factors in the promoters of the genes and ultimately affecting the transcription machinery, or in X chromosome inactivation and genetic imprinting (Bird, 2002; Cheng et al., 2013).

Promoter methylation leads to gene silencing, either by direct inhibition of transcription factor and subsequently RNA
polymerase II binding or through methyl-binding domain (MBD) proteins that recruit chromatin-modifying activities to methylated DNA (Deaton and Bird, 2011). In normal cells, promoter methylation plays a decisive role in cellular differentiation and development (Meissner et al., 2008; Kaaji et al., 2013; Shao et al., 2014). However, aberrant de novo methylation of CpG islands is an important hallmark of cancer (Jones and Baylin, 2002), mostly leading to silencing in TSGs (Phillips et al., 2008), Figure 1.6.

**Figure 1.6 : Mechanism of regulation of gene expression by promoter methylation**

![Diagram](image)

a- CpG island of the promoter is unmethylated, resulting in gene transcription (On); b- CpG island is methylated, resulting in prevention of gene transcription (Off).

b) **Histone Modifications**: Histones, the protein components of chromatin undergo post-translational modifications, thereby affecting chromatin arrangement and transcription of DNA. DNA-histone binding forms a compact and transcriptionally repressed structure (heterochromatin) and vice versa (euchromatin).
Histones can be modified by several means such as acetylation, methylation, ADP ribosylation, phosphorylation and ubiquitination (Goll and Bestor, 2002; Karlick et al., 2010; Bannister and Kouzarides, 2011). Knowledge about these histone modifications amalgamated into the “Histone code hypothesis”, which postulated that these modifications may be inter-dependent and may give rise to specific histone landscapes that provide access to proteins responsible for higher order chromatin organization and gene activation or inactivation (Chi et al., 2010).

c) **RNA-Associated Silencing**: Post-transcriptional gene silencing can be mediated by RNA when it forms antisense transcripts, microRNA (miRNA) or non-coding RNAs. These in turn might lead to formation of the transcriptionally inactive heterochromatin, or act as signals for histone modifications and DNA methylation (Hammond et al., 2001; Egger et al., 2004; Moazed, 2009; Malecová and Morris, 2010).

miRNAs are short, endogenous, single stranded, non-coding RNAs present in diverse organisms, including viruses, plants and animals (Valinezhad Orang et al., 2014). They regulate the protein expression of genes via binding to their complementary target messenger RNA (mRNA) “cognate mRNAs” (Djurandjovic et al., 2012; Valinezhad Orang et al., 2014). This event is achieved by pairing of the miRNA to the 3’ untranslated regions (UTRs) of the target genes, leading to the production of an effector RNA-induced silencing complex (RISC) (Figure 1.7) (Djurandjovic et al., 2012; Fukaya and Tomari, 2012), thereby inhibiting translation of proteins involved in a broad range of pathways including tissue development, cell division, cell proliferation, protein secretion, apoptosis, neuronal asymmetry, metabolism, stem cell properties, differentiation and viral infection (Valinezhad Orang et al., 2014).
miRNAs regulate the expression of the genes either through formation of RNA induced silencing complex (RISC) and direct translational blockage, or through degradation of target mRNAs. Image adapted from Joshi et al., 2011.

d) **Alternative Splicing** : Alternative splicing, whereby different exons of a gene are joined together during processing of RNA transcripts offers a certain degree of plasticity in production of mRNA, leading to production of functionally different proteins (Sebestyen et al., 2015) (Figure 1.8). Although this is a normal and universal cellular process, aberrant splicing often occurs in
cancer cells (Sebestyen et al., 2015; Zhang and Manly, 2013; David and Manley, 2010). Studies have shown that in several genes, many isoforms are specifically associated with the progression and metastasis of cancer (Oltean and Bates, 2013). Further studies might shed light on the transcript isoform signature in different cancers and provide potential molecular targets for prognosis and therapy.

**Figure 1.8 : Alternative splicing**

Alternative splicing offers a plasticity to transcripts produced, thereby leading to production of several proteins from the same transcript and might be an important process in cancer. Image adapted from CRCL, France.
1.4 Multi step progression of cancer

Progression of cancer occurs in multifarious, clinically well-documented steps. Carcinogenesis involves two main phases: i) pre-malignant and ii) malignant stages. Normal tissues, through either genetic alterations or environmental factors gradually develop pre-malignant lesions in the order of inflammation, ulceration and increasing stages of dysplasia (mild < moderate < severe) and finally carcinoma-in-situ (Mariani et al., 2014; Sipponen and Hyvirinen, 1993; Keith and Miller, 2013), Figure 1.9. Subsequent alterations in one/few of the pre-malignant cells convert them into malignant clones which give rise to the primary tumour (Yokota, 2000). Further accumulation of genetic alterations confer to the initial tumour invasive and malignant properties (well differentiated < moderately differentiated < poorly differentiated). Therefore, tumours are biologically and genetically extremely heterogeneous due to differences in genetic alterations in each cancer cell (Yokota, 2000).

Figure 1.9 : Multi-stage progression of cancer

The stages involved during the process of conversion of normal tissues to pre-malignant and ultimately invasive squamous cell carcinoma encompasses several sub-stages in gradual succession.
The multi-step model of progression of cancer was explained as a step-wise accumulation of a critical load genetic alterations at the molecular levels (Ilyas et al., 1999). If the alteration occurs in the initial stage, it leads to an initiating genotoxic hit leading to clonal expansion (N  $\rightarrow$  I). Evidences however indicate that the initiating damage must be fixed genetically and further compounded. If the first event is biallelic and fixed, it would lead to an irreversible multiplication of the initiated clone (I  $\rightarrow$  T2), iteration of which leads to greater survival advantage and enhanced malignancy (T2  $\rightarrow$  T3  $\rightarrow$  T4). On the other hand, if the entry point occur at the later stage, then the tumour T4 is highly dissimilar from the normal tissue with all alterations incurred in the previous stage transitions. Therefore, the pathway through which normal cells attain malignancy consists of a nested subset of genetic alterations which might not always be visible phenotypically or microscopically (Scrale et al., 1990; Bodmer 1997; Farber and Cameron 1980), vide Figure 1.10.

**Figure 1.10 : Multi- step progression of cancer**

Progression of normal epithelium to invasive cancer involves an iteration of multiple smaller steps in succession. Formation of premalignant lesions involves initiation, promotion and progression which finally gives rise to malignant s. Image courtesy Ilyas et al., 1999.
Evolution of tumours follow two models, the inverted pyramid and the nexus (Ilyas et al., 1999). In the inverted pyramid model, the mutations are largely interdependent with early mutations providing internal selection pressures for later mutations (Figure 1.11 a). In contrast, in the nexus model, mutation are neither interdependent nor under selection pressures - once a stage has been passed, the mutations responsible for that stage are no longer necessary and their reversal does not affect the (Figure 1.11 b). However, recent advancements in technology have helped provide a better view into the fine details underlying the transformation of a normal cell to a pre-malignant clone and finally into a malignant tumour.

**Figure 1.11 : Models of the evolution of cancer**

Two models to explain the evolution of tumours: a- Inverted pyramid model, which predicts that alterations of cancer associated genes are interdependent with early mutations conferring selection pressures for later mutations. Image adapted from Ilyas et al., 1999; b- Nexus model, according to which, mutations are neither interdependent nor under selection pressure; after crossing a stage, mutations responsible are dispensable with their reversal not affecting tumour evolution. Image adapted from Ilyas et al., 1999.
1.5 Identification of cancer associated genes

Identification of candidate genes associated with the development and progression of cancer involves a) Identification of gross chromosomal regions with abnormalities, b) Fine genetic and physical mapping of the genes, and c) Narrowing down of candidate gene loci and confirmation of the candidate genes (Figure 1.12).

Figure 1.12 : Identification of cancer associated genes

Cancer associated genes can be identified in a step by step approach involving identification of gross chromosomal regions altered, followed by either molecular cytogenetic analysis or genetic linkage analysis. The results are followed by fine genetic and physical mapping of the altered regions, narrowing down and ultimate confirmation of the candidate genes by different approaches.
Identification of gross chromosomal regions with abnormalities:

a) Molecular cytogenetic studies: Involves visualization and identification of chromosomes and detection of various chromosomal abnormalities such as anomalous chromosome number, amplifications, translocations, deletions, inversions, etc. The different methods used are:

1. **G banding** - It involves harvesting and staining mitotic chromosomes with Giemsa stain (Bickmore, 2001; Moore and Best, 2001).

2. **Fluorescence in situ hybridization (FISH)** - It involves binding of a fluorescence labelled DNA or RNA probe to a target sequence in the genome with high sensitivity and specificity (Bishop, 2010; Kearney, 2012).

3. **Spectral Karyotyping (SKY)** - Similar to FISH but refined in the use of several fluorescent probes covering the entire genome, Fourier spectroscopy, charged coupled device (CCD) imaging and optical microscopy for better, digital visualization and resolution (Guo et al., 2014; Abdel-Rahman et al., 2000).

4. **Comparative genomic hybridization (CGH)** - Provides a global overview of chromosomal gains and losses by quantitatively measuring the green/ red fluorescence ratio of equally mixed test DNA (labelled with a green fluorochrome) and control normal DNA (labelled red), hybridized to normal metaphase preparations. Resolution is about 10Mb (Weiss et al., 1999).

b) **Genetic linkage analysis**: It is used to map genetic loci based on observations of related individuals, when candidate genes are localized based on their proximity to known chromosomal markers (Lobo and Shaw, 2008). It can be applied to both major gene disorders and complex diseases (Teare and Barrett, 2005).
1.5.2 Fine genetic and physical mapping of the genes:
To narrow down specific regions by a) DNA polymorphism studies, b) Array based comparative genomic hybridization (CGH), c) Sequence tagged sites (STS) marker/Expressed sequence tags (EST) and d) Whole genome sequencing.

a) DNA polymorphism studies: An overview of the different markers used to study genetic polymorphisms is given in Table 1.1.

Table 1.1: Different Polymorphic markers used for mapping of the genes

<table>
<thead>
<tr>
<th>POLYMORPHISM</th>
<th>AVERAGE FREQUENCY</th>
<th>NO. OF ALLELES</th>
<th>LOCATION</th>
<th>INFORMATIVENESS</th>
<th>METHOD OF DETECTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA restriction fragment length polymorphism (RFLP)</td>
<td>Enzyme specific, eg., 1 in 2.9 Kb for Taq1.</td>
<td>2</td>
<td>Throughout genome</td>
<td>Low</td>
<td>Southern blotting, PCR</td>
</tr>
<tr>
<td>VNTRs or minisatellites (10-40bp motifs)</td>
<td>Several</td>
<td>Mostly at the chromosomal ends</td>
<td>High</td>
<td>Southern blotting, PCR</td>
<td></td>
</tr>
<tr>
<td>SSRs or microsatellites (1-6 bp motifs)</td>
<td>1 in 6Kb</td>
<td>Several</td>
<td>Throughout genome</td>
<td>High</td>
<td>PCR, Sequencing</td>
</tr>
<tr>
<td>Single nucleotide polymorphism (SNP)</td>
<td>1 in 400 bp</td>
<td>2</td>
<td>Throughout genome</td>
<td>Low</td>
<td>PCR, Sequencing</td>
</tr>
</tbody>
</table>

Different polymorphic markers used in the identification of cancer associated genes are shown, along with their frequency of prevalence in the human genome, number of alleles available, genomic distribution, informativeness and method of their detection.
1. **RFLP marker**
   Spacing: 0.7 cM apart.
   Limitations:
   a. Low polymorphism, so low in informativeness.
   b. Dependent on availability of restriction endonucleases, no whole genome might not be equally represented.
   c. Expensive and time consuming.

2. **VNTR marker**
   Limitation: Tends to cluster near the distal ends of chromosomes.

3. **Microsatellite marker** - Microsatellites, alias simple sequence repeats (SSRs) or short tandem repeats (STRs) are tandemly repeated DNA sequences with repeat length from 2 to 6 base pairs (bp), encompassing a total length up to 60 bp (Weber and Wong, 1993). Microsatellites are numerous, comprising of approximately 3% of the entire human genome (Saway et al., 2013) with about $5 \times 10^5$ in number and uniformly distributed throughout the human genome with an estimated average spacing of 6 kb (Weber et al. 1993). They are found in both coding and non-coding regions with some repeats predominant in specific genomic regions. Microsatellites undergo replicative slippage, offering them a dynamic polymorphic advantage on the evolutionary scale due to their expansion and contraction (Subramanian et al., 2003). These polymorphism are intricately related to the process of mutation as microsatellites have a high mutation rate, about $10^{-3}$ to $10^{-5}$ per generation (Payseur et al., 2010). In the light of the above, microsatellite instability (MIN) is of paramount importance, not only in evolutionary, population (Pemberton et al., 2013; Aimé et al., 2014; Putman and Carbone, 2014) and forensic (Planz and Hall, 2010; Filoglu et al., 2014; Zhang et al., 2014) studies, but
also as an important cause behind several diseases including multiple types of cancers (Kang et al., 2015; McIver et al., 2014; V et al., 2014; Amankwah et al., 2013; Arai et al., 2013; Bennett et al., 2009) and identification of susceptible/risk alleles associated with diseases (Kim et al., 2013).

The polymorphic character of microsatellites also finds widespread use in the identification of putative suppressor genes (TSGs) through allelotyping of the DNA with respect to normal tissues. In this method, PCR based amplification of the normal paired DNA, allelic losses are represented as loss of heterozygosity (LOH), whereas alteration of the size of either one or two alleles are represented as microsatellite size alterations (MAI, MAII respectively) (Figure 1.13) (David et al., 2014). Study of these alterations might be useful in associating such with cancer development.

**Figure 1.13 : Microsatellite markers for study of deletion**

Microsatellite markers are highly heterozygous due to different length in tandem repeats. Hence, different types of alterations can be observed in studies using microsatellite markers, such as loss of heterozygosity (LOH), microsatellite size alterations (MA1, MA2) and loss of heterozygosity in one allele and MA in the other alleles (LMA).
4. **SNP marker** - SNPs account for the 3.2 million differences in the diploid genome, conferring 0.1% dissimilarity between two individuals, arising from single base substitutions (Mayor, 2007). SNPs are the most frequent type of variation in the human genome (Wang et al., 1998) with almost 14-15 SNPs with an average distance of 300 bp (vide db SNP). Although majority of the SNPs are biologically inconsequential, a fraction of these confer diversity among individuals (Collins et al., 1997), Figure 1.14. Due to their polymorphic characteristics, they are important as genetic markers for tracing inheritance among generations (Johnson and Todd, 2000; Risch 2000), as well as for study of many diseases including cancers (Oh et al., 2015; Liu et al., 2014; Kim et al., 2013). However, majority of the diseases including cancers are multigenic, consisting of a combination of several SNPs that give rise to particular disease phenotype (Gray et al., 2000). Determining the types of SNPs associated with different forms of cancers might reveal novel insights into its management and drug development.

**Figure 1.14 : Single Nucleotide Polymorphism (SNP)**

*Single nucleotide polymorphism (SNP) arises due to difference in a single base in a particular locus, conferring genetic heterogeneity. Image adapted from NEI connection.*
b) **Array based CGH technique**: Array based CGH uses DNA microarrays for detecting copy number variation across a few kilobases in the entire genome, depending upon the platform used. Total genomic DNA, isolated from test and reference populations differentially labelled gives a relative hybridization intensity ideally proportional to the relative copy number of the sequences in the test and reference genomes (Pinkel and Albertson, 2005; Galizia et al., 2012). The highest resolution that can be obtained is one probe in 400bp (Mason-Suares et al., 2013).

c) **STS/ EST**: STS are short DNA sequences (200-500bp) of known sequence and location present uniquely in the genome that can be used to map adjacent candidate genes to chromosome physical maps (Rajneesh et al, 2010). EST are the expressed STS, generated from sequencing the 5’ and/or 3’ ends of cDNAs for mapping of candidate genes in the genome (Feichtinger et al., 2014).

d) **Whole genome sequencing**: Alterations mapped by whole genome sequencing help in identifying genes involved in biological processes including diseases like cancer. Sequencing, especially through genome wide association studies (GWAS) helps in providing a bulk yet cost effective method of genetic mapping and provides extensive information such as discovery of SNP, mutations and identification of sequence variants (Leshchiner et al., 2012; Chang et al., 2014).
1.5.3 Narrowing down of candidate gene loci and confirmation of the candidate genes:

Candidate genes can be identified and validated by a) Mutational analysis, b) Methylation analysis and c) Functional analysis.

a) Mutational analysis: Identification of mutations in the gene at different stages during progression of the disease might help in identifying the importance of the gene in that disease. Mutations can be gain-of-function, i.e., the function of the coded proteins might be enhanced or new functions acquired, or loss-of-function, i.e., the protein loses its designated function (Antoniou et al., 2014). Mutations can also be classified as dominant, those that manifest their effect when only a single allele is mutated, or recessive, those which are effective only when both alleles are altered (Schneider et al., 1991).

b) Methylation analysis: Epigenetic modification of the genome, in the form of methylation of the promoter, especially those upstream of suppressor genes play an important role in regulating gene expression. Identification of genes with methylated promoters might help in identifying genes that are associated with disease phenotypes (Shi et al., 2015; Shenker et al., 2015).

c) Functional analysis: Functional analysis of the genes in terms of their expression, activity, stability, etc. in normal and affected tissues such as cancer provides an indication of the association of expression with the disease (Xu et al., 2013; Muller et al., 1992; Baresova et al., 2014). Functional differences between wild type and mutant variants of the protein also provides an idea about its importance in terms of disease progression and therapeutic intervention (Muller and Vousden, 2014)
1.6 Head and Neck Squamous Cell Carcinoma (HNSCC)

Head and neck squamous cell carcinoma (HNSCC) comprises of a group of cancers which develop in the mucosal linings of the upper aero-digestive tract, comprising of 1) the nasal cavity and paranasal sinuses 2) the nasopharynx 3) the hypopharynx, larynx, trachea and 4) the oral cavity and oropharynx, vide Image 1.15 (Jemal et al., 2011).

Figure 1.15: Sites affected by HNSCC

HNSCC comprises of a group of s which behave in a similar manner and affect the mucosal epithelial lining of organs including the nasal cavity, oral cavity, salivary gland, trachea, thyroid, skin, oesophagus, larynx and pharynx. Image adapted from Stadler et al., 2008.
1.6.1 Epidemiology

Head and neck cancers accounts for greater than 95% of all head and neck malignancies (Muir et. al., 1995). Squamous cell carcinoma constitutes greater than 90% of the cancers of the head and neck area (Rezende et al., 2010). HNSCC is the sixth most common cancer worldwide and the fifth leading cause of cancer related deaths (Shi et al., 2015), Figure 1.16. HNSCC presents more than 550,000 cases with around 300,000 deaths each year (Jemal et al., 2011). Globally, The male to female ratio ranges from 2:1 to 4:1 with a very dismal survival rate of 40-50 % (Leemans et al., 2011).

The Indian subcontinent harbours one- third of the global burden of this malignancy (Dikshit et al., 2012). In India alone, 2.5 lakhs new patients are diagnosed annually, among whom about three-fourths are in an advanced stage (Sharma et al., 2015). The age standardized incidence rate of oral cancer in India is 12.6 per 100,000 population with a rapid increase in the incidence rate in recent years (Dikshit et al., 2012). According to GLOBOCAN 2012, cancers in the head and neck are the most common in males and fifth most common in Indian females. Highest age- adjusted incidence rates was found in females in the Indian subcontinent (Sankaranarayanan et al., 1998). Among the HNSCCs, carcinoma of oral cavity and oropharynx predominates in our population, primarily due to the etiological factors associated with this malignancy (Sharma et al., 2015).
Figure 1.16: Global epidemiology of HNSCC

HNSCC is prevalent in all continents and in most countries across the globe. Image adapted from IARC, WHO.
1.6.2 Etiological factors

There are several risk factors associated with the development of HNSCC. Although traditionally 80-90% of the HNSCCs were attributed to the use of tobacco and alcohol (Sturgis and Cinciripini, 2007), the other major risk factors associated with this disease are poor dental hygiene, high-risk sexual behaviours, HPV and genetics (Burke et al., 2014). In cigarette smokers, the risk of developing HNSCC is 5-25 times higher compared to that of non-smokers, with risk 17 times greater for heavy smokers, smoking 80 or more cigarettes per day than others (Lewin et al., 1998). Moreover, the risk increases in a dose-dependent manner with frequency, duration and extent of exposure (Marur and Forastiere, 2008; Curado and Hashibe, 2009). Although there is a decreasing trend in cigarette smoking (CDC, 2009; Choi et al., 2014), the use of smokeless tobacco in the form of chewing tobacco, snuff and khaini manifests a global increase, probably leading to a lifetime addiction (Viswanath et al., 2010; Macha et al., 2011) and emerging as a major risk factor (Secretan et al., 2009; Zhou et al., 2013), especially in India (Dutta et al., 2014). Although heavy alcohol consumption is an independent risk factor for HNSCC, it has been shown to synergistically increase the risk of development of HNSCC in conjunction with tobacco (Hunter et al., 2005; Peters et al., 2005). Heavy frequent usage, defined as three or more drinks per day was associated with increased risk of cancers of the oropharynx, hypopharynx, and larynx (Hashibe et al., 2007). Chewing of areca nut, either alone or in combination with betel quid (paan) is an important risk factor in Asian countries, especially among Indian women (Muttagi et al., 2015). Additionally, HPV is an emerging risk factor, especially for oropharyngeal squamous cell carcinoma, conferring a subset of HNSCC and offering better prognosis than patients without HPV infection (Syrjänen, 2010; Zaravinos, 2014). Oncogenic HPV 16 showed highest prevalence in HNSCC, followed by HPV 18 and 33 (Michaud et al., 2014). The effect of HPV oncoproteins E6 and E7 in the development of HNSCC through inactivation of p53 and RB has been well documented (Rautava and Syrja`nen, 2012; Scheffner et al., 1990), Figure 1.17.
Figure 1.17: Etiological factors associated with the development of HNSCC

Several etiological factors are responsible for the development of head and neck lesions including a- Alcohol; b- Tobacco smoking; c- Betel nut, alone or in combination with paan; d- HPV; e- Chewing of pan masala with/without betel nut/ tobacco; f- Chewing tobacco; g- Poor dental hygiene, etc.
1.6.3 Progression of head and neck lesions

a) Clinical progression

Several studies have indicated that HNSCC arises from a normal cell which undergoes transformation into a pre-malignant progenitor. This transformed cell, through an array of cumulative genetic/epigenetic and phenotypic/histological alterations ultimately expand into clones of malignant cells (Leemans et al., 2011).

Normal oral epithelium consists of three distinctly identifiable zones. The lowermost monolayer is the basal layer, which harbours stem-like undifferentiated and proliferative cells. The parabasal layer, comprising of the next three layer of cells, is made up of partially differentiated and proliferative cells. The remaining top layer of cells make up the spinous layer, consisting of mostly differentiated and dead cells, with or without a keratin layer.

Clinically identifiable pre-malignant lesions of head and neck can be either leukoplakia or erythroplakia. Leukoplakia describes a white lesion in the mucosal membrane, usually associated with mucosal thickening, while erythroplakia describes a red mucosal lesion. Speckled leucoplakia consists of a mixture of red and white lesions (Wenig, 2002).

Histologically, a normal oral epithelium undergoes an array of microscopically visible changes during conversion to pre-malignant and invasive tumours. The first alteration is hyperplastic epithelium with/without keratinization, which converts to increasing grades to dysplastic epithelium (mild, moderate, severe), carcinoma in situ (CIS) and finally progressive grades of invasive squamous cell carcinoma (well differentiated, moderately differentiated and poorly differentiated).

Leukoplakia, mostly not associated with premalignant or malignant lesions includes an array of histological changes ranging from an increase surface keratinization without dysplasia to invasive
keratinizing squamous carcinoma, enclosed within well-defined margins and possessing very little risk of developing a malignancy. In contrast, erythroplakic lesions are generally associated with grievous histopathological alterations, ranging from severe dysplasia, CIS to invasive cancers.

Several studies have demonstrated that on chronic exposure to genotoxic stresses like tobacco and areca nut, the oral mucosa undergoes myriad clinico-pathological manifestations like inflammation, ulceration, oral submucous fibrosis (OSF), lichen planus, haemangioma, etc., which, if unchecked might develop into pre-malignant lesions (Alshadwi and Bhatia, Yardimci et al., 2014; Ali et al., 2014).

**Inflammation** is a complex biological response of the tissues to external agents such as pathogenic microorganisms, chemical irritants, cellular damages, etc. (Ahmed, 2011). The process is initiated by resident immune cells at the affected site, such as macrophages and dendritic cells, which recognize foreign substances through specific receptors on their surface called pattern recognition receptors. Then begins a cascade of reactions including secretion of inflammatory mediators, migration of leucocytes, mainly neutrophils and macrophages and activation of the complement system to counter the irritant. Although inflammation is a systemic response to combat and protect the tissues from external agents, chronic, deregulated inflammation, more often than not plays an important role in the development of cancer (Chai et al., 2015). **Ulceration**, or the corroding away of the mucous membrane may also occur, often accompanied by inflammation, and if prolonged, can contribute to cancer formation (Chu et al., 2013). **Oral submucous fibrosis (OSF)**, a common and chronic insidious infliction affecting people in the Indian subcontinent is a well-known pre-malignant condition (Ali et al., 2014). OSF is associated with juxta-epithelial inflammatory reaction followed by fibroblastic changes in the lamina propria, along with epithelial atrophy, resulting in stiffness of the oral mucosa causing trismus and difficulty
in eating. Histologically, epithelial atrophy, dysplasia and fibrosis are observed (Alshadwi and Bhatia, 2012). **Oral lichen planus** is a chronic inflammatory disease, primarily mediated by T cell mediated autoimmunity. CD8+ T cells, which trigger an apoptotic response in the basal oral epithelium, leading to keratosis and disruption of the basal membrane (Lavanya et al., 2011). Figure 1.18 shows the stages involved.

**Figure 1.18 : Clinical progression of HNSCC**

Progression of normal oral epithelium to invasive squamous cell carcinoma occurs through several stages, shown in the order of their severity. 

- **a** - Normal tongue; 
- **b** - Inflammation; 
- **c** - Ulceration; 
- **d** - Lichen planus; 
- **e** - Oral submucous fibrosis; 
- **f** - Leucoplasia; 
- **g** - Erythroplakia; 
- **h** - Invasive squamous cell carcinoma. 

Picture courtesy Dr. Kabita Chatterjee, Oral and Maxillofacial Pathology Clinic.
b) **Field cancerization**

The concept of field cancerization developed based on the observation that even after surgical resection with margins histopathologically free of tumours recur, either locally or in remote sites, even in the absence of genotoxic stresses (Mohan and Jagannathan, 2014). Field cancerization, initially studied in depth through histological examination of oral tissues by Slaughter et al., 1953, hypothesised that the entire epithelium has an added risk for developing pre-malignant and malignant lesions due to multiple abnormalities along the entire tissue. Alterations in histologically normal tissues adjacent to tumours obtained from HNSCC patients supported this hypothesis, putting forward further questions whether multiple lesions develop independently from each other, or whether they developed from migrating progenitor cells (Oijen and Slootweg, 2000). Moreover, high incidence of recurrence in HNSCC was explained as effect of either molecular events affecting several cells from different locations simultaneously, or molecular events in a single progenitor that undergoes widespread clonal expansion and lateral spread (Aprna et al., 2013). The molecular basis of field cancerization has also been explained, which states that a stem cell, after acquiring genetic alterations expands into a clone of altered daughter cells called a “patch”. These patches can be recognized on the basis of specific alterations, e.g., mutation in TP53. The patch undergoes further expansion into a “field” with the aid of added genetic alterations which provide a growth advantage, thereby gradually displacing the normal mucosa. This extended field later on acts as the ground for local recurrence or new primary tumours (Braakhuis et al., 2003), Figure 1.19.
Field cancerization in oral tumours. a- Recurrence of the after surgical resection or evolution of a second primary explained on the basis of the existence of a “field” of histologically altered tissues. Image courtesy Leemans et al., 2011; b- Demonstration of the existence of field cancerization in tongue tissue through histopathological examination. Black arrow: Verrucous carcinoma; White arrowhead: Well differentiated squamous cell carcinoma; Black arrowhead: Verrucous hyperplasia; Circle: Severe dysplasia and carcinoma in situ; White arrow: Hyperkeratosis and acanthosis without dysplasia. Image courtesy Fortuna and Mignogna, 2011.
c) **Histopathological progression**

The first stage of alteration in pre-malignant lesions is **hyperplasia**. It is accompanied by thickening of epithelium, change of cellular orientation and architecture in the lower three layers from the basal epithelium and increase in cell size and number. Hyperplasia, although deviated from normal tissue architecture is a pathological response to certain tissue stress, albeit with cell growth under regulated conditions (Wenig, 2002). The next stage, atypia or **dysplasia** is encompasses cyto- morphological and maturation abnormalities like proliferation of immature cells with loss of cell polarity, nuclear pleomorphism, increased nuclear/ cytoplasmic index, greater/ uneven nuclear chromatin and large number of mitotic cells with/ without keratosis (Wenig, 2002). Based on the severity of aberration and fraction of abnormal cells, dysplasia is graded into **mild** (affecting lower 1/3 of epithelium), **moderate** (lower 2/3) and **severe** (entire epithelium) (Speight, 2007; Barnes et al., 2005). Severe dysplasia also demonstrates abnormal/ bulbous and branched rete pegs, the earliest sign of epithelial invasion. Although all grades of dysplasia are potentially reversible after cessation of instigating factors such as tobacco (Wenig, 2002), moderate dysplasia often becomes the deciding stage in malignancy (Smith et al., 2009; Dost et al., 2014).

**Carcinoma in situ** is regarded as the most severe form of epithelial dysplasia, characterized by abnormal cytological and architectural changes in the entire thickness of the epithelium, but confined in the intact epithelial layer. Although some pathologists consider CIS as a pre-malignant stage, others consider it as a malignant transformation without actual invasion (Speight, 2007).

The malignant stages include microinvasive and overt squamous cell carcinoma (SCC). Microinvasive SCC can occur by two unrelated mechanisms, one from a CIS, while the other in the absence of CIS (Wenig, 2002). Full blown SCC progresses through three severe grades, identified by microscopically combining three parameters of degree of keratinization, cellular/ nuclear pleomorphism and mitotic activity as
well differentiated, moderately differentiated (both grouped as low grade) and poorly differentiated (high grade) (Pereira et al., 2007). Cellular invasion is the diagnostic criteria, consisting of large cellular aggregates or single cells or small aggregates with desmoplasmic stromal response as a hallmark feature (Wenig, 2002) (Figure 1.20).
1.6.4 Modalities used in treatment of HNSCC

Over the years, multiple modes of treatment are available to combat head and neck cancer other than the traditional methods of surgery, radiotherapy and chemotherapy, either alone, or in combination, depending upon the stage and anatomical site, vide Figure 1.21 (Johnson, 2012). In recent years, newer methods of treatment, such as targeted therapy, immunotherapy, etc. are being used in conjunction with the above (American Cancer Society).

Figure 1.21: Modalities used in the treatment of head and neck lesions

Three widely used modalities of treatment of head and neck lesions, viz. surgery, chemotherapy and radiotherapy, used in various forms, either alone or in conjunction.
a) **Surgery**:

Surgery is mostly the choice of treatment for head and neck cancers. For small tumours, surgery can be used alone, whereas for large tumours, primary chemotherapy is followed by chemo-radiation (Campana and Meyers, 2006). Surgery can be further divided based on the purpose of its use, as:

1. **Preventive/prophylactic surgery**: To remove a tissue that might possibly develop a cancer, e.g. erythroplakia (Kuriakose, 2006).

2. **Diagnostic surgery/biopsy**: To diagnose presence/absence of cancer and determine grade of positive specimens (Richards, 2015).

3. **Curative surgery**: To remove cancer localized in only a single part of the body and subject to complete removal (Kuribayashi et al., 2012).

4. **Restorative/reconstructive surgery**: To restore the function of an organ, or improve the appearance of the person after the primary surgery (Demirkan et al., 1999).

b) **Chemotherapy**:

Chemotherapy, given both intravenously or orally is an important method of treating cancers and may be used to shrink tumours before surgery/radiation, or after surgery/radiation to kill residual cancer cells, or during recurrence of the cancer (Georges et al., 2014; Price and Cohen, 2012). Chemotherapy may be used in the following settings:

1. **Primary settings/neoadjuvant or induction chemotherapy**: Given in locally advanced or unresectable disease to reduce tumour bulk for allowing definitive surgery and preservation of oro-facial organs (Patil et al., 2014) (Figure 1.22).
2. **Adjuvant setting**: In combination with radiation for positive resection margin (Grau et al., 1996; Huang and O’ Sullivan, 2013).

3. **Palliative settings**: To prolong survival and ease symptoms of terminal cancer patients, but not to cure the disease (Kadakia et al., 2012; Archer et al., 1999).

4. **Salvage/ Rescue chemotherapy**: Chemotherapy given to patients when a patient does not respond to most treatment options and all other options are exhausted (Schwartz et al., 2000; Zygulska and Krzemienieck, 2015).

**Figure 1.22: Neoadjuvant chemotherapy**

Neoadjuvant chemotherapy reduces the bulk of the primary, locally advanced head and neck s prior to surgical resection, leading to less adverse surgery and preservation of oro- facial organs. a- Pre- therapy patient with large (indicated by yellow circle); b- Post- therapy image of the same patient with substantially reduced size (indicated by green circle); c- Post surgery, the entire is excised (pink circle). Image courtesy C.N.C.I.
c) **Radiotherapy:**

Radiotherapy in cancer treatment makes use of ionizing radiations in the form of photons (x-rays, gamma rays generated from radioactive Co, Cs or a linear accelerator) or particles (protons, neutrons, electrons, alpha and beta particles) to kill cancer cells. Based on the mode of application, radiotherapy can be broadly divided into the following:

1. **External beam radiotherapy/ teletherapy:** It is the most widely used form of radiotherapy, where a patient sits or lies in a designated area and an external source of radiation is directed to the affected site (Huang and O’ Sullivan, 2012; Wu et al., 2012).

2. **Brachytherapy:** It is an advanced and expensive modality whereby the radiation is applied in or near the tumour itself, thereby allowing a high radiation dose to reach the tumour with minimum exposure to surrounding tissues (Huang and O’ Sullivan, 2012; Petera et al., 2015).

1.6.5 **Molecular progression of HNSCC**

Several studies have identified numerous alterations associated with the development of HNSCC. Preliminary cytogenetic studies have indicated the progressive increase in chromosomal alterations from premalignant to malignant lesions (Hittelman et al., 1993; Sreekantaiah et al., 1994; Van Dyke et al., 1994; Saunders et al., 2000; Jin et al., 1995; Park et al., 2010) including gain of chromosomal arms in chromosomes 1p, 2q, 3q26, 5p14-15, 7p11, 7q, 8q11-12, 8q 24, 11q13, 20q12-13 and loss in chromosomes 1p, 2q, 3p11-12, 3p13-24, 4q21-31.3, 4q35, 5q12-23, 8p22-23, 9p21-24, 10p13-pter, 10q22-26, 11p, 11q23-qter, 12.11-12.12, 13q, 17p, 18q21-23, 19q, 21q11.2-23 (Ha et al, 2002; Wreesmann and Singh, 2005; Golin, 2014). Highest losses were reported in 3p and 9p21-22 (70% loss), while gains were mostly in chromosomes 17q (Glazer et
al., 2009). Array CGH across the entire genome of HNSCC more precisely identified the altered regions—deletion in 2p15, 3p12-14.3, 3p21-26, 4q34.4, 9p, 11q13.3, 16q23.2 and amplification in 3q23, 5p15.2, 7p11.2, 7p12.3-13, 7q21.2 and 7q35 in HNSCC (Baldwin et al. 2005; Garnis et al. 2004; Golin, 2014), Figure 1.23.

Figure 1.23: Molecular progression of HNSCC

Chromosomal regions involved in the molecular progression of HNSCC, arranged according to the stages encompassed. Pink box indicates the studies performed in our laboratory. RED: Deletion/ Methylation; GREEN: Mutation; VIOLET: Amplification.
Considering the genes and chromosomes affected during progression of HNSCC, normal mucosa undergoes 9p21 deletion, p16/ p14 inactivation, trisomy in chromosome 7, EGFR and telomerase activation for conversion to hyperplastic lesions. 3p deletions, 8q gain, 17p13 mutations and tetraploidy convert hyperplastic lesions to dysplasia, with deletions in chromosomes 3p12.3, 3p21.31, 3p22.1, 9p21-22, 9q22.3 and 11q22.3-24 in mild dysplastic lesions obtained from our laboratory (Chakraborty et al. 2003; Dasgupta et al. 2002; Ghosh et al. 2010; Ghosh et al. 2008; Mondal et al. 2003; Tripathi et al. 2003). For conversion of low grade to high grade dysplasia, 3p24, 17p, 9p21.3 and 11q3 deletions were necessary, along with PI3KC, RARβ, RB, p15, APC deletion, COX2 activation and CCND1, INT2, HST1 amplification (Golin et al., 2014, Riaz et al., 2014, Saranath et al., 1993). Further narrowing down the results to identify the candidate genes located in the above chromosomal regions using deletion mapping and epigenetic analysis revealed that LIMD1, ROBO1, ROBO2, RBSP3, SH3GL2, EGFR, SLIT2 and p16 were associated with the development of mild dysplastic lesions and the alterations of CDC25A, LTF and p15 were associated with the development of moderate dysplastic lesions. Further alterations of PHF2, RASSF1 and CACNAD2 were important for the progression from severe dysplasia to invasive squamous cell carcinoma (Ghosh et al. 2008, 2009; 2010; Maiti et al., 2013, 2015). Initial studies with deletion/ methylation had indicated that LIMD1 and RBSP3 were altered at mild dysplasia and CDC25A at moderate dysplasia (Ghosh et al., 2008; Ghosh et al., 2010). However, subsequent studies of protein expression revealed the association of LIMD1 with chronic ulceration and RBSP3 with hyperplastic lesions in pre-malignant lesions exhibiting field cancerization (Maiti et al., 2012). Additional 8p, 5q, 11p, 11q13, 13q21, 14p and 21q deletion, along with gain of 1p and 22q and aneuploidy converted these to carcinoma- in-situ. Final progression of these pre-malignant lesions to malignant squamous cell carcinoma require 4q, 5p, 6q, 10q23, 13q and 18q deletion and 3q26 amplification alongside PTEN inactivation and CMYC, RAS, ETS-1/ COT, uPA system, MMPs, Cathepsin and integrin overexpression, vide Figure 1.24 (Perez-Ordoñez
et al. 2006; Golin, 2014; Saranath et al., 1993). These results were further corroborated by microarray analysis from samples representing different stages of progression, which showed that in the transcriptional progression model, most of the alterations developed in pre-malignant stages, before development (Park et al. 2010). The results were validated by whole genome sequencing of HNSCCs by different researchers, along with providing newer insights into the progression of the disease (Cancer Genome Atlas Network, 2015; Riaz et al., 2014; Zhang et al., 2013).

**Figure 1.23 : Molecular genetic progression of HNSCC**

Genes altered during the molecular progression of HNSCC, arranged according to the stages of alteration. Pink box indicates the studies performed in our laboratory. **RED:** Deletion/ Methylation; **GREEN:** Mutation; **BLUE:** Overexpression; **VIOLET:** Amplification.
Among the genes identified, RBSP3 from 3p22.3 and CDC25A, LIMD1 from 3p21.31 are important, not only due to their myriad functional roles (Table 1.2- 1.4) but also due to their association with early dysplastic lesions of head and neck and alterations in different tumours (Table 1.5- 1.7).

Table 1.2 : Alterations of RBSP3 in different cancers

<table>
<thead>
<tr>
<th>GENE</th>
<th>TYPE OF ALTERATION</th>
<th>CANCER</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBSP3</td>
<td>Deletion/ reduced expression/ promoter methylation</td>
<td>HNSCC</td>
<td>Ghosh et al., 2010; Maiti et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Deletion/ reduced expression/ promoter methylation/ splice variants</td>
<td>Cervical cancer</td>
<td>Anedchenko et al., 2007; Mitra et al., 2010</td>
</tr>
<tr>
<td></td>
<td>Reduced transcript/ promoter methylation</td>
<td>Non- small cell lung cancer</td>
<td>Anedchenko et al., 2008; Senchenko et al., 2010</td>
</tr>
<tr>
<td></td>
<td>Deletion/ promoter methylation/ reduced expression</td>
<td>Breast cancer</td>
<td>Sinha et al., 2008</td>
</tr>
<tr>
<td></td>
<td>Somatic mutations</td>
<td>Multiple cancers</td>
<td>Kashuba et al., 2009</td>
</tr>
<tr>
<td></td>
<td>Splice variants/ mutations</td>
<td>Epithelial cell lines</td>
<td>Kashuba et al., 2004</td>
</tr>
</tbody>
</table>

Studies showing the alterations of RBSP3 in different cancers

Table 1.3 : Alterations of LIMD1 in different cancers

<table>
<thead>
<tr>
<th>GENE</th>
<th>TYPE OF ALTERATION</th>
<th>CANCER</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIMD1</td>
<td>Deletion/ promoter methylation/ mutation/ reduced expression</td>
<td>HNSCC</td>
<td>Ghosh et al., 2008; 2010; Maiti et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Deletion/ promoter methylation</td>
<td>Small cell non-small cell lung cancer</td>
<td>Sharp et al., 2008</td>
</tr>
<tr>
<td></td>
<td>Phosphorylation, altered cellular localization</td>
<td>Breast cancer</td>
<td>Huggins and Andrulis, 2008</td>
</tr>
<tr>
<td></td>
<td>Reduced expression</td>
<td>Acute leukemia</td>
<td>Liao et al., 2015</td>
</tr>
</tbody>
</table>

Studies showing the alterations of LIMD1 in different cancers
Table 1.4: Alterations of CDC25A in different cancers

<table>
<thead>
<tr>
<th>GENE</th>
<th>TYPE OF ALTERATION</th>
<th>CANCER</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC25A</td>
<td>Deletion/ mutation/ reduced expression</td>
<td>HNSCC</td>
<td>Ghosh et al., 2008; Maiti et al., 2012, 2013</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Non-small cell lung cancer</td>
<td>Wu et al., 1998</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Ovarian cancer</td>
<td>BrogGINI et al., 2000</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Breast cancer</td>
<td>Evans, 2000; Cangi et al., 2008</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Thyroid cancer</td>
<td>Ito et al., 2002</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Hepatocellular cancer</td>
<td>Xundi et al., 2003</td>
</tr>
<tr>
<td></td>
<td>Overexpression</td>
<td>Colorectal cancer</td>
<td>Huang et al., 2011</td>
</tr>
</tbody>
</table>

Studies showing the alterations of CDC25A in different cancers

Table 1.5: Functions of RBSP3

<table>
<thead>
<tr>
<th>GENE</th>
<th>FUNCTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>RBSP3</td>
<td>It dephosphorylates RB at serine 807/811, thereby increasing RB-E2F interaction and halting the cell cycle at G1/S boundary</td>
<td>Kashuba et. al, 2004</td>
</tr>
<tr>
<td></td>
<td>Inactivates RNA polymerase-II by preferential dephosphorylation of ‘Ser-5’, thus controlling the transcription machinery</td>
<td>Yeo et. al., 2003</td>
</tr>
<tr>
<td></td>
<td>RBSP3 might function as a transcriptional co-repressor, inhibiting transcription of neuronal genes in non-neuronal cells.</td>
<td>Yeo et. al., 2005</td>
</tr>
<tr>
<td></td>
<td>Might act as a phosphatase of Smad1, Smad2/3</td>
<td>Wu et. al., 2009</td>
</tr>
<tr>
<td></td>
<td>Might act as a phosphatase of Snail.</td>
<td>Sapkota et. al., 2006</td>
</tr>
</tbody>
</table>

Studies showing the functions of RBSP3 in different cellular processes
<table>
<thead>
<tr>
<th>GENE</th>
<th>FUNCTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Binds to pRB and represses E2F-mediated transcription</td>
<td>Sharp et al., 2004; Mayank et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Mediates osteoclast differentiation by binding to Traf6, a regulator of RANK-L regulated osteoclast development.</td>
<td>Feng et al., 2007</td>
</tr>
<tr>
<td></td>
<td>Regulates stress osteoclastogenesis by interaction with p62/sequestosome protein, as well as osteoblast function and osteoblast progenitor commitment</td>
<td>Luderer et al., 2008</td>
</tr>
<tr>
<td></td>
<td>Acts as a molecular scaffold to simultaneously bind prolyl hydroxylases and VHL, enabling HIF-1α degradation</td>
<td>Foxler et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Induces degradation of HIF-1α after dimerization with RHOBTB3 and formation of a RHOBTB3/LIMD1-PHD2-VHL-HIF-1α complex.</td>
<td>Zhang et al., 2015</td>
</tr>
<tr>
<td></td>
<td>Participates in micro-RNA mediated gene silencing</td>
<td>James et al., 2010</td>
</tr>
<tr>
<td></td>
<td>Involved in the phenomenon of multiple drug resistance in colorectal cancer cells.</td>
<td>Chen et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Binds to LATS1, promoted by JNK, thus acting as a link between JNK and Hippo signalling.</td>
<td>Sun and Irvine, 2013</td>
</tr>
<tr>
<td></td>
<td>Regulates mechanical strain through regulation of JNK and Hippo signalling.</td>
<td>Codelia et al., 2014</td>
</tr>
<tr>
<td></td>
<td>Regulates cell migration and cytoskeletal signalling</td>
<td>Bai et al., 2011</td>
</tr>
</tbody>
</table>

Studies showing the functions of LIMD1 in different cellular processes
<table>
<thead>
<tr>
<th>GENE</th>
<th>FUNCTION</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC25A</td>
<td>Dephosphorylates CDK4/6 and promotes G1/S transition of the cell cycle.</td>
<td>Bertero et al., 2013</td>
</tr>
<tr>
<td></td>
<td>Activates both Cdk1–2/Cyclin A and Cdk1/Cyclin B complexes in mitosis</td>
<td>Timofeev et al., 2009</td>
</tr>
<tr>
<td></td>
<td>Activation of a cell cycle pathway-mediated death signal in neurons in conjunction with CHK1</td>
<td>Zhang et al., 2006</td>
</tr>
<tr>
<td></td>
<td>Androgen corepressor, repressing its downstream activity in prostate cancer cells</td>
<td>Chiu et al., 2009</td>
</tr>
<tr>
<td></td>
<td>Regulates both G1-S and G2-M transitions by dephosphorylation of cyclin dependet kinases.</td>
<td>Ray and Kiyokawa, 2008</td>
</tr>
<tr>
<td></td>
<td>DNA damage response and cell cycle arrest mediated by CHK1 and CHK2.</td>
<td>Ray and Kiyokawa, 2008</td>
</tr>
<tr>
<td></td>
<td>Dephosphorylates ERK and inhibits its downstream signalling.</td>
<td>Wang et al, 2005</td>
</tr>
<tr>
<td></td>
<td>Promotes cell survival by stimulating NF-κB activity through IκB-α phosphorylation and destabilization</td>
<td>Hong et al., 2012</td>
</tr>
<tr>
<td></td>
<td>Dephosphorylates EGFR and inhibits its signalling cascade.</td>
<td>Wang et al., 2002</td>
</tr>
<tr>
<td></td>
<td>Dephosphorylates RAF-1 and perturbs its signalling.</td>
<td>Xia et al., 1999</td>
</tr>
<tr>
<td></td>
<td>Enhances Foxo1 stability and promotes tumour metastasis.</td>
<td>Feng et al., 2011</td>
</tr>
<tr>
<td></td>
<td>Down-regulation leads to muscle cell differentiation</td>
<td>Sarkar et al., 2010</td>
</tr>
</tbody>
</table>

Studies showing the functions of CDC25A in different cellular processes
Thus, it is evident from previous reports that these genes undergo frequent alterations (deletion/ promoter methylation/ mutation/ reduced expression) in HNSCC, the results similar to those obtained by other researchers, except for CDC25A. The paradoxical reduced expression of CDC25A in HNSCC compared to others might be due to the influence of etiological factors or due to other regulatory mechanisms yet unknown.

Functionally, the proteins have an array of purposes, although their functions are linked through regulation of RB and monitoring of the cell cycle (Figure 1.25). While RBSP3 is involved in controlling the cell cycle, control of the transcription machinery, prevention of epithelial to mesenchymal transition, etc., LIMD1, besides regulation of the cell, is also involved in osteoclastogenesis and their differentiation, response to hypoxia, gene silencing, maintenance of cytoskeletal structure, etc. CDC25A is similarly involved in regulation of the cell cycle, along with perturbation of different signalling pathways like EGFR, ERK, RAF-1, etc. (Figure 1.26).

In terms of expression of the genes, RBSP3 showed low protein expression in the basal layer of normal oral epithelium with expression increasing towards spinous layers, whereas LIMD1 and CDC25A proteins show expression throughout, with expression of all three genes being lost during development of HNSCC, vide Figure 1.27 (Maiti et al., 2012).
Figure 1.25: Functions of the candidate genes in the cell cycle

RBSP3, LIMD1 and CDC25A are important mediators of the cell cycle. LIMD1 stabilizes RB-E2F interaction in the early G1 phase. For entry into the cell cycle, CDC25A activates Cyclin/CDK complexes by its phosphatase activity, thereby resulting in phosphorylation of RB, freeing of E2F for transcription of cell cycle proteins and subsequent progress of the cell cycle. After completion of the cell cycle, for prevention of untimely re-entry into the next cycle, RB is dephosphorylated by different proteins including RBSP3, thus leading to its reinteraction with E2F and blocking of the cell cycle.
The candidate genes play important roles in different cellular processes, although their functions are linked through their roles in regulation of RB and subsequent control of progression of the cell cycle. The different functions of the genes are shown in the figure.
RBSP3, LIMD1 and CDC25A show distinct patterns of expression in the normal oral epithelial tissues. While RBSP3 showed low expression in the basal, proliferative layer with the expression increasing in the spinous layers, LIMD1 and CDC25A showed expression throughout the entire epithelium.