Chapter 1

General Introduction
1.1 “Cancer”- Molecular basis of the disease

1.1.1 Cancer: A genetic disease

Cancer is a genetic disease—that is, cancer is caused by certain changes to genes that control the way our cells function, especially how they grow and divide. In carcinogenesis, a single cell sequentially accumulates genetic or epigenetic alterations giving rise to a clonal progeny with selective growth advantage and transformed phenotype (Nowell 1976). This evolved population of cells has innate propensities and abilities to disregard the normal controls of proliferation and territory, emerging as a cancer (Ponder 2001). The phenotypes of these neoplastic cells have been well characterized by Hanahan and Weinberg (2000, 2011) (Figure 1), as “hallmark features of cancer” which includes mainly six capabilities upon which the pathogenesis of human cancer cell depends. Changes in a cell, that direct transformation toward a malignant phenotype, include gain-of-function mutations that activate oncogenes, loss-of-function mutations that inactivate tumor suppressor genes and mutations that inactivate stability genes, each conferring a specific growth advantage to the cell, which leads to progressive conversion of normal human cells into cancer cells (Romero-Garcia et al., 2011; Karakosta et al., 2005; Hanahan and Weinberg, 2000). Point mutations, chromosomal translocations, deletions, promoter methylation, gene amplification and numerical chromosomal changes are some of the genetic and epigenetic alterations of the genes associated with cancer. (Barrett, 1993; Suzuki et al., 2008). Recently, non-coding RNAs, such as microRNAs, have also been reported to be involved in the signaling networks associated with cancer. In addition, the results obtained from the Human Genome Project (HGP) and genome-wide associated studies (GWAS), revealed numerous variations in single nucleotide polymorphisms (SNPs) and gene signatures are associated with increased cancer risk (Cao et al 2011).


**Figure 1.1:** The hallmark features of cancer (adapted from Hanahan and Weinberg 2011)

### 1.1.2 Hallmarks of cancer

The cancer cell were characterized by six essential alterations in cell physiology that collectively dictate malignant growth (Hanahan and Weinberg, 2000; 2011). These common and basic mechanisms, widely known as the “Hallmarks of Cancer” (Figure 1.1). These six manipulations of cancer cell genotypes were:

#### 1.1.2.1 Sustaining proliferative signaling

Normal cells can regulate the growth promoting signals and maintain the balance between cell growth and division and ensuring the homeostasis of the cell. Cancer cells by deregulation these signals become ‘Masters of their own destinies’. These disregards to the growth signals can take place by somatic mutation which activates additional downstream pathways and disrupting negative feedback mechanism (Figure 1.1). Activation of growth factors like PDGF (platelet-derived growth factor) and TGF-α.
(tumor growth factor α) (Fedi et al., 1997) and upregulation of receptor tyrosine kinases (RTK) like epidermal growth factor receptor (EGF-R/erbB), HER2/neu receptor (Slamon et al., 1987; Yarden and Ullrich, 1988) are associated with acquiring self-sufficiency in growth signals. They are also called oncogenes.

1.1.2.2 Evading growth suppressors
Cancer cells also disobey the negative signals that regulate the growth and proliferations of the cells. Many of these programs depend on the actions of tumor suppressor genes. They operate as central control nodes within two key complementary cellular regulatory circuits that govern the decisions of cells to proliferate or, alternatively, activate senescence and apoptotic programs (Figure 1.1). Inactivation of TP53, RB, TGF-β are some of the important aspects in tumorigenesis (Hannon and Beach, 1994; Datto et al., 1997; Weinberg, 1995).

1.1.2.3 Resisting cell death
The programmed cell death by apoptosis serves as a natural barrier to cancer development. Tumor cells evolve a variety of strategies to limit or circumvent apoptosis resulting expansion of a population of neoplastic cells (Hanahan and Weinberg, 2011). Cancer cells can circumvent the apoptotic trigger and that acquired resistance toward apoptosis is a hallmark of most and perhaps all types of cancer (Figure 1.1). In tumor, most common event is the loss of TP53 tumor suppressor function, which eliminates this critical damage sensor from the apoptosis-inducing circuitry. Alternatively, tumors may achieve similar ends by increasing expression of antiapoptotic regulators (Bcl-2, Bcl-xL) or of survival signals by downregulating proapoptotic factors (Bax, Bim, Puma).

1.1.2.4 Enabling replicative immortality
Telomere shortening is one of the phenomena that threaten cell viability. Telomerase, the specialized DNA polymerase that adds telomere repeat segments to the ends of telomeric DNA. This is expressed significantly in most of the immortalized cells including the human cancer cells. The presence of telomerase activity in immortalized cells is
correlated with a resistance to induction of both senescence and crisis/apoptosis (Cong and Shay, 2008) (Figure 1.1).

1.1.2.5 Inducing angiogenesis
Like normal tissues, tumors require sustenance in the form of nutrients and oxygen as well as an ability to evacuate metabolic wastes and carbon dioxide. The development of neo-vasculature near the tumor cells is known as angiogenesis, which addresses these needs. During tumor progression, an “angiogenic switch” is almost always activated and remains on, causing normally quiescent vasculature to continually sprout new vessels that help sustain expanding neoplastic growths (Hanahan and Folkman, 1996) (Figure 1.1). The expression of proangiogenic factor VEGF is stimulated by hypoxic intra-tumor microenvironment. The VEGF mediates intricate interplay between various extracellular signaling pathways like notch, hedgehog, SLIT2-ROBO1 etc. (Dickinson and Duncan, 2010; Sheng et al., 2004; Thurston and Kitajewski, 2008), contributing towards angiogenesis. The Hypoxia inducible factor (HIF-1α) is a crucial player of tumor angiogenesis. Its expression is invoked by the SLIT-ROBO pathway via PI3K/AKT activation (Wang et al. 2003).

1.1.2.6 Activating invasion and metastasis
Carcinomas from epithelial tissues having the potential in local invasion and distant metastasis are at the higher pathological grades of malignancy (Figure 1.1). The epithelial cells of the tumor develops a mesenchymal properties by a developmental regulatory program known as the “epithelial-mesenchymal transition” (EMT) and acquire the abilities to invade, resist apoptosis and disseminate (Klymkowsky and Savagner, 2009; Polyak and Weinberg, 2009). The most widely observed alteration in cell-to-environment interactions in cancer involves E-cadherin, a homotypic cell-to-cell interaction molecule ubiquitously expressed on epithelial cells.

Along with these six hallmarks characteristics, Hanahan & Weinberg proposed two enabling characteristics of these hallmark properties by which their acquisition is made possible as well as two emerging hallmark characters of the cancer cells in 2011.
The enabling characteristics

1.1.2.7 Genomic instability and mutability
The defects in genome maintenance and repair are selectively advantageous to tumor progression. Normal cells can repair any error occurred in the genome in any stage of cell cycle. Cancer cells having defects in this repair machinery imparts “mutator phenotype” characterized by elevated frequencies of spontaneous mutations and increased microsatellite instability (MSI) (Loeb and Harris, 2008a). Though environmental parameters determine which mutations are selected. Mutations that have already occurred may influence selection pressures associated with subsequent mutations. The mutations may interact to enhance or neutralize the effects of each other. The collaborative interaction may be represented by two models i.e. inverted pyramid and nexus model (Ilyas et al., 1999). Along with mutations epigenetic modifications like DNA methylation and histone modifications also play important role in inactivation of instability screening mechanisms of the cells (Berdasco and Esteller, 2010; Esteller, 2007; Jones and Baylin, 2007). CGH analysis provides clear evidence of presence of some recurrent aberrations (both amplification and deletion) in the particular sites of the genome which codes for the genes whose alterations prefer neoplastic growths (Korkolka and Gray, 2010). Hence, genomic instability is undoubtedly an enabling characteristic to acquire the hallmark capabilities (Figure 1.1).

1.1.2.8 Tumor promoting inflammation
A second enabling characteristic involves the inflammatory state of premalignant and malignant lesions driven by the immune system cells, which serve to promote tumor progression through various means. The immune cells largely of the innate immune system have a neoplastic progression (DeNardo et al., 2010; Grivennikov et al., 2010; Qian and Pollard, 2010; Colotta et al., 2009). Inflammation can contribute to multiple hallmark capabilities by supplying bioactive molecules to the tumor microenvironment. Some of them are growth factors that sustain proliferative signaling, survival factors that limit cell death, pro-angiogenic factors, extracellular matrix-modifying enzymes that facilitate angiogenesis, invasion, and metastasis, and inductive signals that lead to
activation of EMT and other hallmark-facilitating programs (DeNardo et al., 2010; Grivennikov et al., 2010; Qian and Pollard, 2010; Karnoub and Weinberg, 2006–2007) (Figure 1.1).

The emerging hallmarks

1.1.2.9 Reprogramming energy metabolism
The uncontrolled cell proliferation of neoplastic disease involves not only deregulated control of cell proliferation but also the energy metabolism adjustments of the cells to fuel cell growth and division. Under aerobic conditions, normal cells process glucose, first to pyruvate via glycolysis in the cytosol and thereafter to carbon dioxide in the mitochondria. Under anaerobic conditions, glycolysis is favored and relatively little pyruvate is dispatched to the oxygen-consuming mitochondria. Otto Warburg first observed an anomalous characteristic of cancer cell energy metabolism (Warburg, 1930, 1956a, 1956b). Even in the presence of oxygen, cancer cells can reprogram their glucose metabolism largely to glycolysis, leading to a state called ‘‘aerobic glycolysis.’’ Glycolytic fueling has been shown to be associated with activated oncogenes (e.g., RAS, MYC) and mutant tumor suppressors (e.g., TP53) (DeBerardinis et al., 2008; Jones and Thompson, 2009). Altered energy metabolism is appropriately termed as emerging hallmarks of cancer due to its evident importance as well as unresolved functional independence from the core hallmarks (Figure 1.1).

1.1.2.10 Evading Immune Destruction
The long-standing theory of immune surveillance proposes that cells and tissues are constantly monitored by an ever-alert immune system. The immune surveillance is responsible for recognizing and eliminating the vast majority of incipient cancer cells and thus nascent tumors. According to this logic, solid tumors that do appear have somehow managed to avoid detection by the various arms of the immune system or have been able to limit the extent of immunological killing, thereby evading eradication. Cancer cells may paralyze infiltrating CTLs and NK cells, by secreting TGF-β or other immunosuppressive factors (Yang et al., 2010; Shields et al., 2010). Immunoevasion of
cancer cells is another emerging hallmark whose generality as core hallmarks remains to be firmly established (Figure 1.1).

1.1.3 Genes associated with cancer

Cancer is a genetic disease as discussed above. Different genes interplay in a complicated circuit to acquire the hallmark capabilities in cancer cells. These genes can be broadly classified as:

1.1.3.1 Oncogenes

Oncogenes play an important role in sustaining proliferation signaling in cancer cells and activate anti apoptotic pathways hereby promoting tumor growth. An oncogene is a viral or cellular gene that can induce one or more characteristics of the neoplastic transformation when introduced alone or in combination with another gene, into an appropriate cell type (Varmus, 1989). Oncogenes are derived from proto-oncogenes, which are normal cellular counterparts and most of these proto-oncogenes are critical regulators of cellular proliferation, differentiation, apoptosis etc. It is only when the proto-oncogenes are abnormally activated through mutations, amplifications, translocations, viral integration etc. do they become “oncogenes” (Vogelstein and Kinzler, 2004). The oncogenes have a dominant or “gain-of-function” phenotype i.e. activation of only one allele is sufficient to trigger oncogenic function (Bishop, 1989; Vogelstein and Kinzler, 2004; Tannock, 2005; Tili et al., 2005; Croce, 2008). The well characterized oncoproteins have diverse functions ranging from signal transduction to cell cycle regulation and transcriptional control. However according to their specific functions they have been classified as: growth factors (sis, int2, FGF-5 etc.), growth factor receptors (ErbB2 family, fms etc.), signal transducers (abl, src, raf, mos, ras gene family etc.), transcription factors (myc family, fos, jun etc.), regulators of cell cycle and cell death (cyclin family, mdm2, HPV E6/E7, Bcl-2, Bcl-xL etc.) (Schwab, 1998; Redon et al., 2002; Croce, 2008).
1.1.3.2 Tumor Suppressor Gene

Tumor suppressor genes (TSGs) are a group of regulators capable of suppressing tumor formation, angiogenesis and activate apoptotic pathways. TSGs are also called anti-oncogenes. In general the TSG mutations are recessive at the cellular level, which means that both alleles have to be inactivated in order to observe “loss-of-function” of the TSG protein (Vogelstein and Kinzler, 2004). The inactivation events of TSG are said to act recessively because both copies of the gene must be inactivated before their protective function is lost; this is explained in detail and elaborately in Knudson’s two hit hypothesis. The Knudson’s two-hit hypothesis is based on incidence curves for familial and sporadic retinoblastoma, a childhood tumor of the eye. Knudson postulated that in familial retinoblastoma, one mutated TSG is inherited via the germline cells, the first hit, followed by a second hit in the somatic retinal cells. In sporadic retinoblastoma, both hits are thought to occur in the same somatic cell, which statistically is less likely to occur compared to the familial form of retinoblastoma (Knudson, 1971). This second allelic inactivation in tumors often occurs through deletion or chromosomal rearrangements (Vogelstein and Kinzler, 2004).

Though Knudson’s “two-hit” hypothesis for TSG inactivation has been applicable to all human cancer types, evidences suggesting transcriptional silencing of gene promoters as a result CpG island methylation led to the suggestion that abnormal methylation of TSG promoters might be implicated in carcinogenesis as revised Knudson’s two hit hypothesis (Jones and Laird, 1999). According to the modified Knudson’s two-hit dogma, both alleles inactivation in the TSG may be achieved by I) genetically through (a) mutations in the regulatory regions that cause elimination or suppression of mRNA expression; (b) deletion in part of or entire gene; (c) missense, nonsense, frameshift or splice site mutation resulting in absence of a functional protein. II) Epigenetically by (a) abnormal methylation; (b) deregulation of imprinting; (c) aberrant splicing or epigenetically (methylation/imprinting/splicing) (Figure 1.2).
Figure 1.2: Knudson’s two-hit hypothesis revised. In somatic mutation, among the two active alleles of tumor-suppressor gene, one allele is inactivated by genetically deletion and/or mutation and other by epigenetically like methylation, imprinting or aberrant splicing in the first hit followed by second hit by either genetically (deletion or mutation) or epigenetically (methylation/imprinting/splicing). In Inherited mutation other active allele of TSGs are inactivated by either genetically (deletion or mutation) or epigenetically (methylation/imprinting/splicing). Adapted and expanded from Jones et al. 1999.

Classes of TSGs

TSGs can be classified on the basis of following aspects:

A. On the basis of mechanism of inactivation:

Class I TSGs

The TSGs that are commonly inactivated by ‘two hit’ events i.e. loss of function is from mutation or deletion or methylation of DNA are termed as class-I TSGs e.g. RB, WT1, NF1, APC, VHL, TP53 etc. (Lee et al., 1991; Zou et al., 1994).
Class II TSGs
The TSGs in which loss of function is from regulatory block in expression. Some recently described tumor-suppressor genes have been hypothesized to exert a selective advantage on a cell when only one allele is inactivated and the other remains functional; this phenomenon is called haploinsufficiency. **Haploinsufficiency** is thought to occur because of dosage effect: the heterozygote produce only half as much of the product encoded by the TSG. This type of gene is regulated by a different suppressor gene that lost its function by mutation or deletion (Lee et al., 1991). The expression levels of these genes are usually low in tumors but they decrease tumorigenecity when its normal cDNA is transfected into cancer cell lines (Daigo et al., 1999). Examples are DLC-1, IGSF4, TIG3 (Daigo et al., 1999; Raval et al., 2003).

**B. On the basis of functional importance:**

**Gatekeeper genes**
The ‘gatekeeper’ are the TSGs that directly regulate growth of tumors by suppressing growth or accelerating death and their loss of functions directly initiate tumor initiation. The examples of gatekeeper TSGs are TP53, RB1, APC, VHL etc. (Kinzler and Vogelstein, 1997).

**Caretaker genes**
These TSGs controls the fidelity of the genome through effective repair of DNA damage or prevention of genomic instability (such as microsatellite or chromosomal) (Lengauer et al., 1998). They are also known as stability genes (Friedberg, 2003). Inactivation of the both alleles of this gene is suggested to lead the instability of the genome but not directly initiation of neoplasia. This class of genes include the Double strand break repair (DSBR), mismatch repair (MMR), nucleotide excision repair (NER) and base excision repair (BER) genes responsible for repairing subtle mistakes made during normal DNA replication or induced by exposure to mutagen. ATM, MSH2, MLH1, BUB1 and BUBR1, etc. are examples of this class of genes. Some TSGs (BRCA1, BRCA2) may possess overlapping caretaker and gatekeeper properties (Kinzler and Vogelstein, 1997). Mutations in both alleles of a caretaker gene must precede mutation in both copies of a gatekeeper gene to have any effect on tumor growth. Restoration of caretaker gene
function will not stop tumor growth if mutations in a gatekeeper gene have already occurred.

**Landscaper genes**

Landscaper genes are the TSGs that act by modulating the microenvironment in which the tumor cells grow, perhaps by direct/indirect regulation of extracellular matrix proteins, or secreted growth/survival factors. The examples of landascaper genes are SMAD4, PTEN and LKB1/STKII (Kinzler and Vogelstein 1998). Recently new class of TSGs has emerged that antagonizes tumor growth in vivo but not in vitro (Imreh et al., 2003). HYAL1 is one of the examples of this class of TSGs.

**1.1.3.3 Cancer predisposing genes**

The genes where the primary mutation is present in the germline and suffer another somatic mutation in somatically arising tumors (Li, 1995). A small proportion of cancers are attributable to the effects of mutations in known cancer susceptibility genes. Families with multiple cases of cancer, frequently arising at an early age have been useful for identification of many cancer predisposing genes. The best studied adult cancer syndromes are hereditary breast and ovarian cancer. These genes are classified according to their level of penetrance. BRCA1 and BRCA2 are two of the major highly penetrant genes associated with breast and ovarian cancer syndrome.

**1.1.3.5 Replication fidelity genes**

The shortening of telomere resulting from absence of enzyme telomerase (TERT protein catalytic subunit), enables cells to count the number of replicative generations that they and their ancestors have passed through since early development (Wright and Shay, 1992). Activation of telomerase and the resulting acquisition of replicative immortality represent an essential step in the development of the great majority of the human tumors (Weinberg RA, 1995). Telomerase is an example of this class of gene.
1.1.4 Multi-step progression of cancer

The development from a normal cell to a cancerous cell is usually a multistep process of clonal evolution driven by accumulation of a series of genetic and epigenetic alterations that progressively convert the cell from normal growth to a premalignant state and finally a cancerous state (Yokota et al., 1993).

Figure 1.3: Development of different predicted models explaining stage wise progression of cancer.

Fould L (1957) proposed that cancer represents the phenotypic manifestation of accumulation of a critical load of genetic damage (Figure 1.3). Nowell PC (1976) formulated the clonal evolution concept which proposes that most neoplasms arise from a single cell of origin, and tumor progression results from acquired genetic variability within the original clone allowing sequential selection of more aggressive sublines (Figure 1.3). Then, it was proposed (Bodmer, 1997; Farber and Cameron, 1980; Scrable et al., 1990) that normal cell progressing to fully malignant phenotype might be due to a nested set of aberrations (Figure 1.3).
Scrable et al (1990) furnished two types of predictions (Figure 1.4) depending on the entry point into which the pathway is viewed. If this point is at the beginning, the circuitry can be viewed as an initiating genotoxic event with clonal outgrowth, perhaps due to incurred proliferation advantages (N→I). Large body of evidence indicates that the initial damage must be fixed and then compounded (Scorable et al 1990).

Figure 1.4: Pathway for Clonal Evolution Model. N, Normal cell; I, Inherited cell; T1, T2, T3, T4, refer to progressively more damaged cells in the promotion and progression stages of malignancy. Filled symbol indicate those cells that have undergone lethal mutations. This pathway can begin in single somatic cell in sporadic cases but can initiate in any cell in inherited cases (Adapted from Scorable et al. 1990).

Based on various experimental models, the process of carcinogenesis can be divided into the three stages of initiation, promotion, and progression based on evidence from experimental models (Harris et al., 1992). Clinically, human tumors can be divided into three groups: premalignant lesions (dysplasia, hyperplasia, leukoplakia, and adenoma), primary tumors, and metastases (Yokota et al 1993). Cells in the premalignant lesions (initiated cells) are clonally expanded because of the acquisition of selective growth advantage by genetic events that occurred in the cells (Figure 1.4). A cellular mechanism that causes biallelic fixation of the first event could cause the irreversible conversion to
an expanded initiating clone (I → T1) in sporadic cases (Yokota et al., 1993; Feinberg et al., 2006) (Figure 1.4). In inherited cases, this pathway begins with one allele mutated in the germline and other allele is somatically altered to generate the initiating predisposed clone. Further damage to this clone would then permit (or even force) it to a stage of higher malignancy (T1 → T2). Iteration of this process would then culminate in a fully malignant clone (T2 → T3 → T4) (Figure 1.4). If the entry point is at the end of the pathway, then one is considering the cellular mass (T4) that is greatly removed from normality and carries the genotoxic damage suffered in each of the previous stage transition. Thus, the dissection of the pathway by which a normal cell becomes fully malignant may be viewed as unraveling of a nested set of aberrations (Farber and Cameron, 1980; Scarble HJ et al., 1990; Bodmer, 1996).

Ilyas et al. (1999) proposed an inverted pyramid model, (Figure 1.3) where first mutation predicts the selection of the second mutation and they interact to ensure optimal activity of both. These together form the basis for selection of the third mutation and so on until a number of mutations have been generated necessary for the metastasis of tumor. Thus, the tumorigenic process is like an inverted pyramid.

He also explained tumor progression through nexus model (Figure 1.3) in the nexus model, there may be no strict interdependence of the interaction of mutations. Tumor development could be viewed as a nexus of interconnecting mutations in which selection pressures apply at each point, but previous mutations may no longer be essential for cell survival. The truth probably lies somewhere in between (Ilyas et al., 1999).

Further studies lead to the development of clonal genetic model and epigenetic progenitor model of cancer progression (Baylin and Ohm, 2006; Feinberg et al., 2006) (Figure 1.3). The discovery of dominantly acting oncogenes and recessively acting tumor suppressor genes have provided support for the clonal genetic model of cancer. Clonal genetic model of cancer has been mainly successful in predicting the gate-keeper mutations necessary for earliest stage of tumor development, such as those in APC in colorectal cancer and VHL in renal-cell carcinoma. Epigenetic progenitor model proposes that cancer arises in three steps: an epigenetic disruption of progenitor cells, an initiating mutation, and genetic and epigenetic plasticity.
Finally the advanced **somatic evolution model** of cancer progression was proposed by Vogelstein et al., 2013 in colorectal carcinoma (Figure 1.3). Genome wide association analysis and whole exome sequencing analysis revealed that average 33 to 66 genes display subtle somatic mutations that would be expected to alter their protein products and hence driver mutation in most solid tumors (Vogelstein et al., 2013).

### 1.1.5 Molecular pathways associated with development of cancer

In contrast to normal cells, cancer cells do not obey the signals of growth, differentiation, and homeostasis and become masters of their own destinies. They have evolved to destroy the normal growth control. Multiple signaling pathways are deregulated to get the growth advantage. Due to the recent advancement in molecular biological techniques, whole cancer genome sequencing analysis, whole genome transcriptome analysis and whole genome epigenome analysis are possible. Genome wide analysis categorizes numerous driver and passenger genes which are associated with the development of several human cancers (Vogelstein et al., 2013). These pathways have been further organized into three core cellular processes: a) Cell fate, b) Cell survival and c) Genome maintenance (Figure 1.5).

![Different molecular pathways and their cellular effects in cancer development. (Adapted from Vogelstein et.al 2013)](image_url)
Chapter 1: General Introduction

a) Cell fate-The precise balance between the cell division and differentiation decides the fate of a cell. Many genetic alterations in cancer cells promote division over differentiation offering a selective growth potential through modulation of self-renewal pathways (Perrimon et al., 2012; Hoffman, 2012), chromatin modifications and transcriptional regulations (Rosetto et al., 2012; Vogelstein et al., 2013). The self-renewal pathways include WNT-Adenomous polyposis coli (APC), Hedgehog (HH) and NOTCH signaling pathways, which are well known to control cell fate in organisms ranging from worms to mammals (Perrimon et al., 2012) (Figure 1.5).

b) Cell survival-Cellular survival is controlled by key regulatory genes like EGFR, HER2, FGFR2, PDGFR, TGFbR2, MET, KIT, RAS, RAF, PIK3CA, and PTEN genes which leads to abrogation of TGF-β, MAPK, PI3K, RAS, STAT pathways (Vogelstein et al., 2013) (Figure 1.5). Progression through the cell cycle can be directly controlled by driver genes that directly regulate the cell cycle or apoptosis, such as P16, MYC, and BCL2 are often mutated in cancers. Another gene whose mutations enhance cell survival is VHL, the product of which stimulates angiogenesis through the secretion of vascular endothelial growth factor (VEGF).

c) Genome maintenance-About $10^5$ DNA lesions are produced in a mammalian genome each day as a result of spontaneous decay, replication errors and cellular metabolism in normal cell (Iyama and Wilson, 2013). Among the range of lesions formed, modified bases, various strand breaks, intra and inter-strand cross-links and protein-DNA adducts are common. Persistent DNA damage can induce mutagenesis, such as base substitutions and small insertions/deletions, as well as gross chromosomal rearrangements. Cells respond promptly to the generation of DNA damage by launching a coordinated set of actions, collectively known as the DNA damage response (DDR) (Shiloh, 2001) (Figure 1.5). The extent of damage incurred by the genome is detected by a set of sensors, which initiates the DDR.
Figure 1.6: Representation of timings of different repair process throughout the cell cycle.

Figure 1.7: Activation of different DDR pathways upon different genotoxic stresses.
1.1.6 DNA Damage Response Pathway

Depending on the nature of DNA damage and stage of cell cycle, different repair mechanisms could be activated like, **Double strand break repair (DSB)**, **Mismatch repair (MMR)**, **Nucleotide excision repair (NER)**, **Base excision repair (BER)**, **Direct reversal and Translesion Synthesis (TLS)** (Figure 1.6, Figure 1.7).

### 1.1.6.1 Double Strand Break (DSB)

**DSBs** are one of the most deleterious forms of DNA damage, activating cell death responses if not repaired and promoting genome instability like translocations in case of improper repairation (Bohgaki et al., 2010; Jackson and Bartek, 2009). DSBs can arise endogenously through the action of ROS that are produced by normal cellular metabolism, or during certain scenarios of failed DNA replication. DSBR is divided into two major pathways, **Homologous recombination Repair (HRR)** and **Non-homologous end joining (NHEJ)**. HR operates in dividing cells and in S phase, as it requires a homologous sister chromatid for execution, whereas NHEJ operates particularly during phases of the cell cycle when a homologous sister chromatid is absent, i.e., mainly at G1 (Figure 1.6, Figure 1.7). The MRN complex which is comprised of the MRE11, RAD50, and Nijmegen breakage syndrome 1 (NBS1) proteins, acts as a break sensor in HR and facilitates the subsequent steps of the recombination process. During initiation of NHEJ, the protein Ku70/Ku80 (Ku) heterodimer binds directly to the two DSB ends and recruits the other machineries for the repair process. However although **HRR** is error proof, **NHEJ** process is more error prone and dangerous to cell. The genes associated with these repair pathways are known as “guardian of the genome” and their defects in cancer can indirectly confer a selective growth advantages to the cells having gross chromosomal changes (Vogelstein et al., 2013).

In **HRR** pathway, in response to the DSB by radiation or different chemical reagents, a cascade of molecular events has been taken place to repair the DNA. Upon DNA damage, BRIT1 forms a nuclear foci and recruits different DDR proteins like ATM, MDC1, NBS1, RAD51 and BRCA2 to the damaged site (Lin et al., 2010; Rai et al., 2006) (Figure 1.8). The nuclear foci of BRIT1 colocalize with the foci of γ-H2AX (Jeffer et al., 2008), upon phosphorylation at S139 by ATM (Xie et al., 2010). NBS1 binds to γ-H2AX
through FHA/BRCT domain to activate DSB repair (Kobayashi et al., 2004). The activation of ATM/ATR is followed by the activation of checkpoint kinases CHK1/CHK2 and the other downstream targets to halt the cell cycle either in S phase or G2/M phase for the repairing process of DNA (Rai et al., 2006). The stimulated ATM recruits a multiprotein complex of MRE11-RAD50-NBS1 (MRN) in the vicinity of the DNA double strand breaks (DSB) (Lee et al., 2005). ATM phosphorylates NBS1 at Ser-343 residue (Wu et al., 2000). In this pathway, formation of FA core complex is an important phenomenon. One FA gene subset encodes nine proteins of the FA core complex (FANCA/B/C/E/F/G/L/M/T), which activates the FANCI–FANCD2 heterodimer through monoubiquitination. The remaining eight FA proteins (FANCD1/J/N/O/P/Q/R/S) mediate recombinational and nucleolytic reactions to complete repair (Zhang and Walter, 2014). FANCA and FANCG form a complex in the cytoplasm. FANCC join the complex. This FA complex translocate into the nucleus, where FANCE and FANCF are present. FANCE and FANCF join the complex. The FA complex subsequently activates FANCI–FANCD2 by monoubiquitination during DNA damage (Nakanishi et al., 2005; Kitao et al., 2006) (Figure 1.8). FANCD2 is also activated by ATM by phosphorylation at S222 for S-Phase arrest (Taniguchi et al., 2002; Pichierri and Rosselli, 2004).
Figure 1.8. Homologous Recombination Repair (HHR). DNA double strand break induces nuclear foci formation of BRIT1 which activates ATM. ATM phosphorylates $\gamma$-H2AX to form nuclear foci and NBS1 to arrest cell cycle at S-Phase. Fanconi Anemia genes including FANCC form a complex which interacts with DNA and leads to the mono-ubiquitination of the FANCD2 protein at K561. FANCD2 is also activated by ATM by phosphorylation at S222 which leads to S-Phase arrest. The ATM kinase activated by DNA damaging agents directly activates BRCA1 by phosphorylation at Ser-1387 residue or indirectly by CHEK2. Through an association with FANCD2, BRCA1, BRCA2, RAD51 and DSS1 in damaged nuclear loci leads to activation of the processes that lead to DNA repair. Immediately after DNA repairment the monoubiquitinated FANCD2 gets deubiquitinated by USP1, thereby inactivating the pathway (Kennedy and D'Andrea, 2005).

The ATM kinase activated by DNA damaging agents directly activates BRCA1 by phosphorylation at Ser-1387 residue or indirectly by CHEK2 (Cortez et al., 1999; Zhang et al., 2004). Activated FANCD2 interacts with BRCA1, BRCA2, Rad51 and DSS1 to the DNA damage site for error free repair by homologous recombination (Figure 1.8) (Garcia-Higuera, 2001; Marston et al., 1999; Hussain et al., 2004; Cousineau et al., 2007). The interaction between BRCA2 and RAD51 the two proteins is mediated by eight BRC repeats, which are characterized by a conserved ‘FxxA’ motif (Scott et al., 2016).

### 1.1.6.2 Mismatch Repair (MMR)

MMR system recognizes and repairs base-base mismatches and insertion-deletion loops (IDLs) arise as a result from DNA polymerase mis-incorporation of nucleotides and template slippage respectively (Figure 1.7, Figure 1.9). Mispairs generated by the spontaneous deamination of 5-methylcytosine and heteroduplexes formed following genetic recombination are also corrected via MMR. The protein MSH2 and MSH6 are responsible recognition, the MLH1 and PMS1 are recruited to organize other proteins, such as PCNA, at the damage site. It is well recognized that a major responsibility of MMR in dividing cells (Figure 1.6) is to suppress genetic instability arising from replication errors and the consequent carcinogenesis mainly at S-phase (Iyama and
A defect in this pathway imparts “mutator phenotype” characterized by elevated frequencies of spontaneous mutations and increased microsatellite instability (MSI), which are the hallmarks of cancer (Loeb et al., 2008a).

Several human MMR proteins have been identified based on their homology to E.coli MMR proteins. These include homolog of MutS, MutL, EXO1 single stranded DNA binding protein RPA, proliferating cellular nuclear antigen (PCNA), DNA polymerase δ and DNA ligase I (Figure 1.9). In human cells, the MMR pathway is initiated by recognition of the mismatch or IDL by Mut Sα (MSH2 & MSH6) or Mut Sβ (MSH2 & MSH3) (Figure 1.9). The former predominantly recognizes base-base mismatch and single-base IDL, whereas the later detects larger IDLs (Guo-Min Li, 2008). After identification of the lesion, Mut Sα or Mut Sβ binds ATP, undergoes a shift in conformation and translocates along the DNA away from the site of the lesion until it encounters additional MMR proteins (Gradia et al., 1999, Blackwell et al., 1998). Higher order protein complexes are formed with Mut Lα (MLH1 & PMS2) and replication factors viz; PCNA. The interaction of Mut Lα with PCNA probably helps the MMR-complex to identify the newly synthesized strand at the replication fork (Umar et al., 1996; Guo-Min Li., 2008). Excision and resynthesis of the nascent strand (containing the mismatch or IDL) is performed by a number of factors including PCNA, RPA, RFC, exonuclease I (Exo1), DNA polymerases delta and epsilon, endonuclease FEN1 and additional factors (Buermeyer et al., 1999) (Figure 1.9).

1.1.6.3 Nucleotide excision repair (NER)

The NER pathway resolves numerous DNA lesions, particularly base modifications that distort the normal helical structure of duplex DNA in any stage of cell cycle (Figure 1.6, Figure 1.7). NER substrates include cyclobutane pyrimidine dimers (CPDs) and pyrimidine-(6,4)-pyrimidone photoproducts (6-4PPs) generated by UV radiation, base adducts created by exogenous chemical agents such as cisplatin and benzopyrene, base lesions produced by reactions with endogenous lipid peroxidation products and reactive oxygen species (ROS)-induced base modifications such as the cyclopurines (Iyama and Wilson, 2013).
Figure 1.9. **Mismatch repair (MMR)**. This is initiated by detection of mismatches and strand discrimination, followed by excision of the newly synthesized stand containing the mismatch. Finally, DNA is resynthesized to ensure DNA fidelity (diagram adapted from https://www.rndsystems.com/resources/articles/dna-damage-response).

---

**Figure 1.9. Mismatch repair (MMR).** This is initiated by detection of mismatches and strand discrimination, followed by excision of the newly synthesized stand containing the mismatch. Finally, DNA is resynthesized to ensure DNA fidelity (diagram adapted from https://www.rndsystems.com/resources/articles/dna-damage-response).
1.1.6.4 Base excision repair (BER)
DNA base modifications are common damages caused by oxidation, deamination or alkylation. In fact, there are >100 types of oxidative base modifications that can potentially arise in DNA as a result of attack of ROS, which are mainly generated by normal mitochondrial respiration. It has been estimated that about 180 guanines are oxidized to 8-oxo-dG per mammalian genome per day. Deamination is another potentially harmful spontaneous reaction, producing uracil, inosine and xanthosine from cytidine, adenine and guanine, respectively. Given the frequent nature of oxidation, deamination and spontaneous hydrolysis, BER expectedly operates during all stages of the cell cycle (Figure 1.6, Figure 1.7), and thus, serves a critical function in both dividing and non-dividing cells. Conventional BER is initiated by a lesion-specific DNA glycosylase (mono- or bi-functional), which recognizes and hydrolyzes the N-glycosidic bond of a substrate base, creating an AP site intermediate (Iyama and Wilson, 2013).

1.1.6.5 Direct reversal
Cells are known to eliminate three types of damage to their DNA by chemically reversing it. These mechanisms do not require a template, since the types of damage they counteract can occur in only one of the four bases. Such direct reversal mechanisms are specific to the type of damage incurred and do not involve breakage of the phosphodiester backbone. The formation of pyrimidine dimers upon irradiation with UV light results in an abnormal covalent bond between adjacent pyrimidine bases. The photoreactivation process directly reverses this damage by the action of the enzyme photolyase (Sancar, 2003). Another type of damage is by Methylguanine (O6-meG) bases which is directly reversed by the protein O6-methylguanine DNA methyltransferase (MGMT). There is a significant reduction in MGMT prior to or early in S-phase; followed by a recovery during the G2/S phase (Iyama and Wilson, 2013) (Figure 1.6). The third type of DNA damage reversed by cells is certain methylation of the bases cytosine and adenine.

1.1.6.6 Translesion Synthesis (TLS)
TLS is a DNA damage tolerance process that allows the DNA replication machinery to replicate past DNA lesions such as thymine dimers or AP sites (Waters et al., 2009)
(Figure 1.6). It involves switching out regular DNA polymerases for specialized translesion polymerases (i.e. DNA polymerase IV or V, from the Y Polymerase family), often with larger active sites that can facilitate the insertion of bases opposite damaged nucleotides. The polymerase switching is thought to be mediated by the post-translational modification of the replication processivity factor PCNA. Translesion synthesis polymerases often have low fidelity (high propensity to insert wrong bases) on undamaged templates relative to regular polymerases. However, many are extremely efficient at inserting correct bases opposite specific types of damage.

1.1.7 Role of DDR pathways in development of cancer

The genome is subject to regular and frequent stressors, from both endogenous and environmental agents (Houtgraaf, 2006) (Figure 1.10). Eukaryotic cells rely on a strictly coordinated series of events, termed the DNA damage response (DDR), to cope with genotoxic insults to maintain homeostasis. The DDR includes cell-cycle checkpoint activation, regulation of DNA replication, and DNA damage repair (Polo et al., 2015, Liu et al., 2016).
Chapter 1: General Introduction

Figure 1.10: How the DNA damage response pathways influence steps leading to cancer. The DNA damage response (DDR) pathways promote critical steps in the aetiology of carcinogenesis. A healthy cell has a plethora of DDR processes to protect its DNA from exogenous and endogenously arising DNA-damaging agents, and respond to viral infections. Nonetheless, the processes are not perfect, and an early step in the aetiology of cancer is the generation of one or more mutational changes. This may directly or indirectly result in oncogene activation, which leads to replicative and/or oxidative stress. Genetic predisposition to cancer can arise when one of these DNA repair processes is compromised. However, although enhanced replication stress increases the level of DNA breakage, mutation or rearrangement, a range of responses — for example, the ability to accurately recover replication, the activation of checkpoint arrest or other p53-dependent responses — can prevent the proliferation of damaged cells. Progression from this precancerous state to ongoing proliferation requires the downregulation of these DDR processes, thereby facilitating persistent genomic instability (Adapted from Jeggo et al., 2016).

Among the DSBR, the HRR pathway is the most frequently affected DDR pathway in different cancers. The BRCA genes are considered to be ‘caretakers’, thereby protecting our genome from carcinogenic alterations. The hereditary forms of breast and ovarian cancer (HBOC) caused by monoallelic, that is dominant mutations in either the BRCA1 or BRCA2 genes (Nielsen and van Overeem Hansen, 2016). The lifetime risk of developing breast or ovarian cancer can be up to ~80% in these affected individuals. Different studies showed variable frequencies of deletion and promoter methylation in BRIT1, ATM and BRCA1/BRCA2 genes in pretherapeutic breast cancer (Chunder et al., 2004, Esteller et al., 2002; Bhattacharya et al., 2011). In addition, high deletion and promoter methylation in FANCC gene were seen in this tumor (Sinha et al., 2008a). High frequencies of deletion of BRCA1 and BRCA2 were reported in ovarian cancer (Gunn et al., 2013). Comparable frequencies of deletion and methylation were reported in FANCC in head and neck cancer samples (Ghosh et al., 2012). In cervical cancer, alteration (deletion/methylation) of ATM/CHK1 locus were reported (Majumdar Indra et al., 2011a, 2011b). Mutations in BRCA1/2 also predispose individuals to cancer in other
Chapter 1: General Introduction

organs such as prostate or pancreas. Deleterious germ line and somatic aberrations of DDR genes (BRCA2, BRCA1, ATM, FANCD2 and RAD51C), heterozygous deletion of FANCD2 and RAD51 associated with castration-resistant prostate cancer (Mateo et al., 2016). The pancreatic cancer genome is highly unstable. That is why pancreatic intraepithelial neoplasia lesions and invasive pancreatic cancer lesions have increased expressions of pATM, pCHK2, pCHK1, γH2AX and RAD51 than the normal pancreatic tissue (Osterman et al., 2014). Some familial pancreatic cancers are associated with genetic defects in DNA damage responses and repair machinery, such as TP53, BRCA2, ATM, PALB2 (Rustgi, 2014). D’Andrea lab at Harvard University discovered biallelic mutations in BRCA2 in a subset of patients with the rare childhood hematological disorder, Fanconi anemia (FA) (FA-D1 subgroup) in 2002 (Howlett et al. 2002). The combined effects of these biallelic BRCA2/FANCD1 mutations appear to be hypomorphic. FA is the most prevalent among the inherited forms of bone marrow failure syndrome (Auerbach, 2009) that progressively develop hypoplastic anemia and cancer predisposition that often results in hematological malignancies such as acute myelogenous leukemia (AML) or myelodysplasia (MDS) as well as various solid tumors, especially head and neck squamous carcinoma (Alter, 2014). In pancreatic lesions increased expression of DNA PKs and Ku70 of NHEJ pathway was reported (Osterman et al., 2014).

The MMR pathway is another important pathway associated with development of cancer. The majority of patients develop colorectal carcinoma (CRC) as a result of chromosomal instability but approximately 15% patients develop CRC due to abnormalities in DNA MMR. Deficient MMR can be secondary to sporadic hypermethylation, which silences DNA MMR genes (Gatalica et al., 2016). The four main MMR genes are MLH1, MSH2, MSH6, and PMS2, and germline defect in these genes are characteristic of Lynch syndrome. These genes can form heterodimers in various combinations; the most common are MSH2/MSH6 and MLH1/PMS2 (Peltomaki, 2005). High microsatellite instability in tumors due to sporadic causes, are a result of epigenetic silencing of the MLH1 promoter (Funkhouser et al., 2012, Veigl et al., 1998). Both genetic and epigenetic alterations of MLH1 and MSH2 contribute to genomic instability and
Chapter 1: General Introduction

tumorigenesis in sporadic breast cancer (Murata H. et al., 2002; Sinha et al., 2008b; Staaf et al., 2008; Alkam et al., 2013). Studies on head and neck and cervical cancer showed deletion and methylation in MLH1 (Ghosh et al., 2010; Sengupta et al., 2007; Mitra et al., 2010). Deletion was also reported in lentigo maligna and malignant melanomas in MLH1 and MSH2 loci (Korabiowska et al., 2001). Some familial pancreatic cancers are associated with MMR-related genes (e.g., hMLH1, hMSH2, and hMSH6) (Rustgi, 2014). MMR is also associated with CRPC (Mateo et al., 2016).

XPA gene is essential for assembly of the pre-incision complex during the processing of DNA damage via the NER pathway (Asahina et al. 1994; Kuraoka et al. 1996; Patrick et al. 2002). Alteration of XPA has been reported in autosomal recessive hereditary disorders like xeroderma pigmentosum. Alterations of NER are associated with CRPC (Mateo et al., 2016). XPA alterations (deletion/methylation) were also reported in breast and head and neck cancer (Sinha et al., 2008; Ghosh et al., 2010).

1.1.8 DDR pathway and treatment of cancer

Treatment of cancer includes chemotherapy, radiotherapy along with surgical removal of the tumors. The therapeutic procedures also incorporate some DNA damages to the genome. NHEJ is thought to be the primary means of repair for therapeutically induced DSBs resulting from irradiation, radiomimetics, topoisomerase poisons and ROS-inducing treatments (Curtin, 2012; Mahaney et al., 2009). So, Inhibition of NHEJ pathway could be a therapeutic target.

Differences in the DNA damage response between normal and cancer cells presumably underlie the ability of the therapeutic agents to preferentially kill cancer cells. Since abnormalities in the DNA damage response of cancer cells are becoming more clearly defined, there is growing interest in the development of small molecules that will selectively target the abnormal DNA repair in cancer cells with the hope that these compounds either alone or in combination with DNA damaging agents will effectively kill cancer cells, while minimizing damage to normal cells (Rassool and Tomkinson, 2010; Huhn et al., 2013) (Table 1.1). High DDR focus formation results in tumor resistance to DNA damage-inducing chemotherapy (Ashakawa et al., 2010).
When there is a loss of or defect in DDR due to oncogenic activation or tumor suppressor inactivation, DNA replication may persist to meet the demands of unrestrained proliferation despite the presence of unrepaired DNA lesions, which then leads to “replication stress”. It is a phenomenon unique to cancer cells. Replication stress induces genomic instability and therefore potentiates oncogenic transformation. However, the novel concept of further enhancing replication stress may provide a plausible alternative to treat cancer due to the induction of “mitotic catastrophe”. Replication stress due to the DNA damage by radiation and chemotherapy is an emerging approach to target DDR. (Zhang et al., 2016).

<table>
<thead>
<tr>
<th>Target</th>
<th>Inhibitor</th>
<th>Mono- or combination therapy / clinical study stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATM</td>
<td>KU-55933</td>
<td>IR, etoposide, doxorubicin, camptothecin, in preclinical testing</td>
</tr>
<tr>
<td></td>
<td>KU-60019</td>
<td>IR in preclinical testing using glioma cells</td>
</tr>
<tr>
<td></td>
<td>NU-6027</td>
<td>Hydroxyurea, cisplatin, temozolomide, rucaparib in preclinical testing</td>
</tr>
<tr>
<td></td>
<td>VE-821</td>
<td>Cisplatin in breast and ovarian cell lines</td>
</tr>
<tr>
<td></td>
<td>ETP-46464</td>
<td>IR, gemcitabine in pancreatic cancer cells in preclinical testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Single agent in p53-deficient cancer cells in preclinical testing</td>
</tr>
<tr>
<td></td>
<td>NU-7441</td>
<td>IR, etoposide in preclinical testing of cancer cell lines and tumour xenografts</td>
</tr>
<tr>
<td>DNA-PKcs</td>
<td>NU-7026</td>
<td>IR and combined with AG14361 (PARPi) in preclinical testing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anthracyclines, mitoxantrone, etoposide in preclinical testing using leukaemia cells</td>
</tr>
<tr>
<td></td>
<td>UCN-01</td>
<td>Single agent in Phase II for relapsed T-cell lymphoma (completed)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Single agent in Phase II for metastatic melanoma (completed)</td>
</tr>
<tr>
<td></td>
<td>GDC-0425</td>
<td>Five-fluorouracil in Phase II for metastatic pancreatic cancer (completed)</td>
</tr>
<tr>
<td></td>
<td>MK-8776</td>
<td>Topotecan in Phase II for various forms of ovarian cancer (completed)</td>
</tr>
<tr>
<td></td>
<td>LY-2606368</td>
<td>Topotecan in Phase II for small cell lung cancer (completed)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Olaparib in pre-clinical testing for multiple mammary tumour types</td>
</tr>
<tr>
<td>CHK1/(CHK2)</td>
<td>UCN-01</td>
<td>Single agent or with gemcitabine in Phase I dose-escalation study (recruiting)</td>
</tr>
<tr>
<td></td>
<td>GDC-0425</td>
<td>Single agent or with gemcitabine in Phase I dose-escalation study (completed)</td>
</tr>
<tr>
<td></td>
<td>MK-8776</td>
<td>Single agent or with gemcitabine in Phase I dose-escalation study (completed)</td>
</tr>
<tr>
<td></td>
<td>LY-2606368</td>
<td>Single agent in Phase I study in patients with advanced cancer (recruiting)</td>
</tr>
<tr>
<td>MRE11</td>
<td>mirin</td>
<td>Single agent or with olaparib (PARPi) in preclinical testing using BRCA2-deficient cells</td>
</tr>
<tr>
<td>RPA</td>
<td>MCI13E</td>
<td>Single agent or with cisplatin in preclinical testing</td>
</tr>
<tr>
<td>RAD51</td>
<td>B02</td>
<td>IR, mitomycin C, cisplatin in preclinical testing</td>
</tr>
<tr>
<td></td>
<td>RI-1</td>
<td>Mitomycin C in preclinical testing</td>
</tr>
</tbody>
</table>

Table 1.1: Small molecule inhibitors of DNA damage response factors in preclinical or clinical development for cancer therapy (Adapted from Huhn et al., 2013).
1.2. Molecular progression of Breast cancer

1.2.1 Epidemiology and incidence
Breast carcinoma (BC) is the most frequent cancer in females worldwide with an estimated 1.67 million new cancer cases diagnosed in 2012 (25% of all cancers). Incidence rates vary across the world regions, with rates ranging from 27 per 100,000 in Middle Africa and Eastern Asia to 92 in Northern America (GLOBOCAN 2012). Whereas, the mortality rates are higher in less developed regions (324,000 deaths, 14.3% of total) than the more developed regions (198,000 deaths, 15.4%) (GLOBOCAN 2012) (Figure 1.11). It is now the most common cause of cancer-related deaths in women worldwide, reflecting a shift in trend from the past decade during which cervical cancer was the dominant cause of cancer-death in women (Jemal et al., 2011; GLOBOCAN 2008). In India, breast cancer is the second most common cancer in women though the incidence rates are rising in the metropolitan cities (Hedau et al., 2011). Currently, India reports roughly 100 000 new cases of BC annually. The overall incidence rate of BC in India is estimated to be 80 new cases per 100 000 population per year (Bagchi, 2008). A 2005 study conducted by the International Association of Cancer Research, based in Lyon, France, projected that there would be 250 000 cases of breast cancer in India by 2015, a 3% increase per year (Bagchi, 2008).

A low frequency of male BC has been documented in India (Shukla et al., 1996). Bilateral BC (BiBC) occur in 1-15% of patients worldwide (Abdalla et al., 2000). BiBC is also reported in India, with incidence of metachronous BC at 2.3 per 1000 women (Gajalakshmi et al., 1998). Synchronous bilateral breast tumors on the other hand are a rare event with an incidence of 0.52% worldwide (Ben Hassouna et al., 2008).
Figure 1.11: Global scenario of breast cancer (Adapted from Globocan 2012, WHO)
1.2.2 Etiological factors

There are several risk factors of breast cancer (*Table 1.2*) (Hankinson et al., 2004). Broadly, risk factors contributing to the disease etiology can be categorized into genetic and non-genetic risk factors. The non-genetic risk factors can be further categorized as hormonal and non-hormonal risk factors.

1.2.2.1 Genetic risk factors

Breast carcinomas can be classified as sporadic and familial/hereditary. Most of the BC are sporadic i.e., without any particular genetic predisposition. About 10%–30% of breast cancer cases are attributed to hereditary factors and only 5%–10% of breast cancer cases are identified with a strong inherited component (Peto et al., 2000; Clause et al., 1994; Hall et al., 2009; Newman et al., 1998) which is a strongest risk factors for the disease (Dumitrescu and Cotarla 2005). Single nucleotide polymorphisms (SNPs) identified through genome-wide association studies (GWAS) declared only a small proportion of genetic risk in BC (Stacey et al., 2007, 2008; Kidd et al., 2008; Krepisch et al., 2012; Kuusisto et al., 2013; Long et al., 2013; Michailidou et al., 2013; Sapkota et al., 2013b, 2014).

Cancer predisposing genes can be categorized according to their relative risk of a particular type of cancer. They are **high-penetrant** genes (cancer relative risk > 5) **intermediate-penetrant** genes (cancer relative risk 1.5-5) and **low-penetrant** genes (cancer relative risk around 1.5) (Apostolou and Fostira, 2013).

**a) High penetrant genes**

In BC, **BRCA1** and **BRCA2** are two major high-penetrance genes. Their mutations account for about 16%–25% of the familial risk in the general population (Collins and Pollitopoulus, 2011; Stratton and Rahman, 2008; Peto et al., 1999) with about 80-85% risk of developing the disease (Apostolou and Fostira, 2013) and 50% chance of developing the disease in next generation (Jatoi et al., 2008). Inherited breast cancer from germline mutations in BRCA1 is associated with a number of distinguishing clinical features, such as an early age of onset, a higher prevalence of bilateral breast cancer, triple negative status i.e., lack of expression of human epidermal growth factor 2 (HER2),
estrogen receptor (ER) and progesterone receptors (PR) and risk of developing ovarian, colon and prostate cancers (Fostira et al., 2012).

In addition to BRCA1 and BRCA2, there exist a number of rare high penetrance genes variants such as TP53, PTEN, STK11 and CDH1 that underlie elevated risk (32-90%) of BC (Nathanson and Weber, 2001; Dumitrescu and Cotarla, 2005; Cowin et al., 2005; Collins and Pollitopoulus, 2011; Apostolou and Fostira, 2013). Collectively, mutations in the high penetrant breast cancer susceptibility genes found to date, account for around 25% of the familial risk of BC (Collins and Pollitopoulus, 2011).

b) Intermediate and low penetrant genes

In addition to high penetrance genes, various moderate penetrance gene variants like CHEK2, PALB2, ATM, BRIP1, NBS1, MRE11, RAD50, RAD51 etc. play an important role in conferring higher risk (15-40%) of BC (Bradbury and Olopade, 2007; Cox et al. 2007; Apostolou and Fostira, 2013).

Most of the low-susceptibility loci have been highlighted through GWAS and initially included a number of loci. In the final BC assessment risk, six SNPs showed statistically significant association with MAP3K1, FGFR2, LSP1, TNRC19, H19 and CASP8 (Easton et al., 2007; Ahmed et al., 2009; Thomas et al., 2009; Zheng et al., 2009; Apostolou and Fostira, 2013).

1.2.2.2 Hormonal risk factors

It has been shown repeatedly that prolonged or elevated exposure to estrogen due to early age of menarche, nulliparity and late age of menopause, prolonged use of combined estrogen and progestin menopausal hormone therapy and factors increasing menstrual cycle number are associated with increased likelihood for developing BC. Nulliparity and late maternal age at first birth both increase the lifetime incidence risk of developing the disease. Finally, there is an association between obesity and an increased breast cancer risk. The major source of estrogen in postmenopausal women is from the conversion of androstenedione to estrone by adipose tissue; thus, obesity is associated with a long-term increase in estrogen exposure (McPherson et al., 2000)
1.2.2.3 Non-hormonal risk factors

A number of non-hormonal risk factors are associated with the development of breast cancer. Exposure to high doses of ionizing radiation especially during adolescence i.e. the active period of development of breast showed increase in cancer risk. Also, alcohol consumption, high fat diet could increase the estradiol levels in serum suggesting their association with BC development by increasing exposure to estrogen (Martin and Weber, 2000).

<table>
<thead>
<tr>
<th>Etiology of breast cancer</th>
<th>Direction of effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family history in first-degree relative or genetic predisposition (e.g. BRCA1)</td>
<td>↑↑</td>
</tr>
<tr>
<td>Height</td>
<td>↑</td>
</tr>
<tr>
<td>Benign breast disease</td>
<td>↑↑</td>
</tr>
<tr>
<td>Mammographically dense breasts</td>
<td>↑↑</td>
</tr>
<tr>
<td>Parity</td>
<td>↓</td>
</tr>
<tr>
<td>Age at first birth &gt; 30 years versus at &lt; 20 years</td>
<td>↑↑</td>
</tr>
<tr>
<td>Lactation (longer durations)</td>
<td>↓</td>
</tr>
<tr>
<td>Menopause at &gt; 54 years versus at &lt; 45 years</td>
<td>↑↑</td>
</tr>
<tr>
<td>High endogenous estrogen levels</td>
<td>↑↑</td>
</tr>
<tr>
<td>Postmenopausal hormone use</td>
<td>↑</td>
</tr>
<tr>
<td>Ionizing radiation exposure in childhood</td>
<td>↑↑</td>
</tr>
<tr>
<td>Menarche at &lt; 12 years versus at &gt; 14 years</td>
<td>↑</td>
</tr>
<tr>
<td>High body mass index (postmenopausal)</td>
<td>↑</td>
</tr>
<tr>
<td>High body mass index (premenopausal)</td>
<td>↓</td>
</tr>
<tr>
<td>Alcohol use (~ 1 or more drinks/day)</td>
<td>↑</td>
</tr>
<tr>
<td>Current oral contraceptive use</td>
<td>↑</td>
</tr>
<tr>
<td>Physical activity</td>
<td>↓</td>
</tr>
<tr>
<td>High prolactin levels</td>
<td>↑↑</td>
</tr>
<tr>
<td>High premenopausal insulin-like growth factor I levels</td>
<td>↑↑</td>
</tr>
<tr>
<td>In utero exposures</td>
<td>↑</td>
</tr>
<tr>
<td>Nonsteroidal anti-inflammatory drug use</td>
<td>↓</td>
</tr>
</tbody>
</table>

Table 1.2: Etiological factors of BC. Arrows indicate approximate magnitude of the relationship: ↑, slight to moderate increase in risk; ↑↑, moderate to large increase in risk; ↓, slight to moderate decrease in risk; ↓↓, moderate to large decrease in risk (Adapted from Hankinson et al., 2004).
1.2.3 Classification of breast tumors

1.2.3.1 Histological

BC is a malignant epithelial tumor originating in the terminal ducts or lobules in the mammary gland and is histologically classified either as non-invasive (carcinoma in situ) or invasive (infiltrating). The ductal and lobular subtypes constitute the majority of all breast cancers worldwide, with the ductal subtype accounting for 40–75% of all diagnosed cases (Bombonati and Sgroi 2011). Apart from cancer of breast, there are other types of breast lesions such as (a) benign fibrocystic disease including fibroadenosis, ductal ectasia, atypical ductal hyperplasia, atypical lobular hyperplasia etc., (b) fibroepithelial diseases including fibroadenoma (benign), (c) intraductal papilloma (benign), and phylloid (benign and malignant) (Lundin et al., 1998).

1.2.3.2 Stage and grade

Breast tumors can also be graded according to their differentiation. Grade 1: Well differentiated; Grade 2: Moderately differentiated; and, Grade 3: Poorly differentiated. Breast tumors can also be classified depending on the stage at the time of diagnosis (Bunting, 1976). The TNM-staging is based on tumor size (T), regional lymph node involvement (N) and distant metastasis (M). T0 means there isn't any evidence of the primary tumor. This means the cancer is "in situ" (the tumor has not started growing into healthy breast tissue). The numbers T1, T2, T3, and T4 are based on the size of the tumor and the extent to which it has grown into neighbouring breast tissue. The higher the T number, the larger the tumor and/or the more it may have grown into the breast tissue. Similarly, N0 means nearby lymph nodes do not contain cancer while N1, N2, N3 denote the number of lymph nodes involved and how much cancer is found in them. M1 means that distant metastasis is present. Once the T, N, and M are determined, they are combined, and an overall "Stage" of I, II, III, IV is assigned (International Union against Cancer, UICC-TNM classification). The most favorable 10 year survival in patients with a Stage 0 (carcinoma in situ) or Stage I is found to be 95% and 88%, respectively, followed by Stage II tumors (66%), and only 36% and 7% with a Stage III and Stage IV tumors, respectively (Bland et al. 1998). In most Asian countries including India, the
majority of breast cancers continue to be diagnosed at a relatively late stage (Agarwal et al., 2007).

1.2.3.3 Laterality
Depending on its appearance in either of the breasts tumors may be unilateral or bilateral (biBC). When two tumors appear simultaneously in both breasts then the tumors are said to be of bilateral type. Bilateral tumors may be synchronous or metachronous. Synchronous tumors are diagnosed simultaneously within one year while, metachronous tumors diagnosed if the time exceeds one year (Heron et al., 2000). The lifelong probability of biBC is greater in younger women with BC due to longer overall life expectancy. Family history of BC is also considered to be an indicator of elevated risk of biBC (Dawson et al., 1998).

1.2.3.4 Age of onset
BC can be further subdivided into 2 groups; early-onset (< 40 years) and late-onset (> 40 years) depending on the age at onset of tumors at the time of diagnosis (Chunder et al., 2004) Higher proportions of breast cancer patients in developing Asian countries like India are younger (age <40) than patients in developed Asian and western countries (Agarwal et al., 2007; Foo et al., 2005). Comparative analysis has shown that younger patients with BC exhibit more aggressive pathological features that include lymphovascular invasion, grade 3 histology, extensive intra-ductal component, necrosis, overexpression of the HER-2 oncogene, absence of the estrogen receptor, higher S-phase fractions, and more abnormal expression of p53 (Zhou and Rhet, 2004; Wenger et al., 1993; Bonnier et al., 1995). Also, patients with invasive breast carcinoma who are aged 35–40 years or younger at the time of diagnosis have been found in several studies to have worse prognosis compared with older patients (Zhou and Rhet, 2004). Furthermore, several studies have indicated that younger patients treated with breast conserving therapy (BCT) have a significantly higher rate of local recurrence than older patients (Zhou and Rhet, 2004). Thus early-onset BC might be considered a separate disease and biologically different from older patients with BC (Chung et al., 1996; Chunder et al., 2004). In general, familial breast cancers are known to be predominantly of early onset type (Hemminki et al., 2001). Recent studies indicate that women with mother or sister
affected by breast cancer are diagnosed and die at earlier ages than do women without family history (Brandt et al., 2010).

1.2.3.5 Molecular subtypes

Molecular subtypes are a key parameter in classifications of BC. It is now well established that determination of ER, PR and HER2 status of invasive breast carcinoma are useful prognostic and predictive markers for the disease (Polyak, 2007). While ER Positivity predicts for response to endocrine therapy such as anti-estrogen (tamoxifen) administration and HER2 positivity is useful for selecting targeted therapy with monoclonal antibody (trastuzumab) against HER2. Comprehensive gene expression profiling of large sets of tumors by multiple independent groups and technologies have revealed five major molecular subtypes of BC: Triple negative (TNBC), luminal A, luminal B, HER2+ enriched, Claudin-low and normal breast–like (Figure 1.12). (Polyak, 2007; Ma et al., 2010). The molecular differences result in distinct clinical outcomes and responses to treatment as in general, the Triple negative (TNBC) subtype has the worst and luminal A-type tumors the best, prognosis (Perou 2010).

Triple-negative BCs (TNBCs) do not respond to chemotherapy and there is a high risk for recurrence and disease progression with these tumors. TNBCs can further be divided into BLBCs and QNBC/5NPs according to the basal-markers (Figure 1.12). TNBC expressing either EGFR or CK5/6 is defined as BLBC (ER−, PR−, HER2−, CK5/6+, and/or EGFR+). Breast tumors which are Triple negative and express neither CK5/6 nor EGFR are defined as ‘quintuple-negative breast cancer’ (QNBC/5NP) (ER−, PR−, HER2−, CK5/6−, and EGFR−) (Choi et al., 2010).

In humans, the most differentiated tumor subtypes of BC are the luminal A and luminal B. Her-2 enriched subtype states are usually linked to late luminal/progenitor cell populations. The Claudin-low subtype and the basal-like cell hormone receptor-negative subtype of BC most resemble the breast stem cell population and luminal progenitor cell populations respectively. A BRCA1 mutation is usually linked to basal-like/luminal progenitor phenotype, and somehow loss of BRCA1 may block further differentiation in a cell and keep the cell in this step of development. (Incassati et al., 2010) (Figure 1.12)
The intrinsic subtypes of breast tumor may be reflective of arrest at different stages of epithelial cell development (Perou, 2010)

**Figure 1.12:** Different molecular subtypes of BC with their differentiation status. The Claudin low is the least differentiated subtype represents the mammary stem cell population and the Liminal A and B represents the differentiated, late luminal/progenitor cells.
Intra-tumor heterogeneity in breast cancer

Breast cancer is a very heterogeneous disease at both the histological and molecular levels. Molecular classification of breast cancer based on complex patterns of gene expression provides a link between the molecular biology of breast cancer and the behavior of cancer cells in the corresponding subtypes (Ellis et al., 2013). In 2012, the Cancer Genome Atlas Network published results related to analyses of gene expression patterns, gene mutations, DNA copy number, DNA methylation, and miRNA expression patterns among a large cohort of approximately 800 breast cancers (Cancer Genome Atlas, 2012). This study demonstrated clearly that breast cancer is a heterogeneous disease with multiple distinct molecular subtypes and that there is great diversity among the recognized major molecular subtypes (Cancer Genome Atlas, 2012). Many breast cancers are not classifiable because of mixed biomarker expression (eg., ER+/PR- or ER-/PR+) or lack of expression of all five markers. The ER+/PR+/HER2- subset contained 100/111 (90%) of the luminal A breast cancers, but this represents only 49% of the cancers that were classified as ER+/PR+/HER2- (Figure 1.13) (Rivenbark et al., 2013). Thus, more than half of patients with ER+/PR+/HER2- breast cancers have disease that will behave clinically in a manner that is more similar to the molecular subtypes with poor prognosis (luminal B, HER2 enriched, basal-like, and/or claudin-low) (Rivenbark et al., 2013). The available genomic, transcriptomic, and proteomic data could be interpreted to suggest that every breast cancer is a distinct entity.

Increasing body of evidence indicate that breast cancer is driven by a component of tumor initiating cells that retain stem cell-like properties (Visvader, 2008). Thus, like in various other cancers, the small population of mammary epithelial stem cells could be the target cells for malignant transformation which drives carcinogenesis, as well as differentiation, that contributes to tumor cellular heterogeneity (Behbod and Rosen, 2005).
1.2.4 Breast cancer progression

Clinical cancer develops over a long period of time, requires multiple molecular alterations, and involves evolution of cellular populations with increasingly aggressive phenotypic characteristics (Kurose et al., 2001; Wu et al., 2012). The initiation of breast cancer is due to transforming (genetic and epigenetic) events in a single cell. Subsequent tumor progression is driven by the accumulation of additional genetic changes combined with clonal expansion and selection (Polyak, 2007).

1.2.4.1 Histological Progression

Normal breast ducts are composed of the basement membrane and a layer of luminal epithelial and myoepithelial cells. The transition from normal breast epithelium to a malignant tumor occurs in several stages. The classical model of breast cancer progression of the ductal type proposes that neoplastic evolution initiates in normal
epithelium (normal), advances to atypical ductal hyperplasia (ADH), evolves to ductal carcinoma in situ (DCIS), culminates as invasive ductal carcinoma (IDC) and finally into metastatic disease (Figure 1.14) (Polyak, 2007; Bombonati and Sgroi, 2011; Karakosta et al., 2005; Rivenbark et al., 2013). ADH is a premalignant lesion characterized by abnormal cell layers within the duct or lobule. ADH is thought to be the precursor of DCIS (Polyak, 2007). DCIS is a commonly diagnosed breast lesion that accounts for 25% of breast neoplasm (Virnig et al., 2010; Jemal et al., 2010, 2011). DCIS is by definition noninvasive, but can vary from low grade to high grade lesions. High grade DCIS, is a risk factor for development of invasive breast cancer (Allegra et al., 2009). Early alterations in the breast epithelium leading to the development of pre-invasive DCIS lesions may determine the severity of the invasive breast cancers that are subsequently develop (Allred et al., 2008). So, DCIS is thought to be a precursor of invasive ductal carcinomas. The lymph nodes are the primary site for breast cancer metastasis.

1.2.4.2 Molecular Progression

Cytogenetic analysis
Molecular studies provide further evidence of recurring chromosomal breakpoints in breast cancer (Hainsworth and Garson, 1990]. Several non-random abnormalities, breakpoints in chromosomal arms 1p/q, 2p/q, 3p/q, 4p/q, 5p/q, 6q, 7p/q, 8p/q, 9p, 11p/q, 12p/q, 13q, 14q, 15q, 16p/q, 17p/q, 18p/q, 19p/q, 20p/q, 21q, 22q and Xp/q (Table 1.3) were observed in BC and the most common numeric changes (trisomy) were observed in chromosomes 1, 3, 7, 8, 11, 12, 18, 19, 20, Y and X (Heim and Mitelman, 1995; Mitelman et al., 1997; Trikkonen et al., 1998; Teixeira et al.. 2002). CGH study in human breast cancer cell lines showed chromosomal gains were at chromosomes 8q, 1q, 20q, 7p, 3q, 5p, 7q, 17q, 1p, and 20p, and the most common losses were: 8p, 18q, 1p, Xp/q, 4p, 11q, 18p, 10q, and 19p (Forozan et al., 1999; 2000).
<table>
<thead>
<tr>
<th>Structural Abnormalities</th>
<th>Chromosomal location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deletion</td>
<td></td>
</tr>
<tr>
<td>1p36, 1p34, 1p33, 1p32, 1p31, 1p22, 1p21, 1p13, 1p12, 1q11, 1q12, 1q21, 1q23, 1q25, 1q32, 1q42, 1q44, 2p21, 2q13, 2q33, 3p25, 3p21, 3p14, 3p13, 3p12, 3p11, 3q11, 3q12, 3q13, 3q21, 4p13, 4q13, 4q22, 5p15, 5p13, 6p21, 6q11, 6q12, 6q13, 6q14, 6q15, 6q16, 6q21, 6q22, 6q23, 6q24, 7p11, 7q21, 7q22, 7q32, 8p22, 8p12, 9p21, 9p13, 9q22, 9q31, 9q32, 10q22, 10q24, 11p13, 11p12, 11q13, 11q14, 11q22, 11q23, 12p12, 12p11, 13q14, 15q24, 16q13, 16q21, 16q22, 16q23, 17p12, 17p11, 18p11, 20p12, 22q12, Xp11, Xq21, Xq24</td>
<td></td>
</tr>
<tr>
<td>Isochromosome</td>
<td>1p10, 1q10, 2p10, 5p10, 6p10, 7q10, 8q10, 10q10, 11q10, 13q10, 14q10, 17q10, 21q10</td>
</tr>
<tr>
<td>Duplication</td>
<td>1q21-1q44, dup(1)(q21; q44), 6p21-6p25, dup(6)(p21; p25), 11q11-11q25, dup(11)(q11; q25)</td>
</tr>
<tr>
<td>Translocation</td>
<td>1p36 der(1)(t;1;1)(p36; q12), 1p13 der(1)(t;1;8)(p13; q11), 1p10 der(1;16)(q10; p10), 1q12 der(1)(t;1;1)(p36; q12), 8q11 der(8)(t;1;8)(p13; q11), 13q10 der(13;13)(q10; q10), 14q10 der(14;15)(q10; q10), 15q10 der(14;15)(q10; q10), 16q10 der(1;16)(q10; p10)</td>
</tr>
<tr>
<td>Inversion</td>
<td>7q22 inv(7)(p22; q11), 7q11 inv(7)(p22; q11), 7q11 inv(7)(q11; q32), 7q32 inv(7)(q11; q32)</td>
</tr>
</tbody>
</table>

**Table 1.3:** Different chromosomal alterations associated with BC (Extracted from Mitelman et al., 1997)
Molecular alteration

Molecular progression of BC involves several chromosomal alterations like deletion, amplification, mutations, translocation, isochromosome formation etc. (Mitelman et al., 1997; Garnis et al., 2004). Among the changes, deletion in 1p, 2p, 3p, 4q, 6q, 8p, 9p/q, 11p/q, 13q, 14q, 16q, 17p/q, 18q etc. are significant (Garnis et al., 2004; Mitelman et al., 1997; Kurose et al., 2001) (Table 1.3).

These molecular alterations lead to upregulation/downregulation of several candidate genes present in those regions and leads to progression of breast carcinoma (Figure 1.14). Though the molecular progression through different histological stages in BC is not well characterized, a number of genetic and epigenetic changes of these candidate genes have been seen for the development of ADH and DCIS indicating the alteration of different cellular pathways in the development of these tumors. In ADH, deletion of BRCA1, APC, CDH1 and SFRP2 genes were observed in a sample (Mukherjee et al., 2015) (Figure 1.14). Common gene amplification of MYC, HER2, FGFR1 and CCND1 was reported in DCIS and synchronous adjacent IDC (Jang et al., 2012; Latta et al., 2002) (Figure 1.14). In invasive hereditary tumors, as discussed earlier (section 1.2.2.1), deletion and mutation in BRCA1 and BRCA2 genes have been well documented (Trikkonen et al., 1997; Eiriksdottir et al., 1998; Ingvarsson et al., 1998; Savelyeva et al., 2001; Welsch and King, 2001). Cells lacking these two genes have been reported to accumulate different chromosomal abnormalities such as breaks, severe aneuploidy and centrosome amplification (Deng and Scott, 2000). Frequent LOH in BRCA1 and BRCA2 as well as in other genes like ATM, RB1, TP53, FANCC, CHEK1, CHEK2, BRCAX, BRIT1, MLH1, MSH2, XPA, APC, MCC, SFRP1, SFRP2, CDH1, LOH11CR2A, PTC1 SH3GL2, SLIT2, ROBO1, ROBO2 etc. were observed in BC (Hamann et al., 1996; Kerangueven et al., 1997; Laake et al., 1997; Phelan et al., 1998; Hanby et al., 2000; Seitz et al., 2001; Chunder et al., 2004; Esteller et al., 2002; Sinha et al., 2008a, 2008b, 2008c, 2011; Staaf et al., 2008; Bhattacharya et al., 2012; Mukherjee et al., 2011, 2016; Bhattacharya et al., 2011) (Figure 1.14). Promoter methylation of these genes causing their reduced mRNA and protein expression were also reported (Thompson et al., 1995, Shirode et al., 2014, Lee et al., 1999, Yoshikawa et al., 1999; Van der Groep
et al., 2008; Roa et al., 2004; Viswanathan et al., 2006; Naqvi et al., 2008; Sinha et al., 2008a,b,c; Alkam et al., 2013; Louise H. et al., 2006). Amplifications of CCND1, MYC and overexpression of β-Catenin were found in BC (Mukherjee et al., 2012; 2016) (Figure 1.14). CD44, ALDH1, TWIST1, SNAI1, CDH1, VIM and cytokines like ILF2, LIF, CXCL1 and CCL2 etc. were down regulated in DCIS (Gudjonsson et al., 2005; Ginestier et al., 2007; Mani et al., 2008; Abba et al., 2004) (Figure 1.14).

Studies based on epidemiological and histological analysis have demonstrated an elevated long term risk of BC in women with benign breast tumors like fibroadenoma and phyllodes tumor (Yang et al., 2014; Dupont et al., 1994; Kasami et al., 1998; Cichon et al., 2010). In this benign breast diseases, alterations (deletion/methylation) were found in stem cell renewal pathway (MCC, APC, CDH1 etc.), HRR and MMR pathway (BRCA1, BRCA2, FANCC, TP53, ATM and MLH1), G1-S checkpoint (RB, CTDSPL, p16INK4a etc.) and cell signaling (SH3GL2) genes (Mukherjee et al., 2015).
Figure 1.14: Histological and molecular progression of BC. Minus sign (-) indicates loss/deletion, arrows; ↑indicate upregulation or amplification, ↓indicate downregulation or inactivation.
Microarray based analysis also showed alterations in different molecular pathways in the development of breast tumors (Hannemann et al., 2006; Tripathi et al., 2008). Thus, cumulative analysis of the above mentioned alterations of the candidate genes along with the microarray analysis lead to the identification of several cellular pathways like DDR (HRR, MMR, NER etc.), self-renewal and EMT pathway (WNT, HH etc.), cell cycle and apoptosis (CCND1, RB1, LOH11CR2A, P53 etc.), cell signaling (G-protein coupled signaling, chemokine receptor activity, MAPK cascade, PI3K etc.) which play a crucial role in breast cancer progression (Figure 1.15).

Figure 1.15: Different pathways involved in the development of BC (Shaded in Blue). The alterations of these pathways are associated with different cellular functions to acquire the hallmark capabilities (Shaded in Green).
1.2.5 Importance of DDR pathways in breast cancer

Among the different molecular pathways, DDR pathways are most important in BC due to the presence of several BC susceptible genes. In section 1.1.7, we have already discussed the roles of different DDR pathways in development of different cancers. Among them the HRR and MMR pathways are most important.

**HRR** pathway is very important in breast tumorigenesis as it includes most of the BC susceptible genes like BRCA1, BRCA2, TP53, ATM etc. Their germline mutation frequencies were already discussed in familial BC. In sporadic BC mutation is a rare incidence. In pretherapeutic sporadic BC, different studies showed variable frequencies of deletion (22-70%), promoter methylation (10-60%) and mutation (4-10%) in BRCA1/BRCA2 genes (Chunder et al., 2004; Esteller et al., 2002; Peto et al., 1999; Malone et al., 2000; Dobrovic et al., 1997). In addition, 40-65% deletion and promoter methylation in FANCC gene were seen in this tumor (Sinha et al., 2008a). Reduced mRNA and protein expression of BRCA1 was reported in BC (Thompson et. al., 1995, Shirode et. al., 2014, Lee et. al., 1999, Yoshikawa et al., 1999). Reduction in mRNA and protein expression of FANCC were also observed in BC (Sinha et al., 2008a). In BC, reduced expression of BRCA2 and FANCD2 proteins were also found (Shirode et. al., 2014, Van der Groep et. al., 2008). About 41-47% alteration (deletion/methylation) and 29-68% reduced expression of BRIT1 were reported in BC samples and cell lines (Bhattacharya et al., 2004, 2012, Rai et al., 2006; Lin et al., 2010; Richardson et al., 2011). Frequent deletion (33-59%), methylation (50-81%) and its reduced mRNA/protein expression (50-88%) of ATM were reported in BC (Sandra Angele et al., 2000; Kerangueven et al., 1997; Laake et al., 1997; Chunder et al., 2004; Ding et al., 2007; Prokopcova et al., 2007; Treilleux et al., 2007; Ye et al., 2007). LOH rate in TP53 was found to be 80% in BC (Rhiem K et al., 2010). TP53 mutation was reported in 19-57% BC samples (de Jong et al., 2002; Pezeshki et al., 2001; Yallowitz et al., 2015; Kim et al., 2016). Deletion and mutation and aberrant expression of CHEK2 are associated with high risk of BC (Borun et al., 2015; Muranen et al., 2011). In CHEK1, 31-45% deletion and methylation were also reported (Sinha et al. 2011).
MLH1 and MSH2 are the most important genes in MMR pathway. Both genetic and epigenetic alterations of hMSH2 and especially of hMLH1 contribute to genomic instability and tumorigenesis in sporadic BC (Murata H. et al., 2002). About 32-38% of the samples showed deletion in MLH1 and MSH2 loci in pretherapeutic BC (Sinha et al., 2008b; Staaf et al., 2008). Variable frequencies (11-45%) of promoter methylation of MLH1 and MSH2 were reported in BC (Roa et al., 2004; Viswanathan et al., 2006; Naqvi et al., 2008; Sinha et al., 2008b; Alkam et al., 2013). Reduced expression of MLH1 and MSH2 was also observed in BC (Murata et al., 2005; Alkam et al., 2013).

Apart from HRR and MMR pathway, other repair genes like XPA was also reported to be altered in BC (Sinha et al., 2008a).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Alteration</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deletion</td>
</tr>
<tr>
<td>BRCA1, BRCA2</td>
<td>22-70%</td>
</tr>
<tr>
<td>FANCC</td>
<td>40-65%</td>
</tr>
<tr>
<td>FANCD2</td>
<td></td>
</tr>
<tr>
<td>ATM</td>
<td>33-59%</td>
</tr>
<tr>
<td>BRIT1</td>
<td>41%</td>
</tr>
<tr>
<td>TP53</td>
<td>24-80%</td>
</tr>
<tr>
<td>CHEK1</td>
<td>31-35%</td>
</tr>
<tr>
<td>MLH1, MSH2</td>
<td>32-38%</td>
</tr>
<tr>
<td>XPA</td>
<td>20-21%</td>
</tr>
</tbody>
</table>

*Table 1.4: The frequencies of alterations of different DDR genes in BC.*
1.2.6 Breast cancer therapy and DDR pathways

The conventional therapy apart from surgery for breast cancer are chemotherapy, radiotherapy, hormonal therapy and immunotherapy. The recent trend in therapy of breast cancer is the administration of neoadjuvant therapies among which neoadjuvant chemotherapy (NACT) is important for locally advanced tumors for better prognosis of the disease (Gampenrieder et al., 2013).

![Therapy diagram]

**Figure 1.16:** Different strategies of breast cancer therapy.

![Mammogram diagram]

**Figure 1.17:** Mammogram of a same patient before and after NACT. NACT shrinks the locally advanced tumor leaving the chemo-tolerant residual tumor.
The major chemotherapeutic drugs used for BC are 5'-fluorouracil, doxorubicin/epirubicin with alkylating agents like cyclophosphamide (FAC/FEC) (Trudeau et al., 2005, Clarke et al., 2005; Tabchy et al., 2010; Mulligan et al., 2014). Of these, anthracycline-based chemotherapy like doxorubicin, potent DNA intercalator, induces DSBs (Siitonen et al., 2016; Capranico et al., 1989, 1990), the most cytotoxic DNA lesion that leads cells into apoptosis especially when relevant repair pathways like HRR are perturbed (Sancar et al., 2004). It is important to note that DNA damage repair competence varies among individual breast tumors and closely correlates with chemosensitivity. For example, secondary mutations of BRCA1 or BRCA2 (essential factors in the HR pathway) caused by chemotherapy using cisplatin or poly (ADP-ribose) polymerase inhibitor in BRCA1/2-mutated cancers restore the wild-type reading frame and, therefore, the tumor acquires resistance to these drugs (Sakai et al., 2008; Edwards et al., 2008; Swisher et al., 2008). Cell deficient in HR are sensitive to platinum based drugs (cisplatin and carboplatin) and mitomycin C (Bhattachayya et al., 2000; Tutt et al., 2005; Liu et al. 1998). BRCA-deficient cells have also shown hypersensitivity to etoposide, a topoisomerase II inhibitor (Rowe and Glazer, 2010). Based on this evidence it has been suggested that HR competence could be a potential biomarker for chemosensitivity (Livingston and Silver, 2008).

Disruption of the FA pathway leads to increased cisplatin sensitivity in tumor cell lines. This has been accomplished by using a gene therapy approach (Ferrer et al., 2004). Different small molecules targeting DSB repair pathway were listed in Table 1.1.

In contrast to other DNA repair systems, a functional MMR pathway actually enhances the cytotoxicity of a variety of chemo therapeutic agents. Following administration of chemo therapeutic agents such as temozolomide (TMZ) or 6-thioguanine (6-TG), MMR-proficient cells repeatedly and unsuccessfully attempt to process chemically induced mispairs. This futile cycling of the MMR pathway is believed to signal a G2 checkpoint arrest and apoptosis. Damage induced by IR is also recognized by MMR, resulting in MMR-mediated cytotoxicity (Zeng et al., 2000; Fritzell et al., 1999). Thus, MMR-deficient cells can be resistant to both chemotherapy and radiotherapy.
1.2.6.1 Mechanism of tolerance

There are several mechanisms of chemo-tolerance in breast carcinomas. Some of them are discussed below.

a) ABC Transporters

Multi drug resistance (MDR) is a phenomenon that causes resistance to several kinds of drugs simultaneously. One of the main mechanisms of MDR is the overexpression of drug transporters called ATP binding cassette (ABC) transporters, also called as efflux pumps (Cole et al., 1998; Balaji et al., 2016). There are 48 members of this superfamily that have been categorized into 7 families as ABC A-G, and majorly 16 ABC transporters are involved in human diseases. Among these, ABCB1 (p-glycoprotein/MDR1), ABCC1 (MRP1) and ABCG2 (BCRP1) are the major drug transporters which have been widely implicated in drug resistance in several cancers. These transporters are found to be overexpressed in several cancers including lung, breast and pancreatic cancers, and their overexpression is inversely correlated with patient survival. Several reports have shown that treatment with chemotherapeutic drugs increases the expression of ABC transporters in vitro (Saxena et al., 2011) and a critical dose the drug is required in the cells to develop drug resistance (Tsou et al., 2015). Consistent with this, chemotherapy-treated patient samples have been shown to have elevated expression of ABC transporters. Indeed inhibition of ABC transporters leads to chemo-sensitization, suggesting that inhibitors of ABC transporters can improve treatment outcome. ABCC3 is overexpressed in BC samples and cell lines and its knockdown increases the doxorubicin retention capability of the BC cells (Balaji et al., 2016). Recent reports suggest that overexpression of ABCB1 leads to increase in the gene expression of CD44 in breast cancer cells (Tsou et al., 2015). Therefore, understanding the expression and functional significance of specific ABC transporters will help in the development of novel treatment strategies aimed at targeting or reducing their expression in order to achieve better treatment response.

b) Hypoxia

Hypoxic cells are known to be more resistant to radiotherapy and chemotherapy than normoxic cells are (Shannon et al., 2003). Hypoxic cell populations within tumors are
believed to be a significant reason for radiotherapy failures, and, indeed, the clinical targeting of hypoxic cell populations is associated with improved loco-regional control and overall survival (OS) (Overgaard, 2007). Intra-tumoral hypoxia has been identified as an adverse indicator for patient prognosis independent of all of the histopathological parameters of BC (Vaupel et al., 2009). Hypoxia within a solid tumor arises from an increase in O2 utilization due to an increase in rapidly dividing cancer cells, and a decrease in oxygen availability due to structurally and functionally abnormal vessels of solid tumors (Semenza et al., 2010). Hypoxia not only mediates resistance to therapy, it also promotes genetic instability and aggressive mutagenesis, in part by impairing DNA repair pathways in tumor cells. Recent data support the idea that, under acutely hypoxic conditions, the check point kinases ATM and ATR are activated and limit DNA damage through cell cycle arrest (Bindra et al., 2007). Chronic hypoxic responses developed by tumor cells, downregulate MMR and HR pathways genes like MLH1, MSH2, BRCA1 and Rad51 (Bindra et al., 2004, 2005; Koshiji et al., 2005; Shahrzad et al., 2005; Mihaylova et al., 2003). Moreover, chronically hypoxic cells display increased sensitivity to crosslinking agents cisplatin and mitomycin C (Chan et al., 2008).

Cancer cells respond to decreased oxygen availability by increasing the activity of the hypoxia-inducible factors (HIFs), HIF-1 and HIF-2 (Semenza et al. 2012). HIF target genes encode proteins involved in cell survival, angiogenesis, metabolic reprogramming, immortalization, EMT, stem cell maintenance, resistance to radiation and chemotherapy, invasion and metastasis (Wang et al., 1995).

c) Epigenetic modulation

Accumulating evidences have implicated that epigenetic mechanisms plays an important role in acquiring drug resistance which does not necessarily need a stable heritable genetic alteration (Glasspool et al., 2006). A subpopulation of cancer cells that transiently exhibit a distinct phenotype characterized by the engagement of IGF-1R activity, hypersensitivity to HDAC inhibition, altered chromatin, and an intrinsic ability to tolerate drug exposure, which does not involve drug efflux (Sharma et al., 2010). Reversible drug tolerance appears to reflect dynamic heterogeneity within a cancer cell population that can be established even following the clonal expansion of single drug-sensitive cells.
Such phenotypic heterogeneity has been observed in some clonally derived normal mammalian cells, such as stem cells (Chang et al., 2008; Stewart et al., 2006), and has been implicated in cancer cell fates following drug exposure in culture (Cohen et al., 2008; Gascoigne and Taylor, 2008). A transiently maintained drug-tolerant state could provide a mechanism that allows a small subpopulation of tumor cells to withstand an initial onslaught of drug or other stressful stimuli to enable their survival for a period of time until more permanent resistance mechanisms can be established. This is highly reminiscent of the properties of antibiotic-tolerant bacterial subpopulations, also called “persisters,” which similarly exhibit a transient ability to endure potentially lethal stresses (Balabanet al., 2004; Dhar and McKinney, 2007). These are slower growing cells, whose survival within a more rapidly proliferating cell population is ensured by the fact that they can readily revert to a non-persister state via epigenetic mechanisms (Lewis, 2007). Reversible drug tolerance may account for accumulating clinical reports demonstrating that cancer patients treated with a variety of anticancer drugs can be successfully re-treated with the same drug after a “drug holiday.” The detection of a distinct chromatin state in drug-tolerant cancer cells and consequent hypersensitivity to HDAC inhibitors potentially yields a therapeutic opportunity to prevent the development of stable drug resistance (Sharma et al., 2010).

d) Intra-tumor heterogeneity
Breast cancer is a heterogeneous disease characterized by variant pathological features, disparate responses to therapeutics and substantial differences in long term patient survival (Polyak, 2007). As discussed earlier (Section 1.2.3.5), intra-tumor genotypic and phenotypic heterogeneity is hence a defining characteristics of human breast tumors. Certain molecular subtypes of breast cancer were found to be associated with specific genetic alterations. For instance, HER2+ and basal-like breast cancers exhibit a high rate of somatic mutation in the TP53 tumor suppressor gene (72% to 80%), whereas other molecular subtypes exhibit TP53 gene mutations much less frequently (12% to 29%) (Cancer Genome Atlas 2012). Luminal A, luminal B, and HER2+ subtypes exhibited significant rates of mutation in the PIK3CA gene (45%, 29%, and 39% respectively), whereas basal-like breast cancers are rarely associated with mutation of this gene (9%) (Cancer Genome Atlas 2012). Copy number variations reflecting gene deletions and
amplifications were found to affect numerous genes and gene regions, including amplifications in PIK3CA and ERBB2 chromosomal regions and deletions in TP53 and MAP2K4 chromosomal regions, among others (Cancer Genome Atlas 2012). Clonal diversity may be critically important when tumor cells encounter selective pressures, including the immune system, hypoxia, nutrient deprivation, geographic isolation, and, perhaps strongest of all, chemotherapy (Burrell et al., 2013; Merlo et al., 2006). Although several studies have investigated intra-tumor heterogeneity prior to chemotherapy and often report that increased diversity correlates with resistance (Chen et al., 2012; Mroz et al., 2013), few have analyzed the diversity of tumor cells in post-treatment samples. The study and treatment of cancer is complicated by this heterogeneity, as small tissue samples, typically obtained by biopsy, may not be representative of the whole tumor (Gerlinger et al., 2012) and a treatment that targets one tumor cell population may not be effective against another (Turner and Reis-Filho, 2012; Yap et al., 2012). Challenging a population of tumor cells with a drug could have several outcomes on the diversity of the population. The therapy could completely eradicate the tumor cell population or lead to a population bottleneck followed by the reconstitution of the tumor mass by the resistant clones. Alternatively, the drug could have no effect on the diversity of the tumor cell population (Navin, 2014). The current paradigm holds that tumor cells are diverse populations that harbor resistance mutations in rare clones (Merlo et al., 2006; Nowell, 1976). In contrast to genetic diversity, Almendro et al. (2014) did observed significant changes in phenotypic diversity. In most of the patients, the tumor cells shifted from a differentiated phenotype (CD24+/CD44-) to a mesenchymal phenotype (CD24-/CD44+) after therapy. This pattern occurred in all of the subtypes except for the Her2 tumors. The change in phenotypes was also accompanied by a decrease in overall cell proliferation (Ki67 staining), suggesting that the chemotherapy may have eliminated most of the rapidly dividing cells while leaving the slowly proliferating cells intact. These data suggest that cell phenotypes play a critical role in developing resistance to neoadjuvant therapy in breast cancer. These studies highlight the importance of cellular phenotypes in developing resistance to chemotherapy and the clinical value of measuring diversity to predict complete response.