1 GENERAL INTRODUCTION

1.1 Cancer

1.1.1 Epidemiology

According to the World Health Organization, cancer is the second leading cause of death in developed countries and is among the three leading causes of death for adults in developing countries (WHO). A steadily increasing proportion of elderly people in the world will result in approximately 16 million new cases of cancer by the year 2020 (IARC).

1.1.2 Multi-stage theory of carcinogenesis

The current view of cancer is that a malignancy arises from a transformation of the genetic material of a normal cell, followed by successive mutations, ultimately leading to the uncontrolled proliferation of progeny cells (Renan 1993). This view can be traced to seminal works of many scientists: Boveri formulated the “somatic mutation” hypothesis of the origin of cancer (Boveri 1929), Berenblum and Shubik demonstrated the multistep, sequential nature of carcinogenesis (Berenblum and Shubik 1949), Knudson demonstrated that an inherited defect in one allele of a protective gene (tumor suppressor) can predispose a person to cancer (Knudson 1985), and Nowell proposed the clonal evolution model of tumor progression which postulated that when a cell acquires a specific genetic alteration it may develop a

proliferative advantage and clonal expansion of the cell, driven by successive mutation, could lead to tumor progression (Nowell 1976).

Two hypotheses about the mechanism of carcinogenesis were proposed in the early 1950s, derived from analysis of cancer mortality statistics (Armitage and Doll 1954). The first hypothesis by Fisher and Hollomon postulated that the relationship between age and mortality could result if a colony of six or seven cancer cells was a critical size below which independent growth was not sustained (Fisher and Hollomon 1951). Fisher and Hollomon also hypothesized that cancer incidence should be proportional to the fifth or sixth power of the concentration of the effective carcinogen. However, experimental data suggests that tumor incidence and concentration of the carcinogen vary in proportion (Armitage and Doll 1954). The second hypothesis by Nordling suggested that the observed relationship between the logarithm of death rate and logarithm of age in cancer, could be explained if a cancer cell was the end-result of seven successive mutations (Nordling 1953). This hypothesis holds true only if the probability of occurrence of each mutation remains constant throughout life, and as long as the occurrence of each mutation is a relatively rare event (Armitage and Doll 1954).

Those hypotheses have been mathematically modeled (Armitage and Doll 1954, Renan 1993). The results strongly support mutation as one of the dominant factors in setting rate-limiting steps in tumor progression (Spencer et al 2006). Tumorigenesis is thought to require four to six stochastic rate-

limiting mutation events to occur in the lineage of a single cell (Nowell 1976, Armitage and Doll 1954, Renan 1993).

1.1.3 A genetic model for colorectal tumorigenesis

Based on the work of the afore-mentioned scientists, new genetic models of tumorigenesis have emerged and enhanced our current understanding of cancer genetics. One of these models is the first genetic model for colorectal tumorigenesis (Fearon and Vogelstein 1990, Figure 1-1).

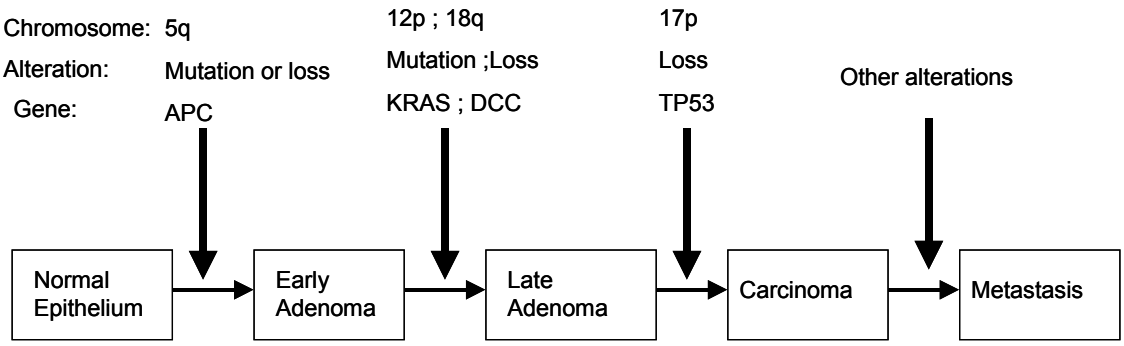

Figure 1-1. A genetic model for Colorectal Tumorigenesis. Tumorigenesis proceeds through a series of genetic alterations involving oncogenes (KRAS) and tumor suppressor genes (APC, TP53; particularly those on chromosomes 5q, 17p, and 18q). Adapted from Fearon and Vogelstein (1990).

The principles of Fearon and Vogelstein's genetic model for colorectal tumorigenesis are as follows: first, colorectal tumors appear to arise as a

result of the mutational activation of oncogenes coupled with the mutational inactivation of tumor suppressor genes; the latter changes predominate. Second, mutations in at least four to five genes are required for the formation of a malignant tumor. Third, although the genetic alterations often occur according to a preferred sequence, the total accumulation of changes, rather than their order, is responsible for determining the tumor's biologic properties. Fourth, in some cases, mutant tumor suppressor genes appear to exert a phenotypic effect even when present in the heterozygous state; thus some tumor suppressor genes may not be "recessive" at the cellular level. The general features of this model may be applicable to other common epithelial neoplasms (Fearon and Vogelstein 1990).

1.1.4 The Hallmarks of Cancer

A decade after Fearon and Vogelstein's genetic model of colorectal tumorigenesis, Hanahan and Weinberg presented general rules that govern the transformation of normal human cells into malignant cancers (Hanahan and Weinberg 2000, Figure 1-2).

Hanahan and Weinberg suggest that the vast catalog of cancer cell genotypes is a manifestation of six essential alterations in cell physiology that collectively dictate malignant growth (Figure 1-2): self-sufficiency in growth signals, insensitivity to antigrowth signals, evasion of apoptosis, limitless replicative potential, sustained angiogenesis, and tissue invasion and metastasis (Hanahan and Weinberg 2000). Hanahan and Weinberg also propose that these six capabilities are shared in common by all types of

human tumors and that the multiplicity of defenses may explain why cancer is relatively rare during an average human lifetime (Hanahan and Weinberg 2000).

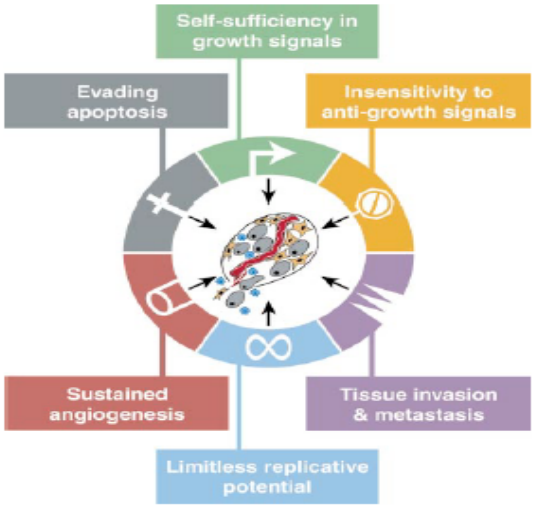

Figure 1-2. Acquired Capabilities of Cancer. Most if not all cancers have acquired the same set of functional capabilities during their development, albeit through various mechanistic strategies. Adapted from Hanahan and Weinberg (2000).

Hanahan and Weinberg propose that although virtually all cancers must acquire the same six hallmark capabilities, their means will vary both mechanistically and chronologically across tumor types. It is also possible that in some tumors, a particular genetic lesion may confer several

capabilities simultaneously, decreasing the number of distinct mutational steps required to complete tumorigenesis (Hanahan and Weinberg 2000).

1.1.5 Cancer Genes

1.1.5.1 Definition

A cancer gene is defined as a mutated gene causally implicated in oncogenesis (Futreal et al. 2004).

1.1.5.2 Mutations in Cancer Genes

1.1.5.2.1 Dominant Cancer Genes

Proto-oncogenes are genes which function to regulate normal cell proliferation and differentiation (Weinberg 2007). Alterations of these genes, gain-of-function mutations, can lead to overly active growth-promoting genes, which appear in cancerous cells as activated oncogenes (Vogelstein and Kinzler 2004). Proto-oncogenes are usually dominantly acting genes at the cellular level (Futreal et al. 2004). The somatic mutations that cause activation of oncogenes are characterized by mutations that cause structural changes to the encoded protein, such as point mutations and chromosomal translocations (Vogelstein and Kinzler 2004). Proto-oncogenes can also be transformed to oncogenes by elevated expression through gene amplification or chromosomal translocations (Weinberg 2007).

1.1.5.2.2 Recessive Cancer Genes

Tumor suppressor genes (TSG) operate to suppress cell proliferation through many biochemical mechanisms, and are often inactivated in various ways in cancer cells (El-Deiry 2003). An inherited mutant copy of a TSG increases susceptibility to specific types of cancer (Weinberg 2007). The Knudson hypothesis of TSG inactivation postulates that mutant alleles of TSG are recessive at the cellular level. Therefore both alleles of a TSG must be inactivated, loss-of-function mutations, in the transformation of normal cells to cancerous cells (Knudson 1985). The loss of TSG function can occur either by genetic mutation or epigenetic silencing of genes via promoter methylation (Baylin 2005). Inactivation of TSG by mutation or methylation of one allele may be followed by other mechanisms that facilitate loss of the second copy (Knudson 2002), such as loss of heterozygosity (LOH) at the TSG locus (Sherr 2004).

1.1.5.3 Cancer Gene Census

A census from the literature of reported cancer genes was recently compiled and it indicates that mutations in more than 1% of genes contribute to human cancer (Futreal et al. 2004). Of the then 291 reported cancer genes, 90% show somatic mutations that are acquired in cancer, 20% show germline mutations that predispose to cancer, and 10% show both (Futreal et al. 2004, Figure 1-3).

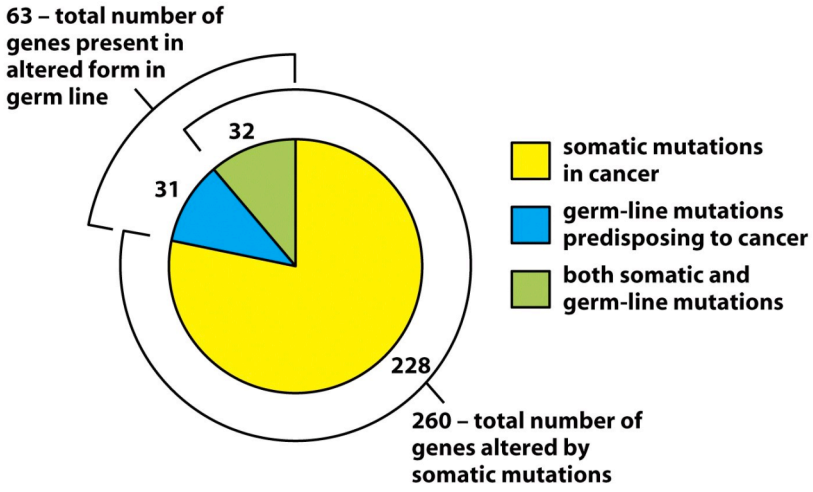

Figure 1-3. Schematic of the relative number of cancer genes with reported germline and somatic mutations in cancer. The Biology of Cancer (© Garland Science 2007).

The census reports that the most common mutation class among known cancer genes is a chromosomal translocation (Futreal et al. 2004). Seventy percent of the cancer genes, altered by chromosomal translocations, are implicated in leukemias, lymphomas and sarcomas but only represent 10% of human cancer. Non-solid tumors such as leukemias are easier to analyze with cytogenetic techniques compared to solid tumors, which in part accounts for this bias.

Solid tumors have many translocations but few have been analyzed in great detail, except for sarcomas. The remaining 90% of cancers of epithelial

origin, have been shown to be altered by other types of mutations: base substitutions that lead to missense amino-acid changes, nonsense changes, alterations in conserved splice site positions, insertions or deletions in coding sequences or splice sites that may cause in-frame or frameshift alterations of the protein.

Another finding from the census indicates that the most common protein family domain (Pfam domain) encoded by cancer genes is the protein kinase. Following the protein kinases, the second and third most common Pfam domains encoded by cancer genes are the Pfam domain broadly involved in transcriptional regulation, and the Pfam domain involved in DNA maintenance and repair, respectively.

1.1.5.4 Twenty-four cancer genes analyzed

More than sixty cancer genes are causally implicated in cancer through the acquisition of somatic small intragenic mutations (Futreal et al. 2004). These mutations include base substitutions that can cause missense, nonsense or splice site changes. They also include small insertions or deletions that may be in frame or out of frame causing premature termination codons. Because analysis by sequencing is laborious, compilation of these types of change in human cancer cell lines has not been carried out systematically. To fill this gap, twenty-four of the cancer genes activated or inactivated by small intragenic mutations were selected for sequence analysis based on mutation frequency, biological interest, and because some

represent druggable mutated targets. Below, I present a discussion of the twenty-four cancer genes analyzed in this thesis.

1.1.5.4.1 *APC* : adenomatosis polyposis coli

APC is a tumor suppressor gene located on chromosome 5q21-q22 and is composed of 15 exons. The predicted encoded protein has 2843 amino acids, and is involved in the WNT signaling pathway and intercellular adhesion (Fearnhead et al. 2001). Familial adenomatous polyposis (FAP) is a hereditary colon cancer predisposition syndrome characterized by adenomatous polyps of the colon and rectum. Gardner syndrome is a hereditary colon cancer predisposition syndrome predominantly characterized by the development of extracolonic features such as adenomatous polyps of the upper gastrointestinal tract, congenital hypertrophy of the retinal pigment epithelium (CHRPE), mesenteric fibromatosis, hepatoblastoma, and papillary carcinoma (OMIM 175100). Turcot syndrome is associated with multiple colorectal polyps and cerebellar medulloblastoma (Fearnhead et al. 2001). The gene responsible for these syndromes, adenomatous polyposis coli (*APC*) was mapped by genetic linkage analysis of families affected by FAP directed by an interstitial deletion on chromosome 5q (El-Deiry 2003). Subsequent positional cloning and genetic polymorphic screens led to the identification of truncating mutations in the *APC* gene (El-Deiry 2003).

Further studies showed that *APC* is somatically mutated in up to 90% of all sporadic colon adenocarcinomas (Miyaki et al. 1994), 20% of gastric cancers, 16% of pancreatic cancers,15% of duodenal cancers, 14% of liver

cancers, 11% of sarcomas, and 9% of ovarian cancers (COSMIC). Both alleles of the *APC* gene are inactivated in sporadic tumorigenesis. However one inactivated allele is present in the germline of hereditary syndromes such as FAP and Gardner's syndrome, with a somatically acquired inactivation of the second allele during formation of colorectal tumors (Fearnhead et al. 2001). The loss of function mutations are generally characterized by frame-shift or single base substitutions leading mostly to truncated forms of the APC protein (Figure 1-4). These loss-of-function mutations of *APC* are present in the early stages of colon cancer development and precede other alterations observed during colon cancer development (Vogelstein et al. 1988, Powell et al. 1992). The mutations observed in germline and sporadic colon cancer are found most frequently in the 5' end of exon 15, termed the mutation cluster region (MCR) (Miyoshi et al. 1992, Beroud and Soussi 1996). The MCR is implicated in β-catenin (CTNNB1) interactions.

The WNT pathway controls cell fate during embryonic development and serves as a key regulator of homeostasis in adult self-renewing tissues (Gregorieff and Clevers 2005). Mutational deregulation of the WNT cascade is closely associated with malignant transformation (Gregorieff and Clevers 2005). The WNT signaling pathway genes were discovered in Drosophila and shown to be conserved in all metazoans (Wodarz and Nusse 1998). The interaction of *WNT* genes with transmembrane proteins such as Frizzled (FZ) are characterized as the initial steps in the canonical pathway, leading to the formation of nuclear TCF/CTNNB1 complexes (Gregorieff and Clevers 2005).

APC protein modulates the oncogenic WNT signal transduction cascade through its cellular effects on CTNNB1 (Goss and Groden 2000).

The role of APC, in the absence of Wnt signals, is to recruit CTNNB1 to the APC-CTNNB1-Axin complex where GSK3β phosphorylates all three proteins leading to CTNNB1 degradation by ubiquitin ligases. Truncation of the APC protein, including the CTNNB1 binding region, renders it unable to recruit CTNNB1 for degradation by ubiquitin ligases. This leads to accumulation of CTNNB1 in the nucleus (Goss and Groden 2000). This accumulation of CTNNB1 is associated with changes in the transcriptional activation by TCF/LEF transcription factors and expression of genes involved in cellular proliferation and differentiation, including *MYC* and *CCND1* (Nathke 2004).

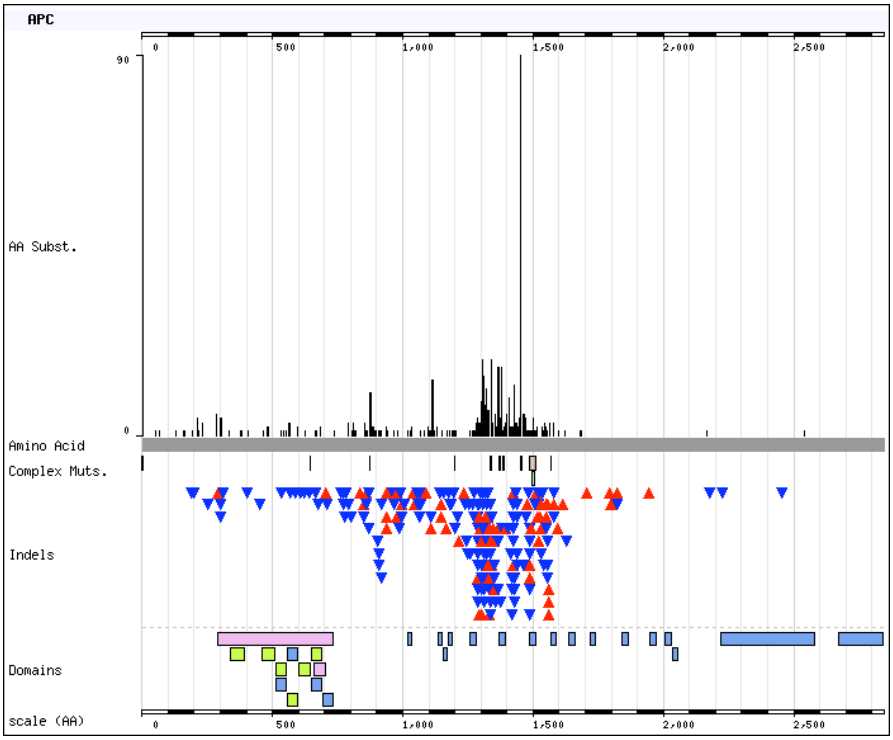

Figure 1-4: Distribution of somatic mutations in APC. A schematic of the APC protein displays base substitutions most of which cause nonsense mutations. Complex mutations, insertions (red triangle) and deletions (blue triangle) result in frame-shift mutations leading to truncation of APC. A similar legend applies for the representation of somatic mutations in the rest of the genes discussed in this section. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.2 *CTNNB1:* catenin (cadherin-associated protein), beta 1

CTNNB1 is a dominantly acting cancer gene located on chromosome 3p22-p21.3 and is composed of 16 exons. The predicted encoded protein has 781 amino acids, and is an adherins junction protein critical for the establishment and maintenance of epithelial layers (OMIM 116806). CTNNB1, a cytoplasmic protein binds to the epithelial cell-cell adhesion

molecule, E-cadherin. *CTNNB1* was cloned in Xenopus laevis and identified as a homolog of the armadillo cell polarity protein in Drosophila and shown to be associated with E-cadherin (McCrea et al. 1991). By fluorescence in-situ hybridization (FISH), *CTNNB1* was mapped to 3p21, a region affected by somatic alterations in a variety of tumors (Kraus et al. 1994, Trent et al. 1995). *CTNNB1* partial deletion was observed in the investigation of the mechanism of E-cadherin- dependent-cell-cell adhesion dysfunction in gastric carcinoma cells (Kawanishi et al. 1995).

Somatic mutations of *CTNNB1* are dominantly acting and are characterized by either missense or in-frame deletions (Figure 1-5). The majority of *CTNNB1* mutations cluster in exon 3, corresponding to the phosphorylation sites on the protein important for GSK3β mediated degradation in the canonical WNT pathway (Gregorieff and Clevers 2005). The somatic mutations usually affect only one of the two *CTNNB1* alleles (Morin et al. 1997). *CTNNB1* mutations occur most frequently in thyroid cancers (27%), duodenal cancers (23%), pancreatic cancers (22%), endometrial (21%), and liver (21%) cancers (COSMIC). *CTNNB1* mutations are infrequent in colon cancer and occur in approximately 6% of colon cancers (COSMIC). In colorectal cancer, *CTNNB1* mutations are found where there is no *APC* mutation (Morin et al. 1997).

The somatic mutations of *CTNNB1* often render CTNNB1 insensitive to APC/CTNNB1/GSK-3β mediated degradation (El-Deiry 2003). Consequently an increase in CTNNB1:TCF-mediated transcription (Chung 2000) results in

the overexpression of *CCND1* and *MYC*, genes involved in proliferation and growth (He et al. 1998, Tetsu et al. 1999, Mann et al. 1999).

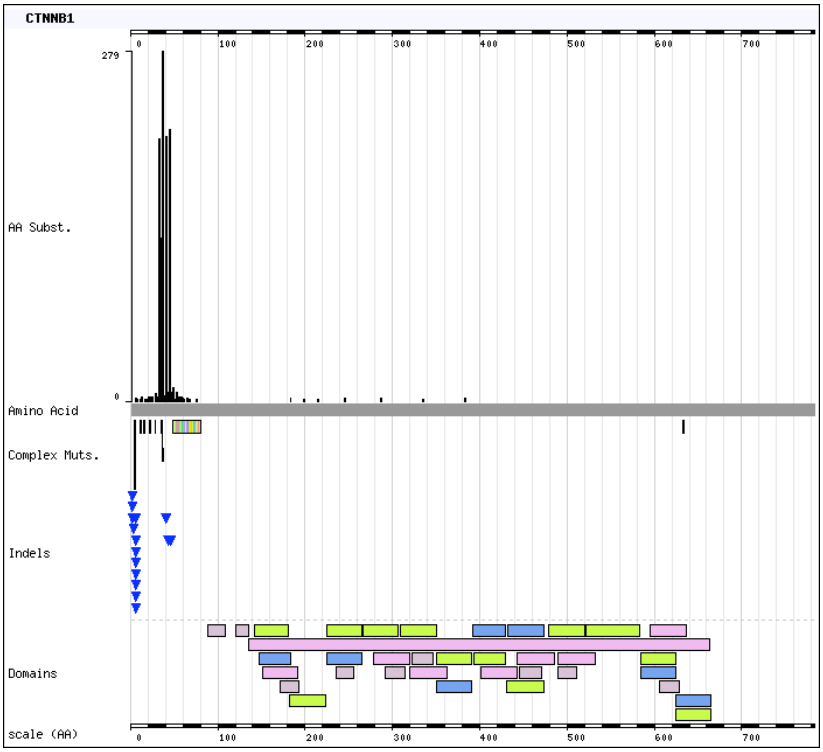

Figure 1-5: Distribution of somatic mutations in CTNNB1. A schematic of the CTNNB1 protein displays mostly missense amino acid substitutions and in-frame deletion mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.3 *MADH4:* Mothers Against Decapentaplegic, Drosophila, Homolog Of , 4; SMAD4

MADH4, a tumor suppressor gene, is located on chromosome 18q21.1 and is composed of 11 exons. The predicted encoded protein has 552 amino acids and plays a pivotal role in signal transduction of the transforming growth factor beta (*TGFβ*) superfamily cytokines (Maurice et al. 2001). *MADH4* was initially mapped on the basis of LOH at the 18q region in more than 50% of pancreatic cancers pointing to the presence of a tumor suppressor gene in that region (Hahn et al. 1996). Fine mapping of the 18q region suggested that MADH4 is also deleted in colorectal cancers (Thiagalingam et al. 1996).

Familial juvenile polyposis (FJP) is an autosomal dominant gastrointestinal hamartomatous polyposis syndrome associated with an increased risk of gastrointestinal cancer and colon cancer (OMIM 174900). A locus for FJP mapped to 18q21.1 by genetic linkage analysis of families. Since *MADH4* was in the region defined by linkage analysis it was analyzed and nonsense, frame-shift, LOH and missense changes in *MADH4* were found in affected individuals (Howe et al. 1998). Therefore, *MADH4* is responsible for FJP.

The MADH4 protein was found to bear approximately 85% similarity to the Drosophila Mad protein and Caenorhabditis elegans sma-2, -3, and –4 proteins (Hahn et al. 1996). *MADH4* somatic mutations are found in 23% of pancreatic cancers, 22% of thyroid cancers, 19% of duodenal cancers, and 10% of sporadic colorectal cancers (COSMIC). The distribution of somatic

mutations of *MADH4* (Figure 1-6) is similar to that of germline (Thiagalingam et al. 1996, Schutte et al. 1996).

TGF-β is involved in apoptotic signaling. It binds to and activates specific cell surface receptors with intrinsic serine/threonine kinase activity (ten Dijke and Hill 2004). In turn, these activated receptors stimulate phosphorylation of MADH1 and MADH2 proteins, which form complexes with MADH4. The complex of MADH1/MADH2/MADH4 also bind another partner, FAST-1, accumulate in the nucleus and regulate the transcription of target genes (Liu et al. 1997). In particular, MADH4 promotes the binding of MADH2/MADH4/FAST-1 complex to DNA as well as activating MADH1 and MADH2 to stimulate transcription (Liu et al. 1997). In the absence of SMAD proteins epithelial cancers evade the growth inhibitory actions of TGF-β (Weinberg 2007).

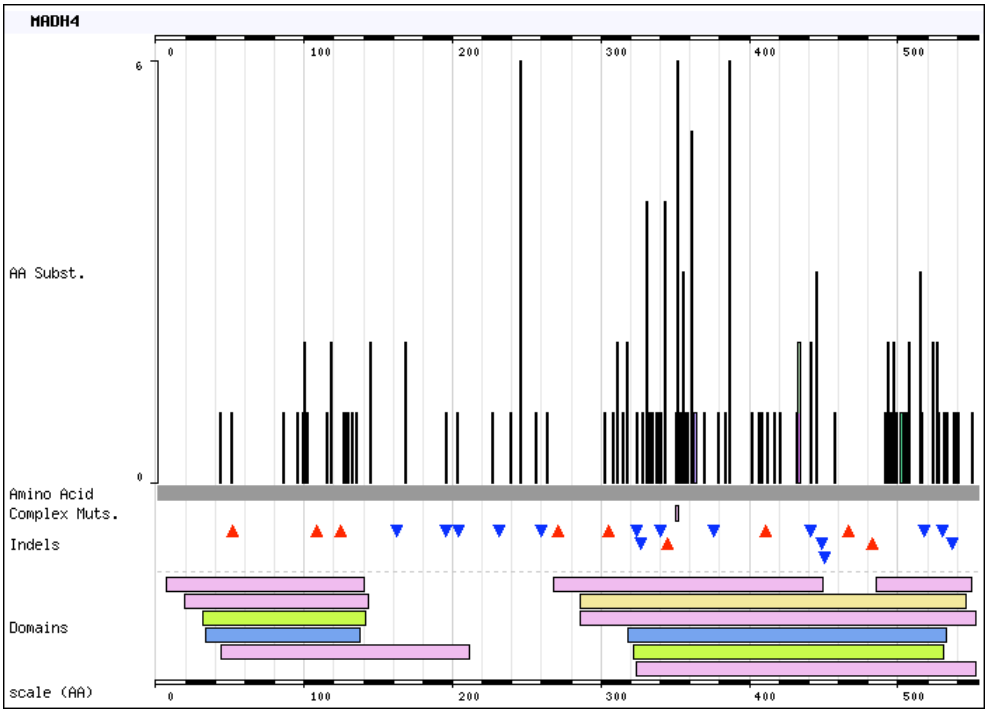

Figure 1-6: Distribution of somatic mutations in MADH4. A schematic of the MADH4 protein displays missense amino acid substitutions, frame-shift and nonsense mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.4-6 *RAS* : rat sarcoma viral oncogene

The *RAS* family of genes is composed of *HRAS, KRAS, and NRAS*. The *RAS* genes are dominant cancer genes and encode proteins involved in cellular signaling (Malumbres and Barbacid 2003). Intensive study of *RAS* genes began with the observation that a preparation of a murine leukemia virus from a leukemic rat, was able to induce sarcomas in new born rodents (Harvey et al. 1964). Another viral oncogene, Kirsten-MSV was obtained by serial passage of murine leukemia viruses through Wister-Furth rats (Kirsten et al. 1966). These mouse leukemic cancer cells containing viral oncogenes,

v-h-ras and v-k-ras were shown to be transforming in NIH-3T3 mouse fibroblasts (Shih et al. 1979).

The human oncogenes *HRAS* and *KRAS* were identified by the use of retroviral oncogene probes to hybridize DNA isolated from NIH-3T3 cells transformed with various human cancer cells (Malumbres and Barbacid 2003). The transforming human gene, *HRAS* was isolated by molecular cloning of DNA sequences from a human bladder cancer cell line (Der et al. 1982). The human *HRAS* sequence hybridized to the viral v-h-ras oncogene (Malumbres and Barbacid 2003). *KRAS* was also isolated by molecular cloning of DNA from a human lung cancer cell line (Der et al. 1982). The human *KRAS* sequence hybridized to the viral v-k-ras oncogene (Malumbres and Barbacid 2003). The transforming ability of a third gene, *NRAS* was demonstrated in NIH-3T3 cells transformed with human cancer cells. *NRAS* was isolated by molecular cloning from a human neuroblastoma cell line (Shimizu et al. 1983a). Analysis of the isolated DNA sequence revealed that the transforming gene of the neuroblastoma cell line was related to the previously identified *HRAS* and *KRAS* oncogenes (Shimizu et al. 1983b, Der et al. 1982, Parada et al. 1982). Therefore the new oncogene was termed *NRAS* (Malumbres and Barbacid 2003).

HRAS is a dominantly acting cancer gene located on chromosome 11p15.5 and is composed of 4 exons. Its predicted encoded protein has 189 amino acids. *HRAS* was the first reported human cancer gene, and was found altered by a single heterozygous missense mutation (Tabin et al.

1982). Dominantly acting mutations in *HRAS* were the first report of cancer causing mutations in human cancer (Weinberg 2007). The missense mutations of the *HRAS* gene occur in less than 5% of all cancers. Missense mutations of *HRAS* are most prevalent in salivary gland, bladder, and cervical cancers (COSMIC).

Recently, germline mutations in *HRAS* were identified as the cause of Costello syndrome, a multiple congenital anomaly and mental retardation syndrome (Aoki et al. 2005). Costello syndrome is a rare abnormality associated with short stature, redundant skin of the neck, palms, soles, and fingers, and nasal papillomata (OMIM 218040). Malignancies of Costello syndrome individuals include bladder cancer and rhabdomyosarcoma (Schubbert et al. 2007). The distribution of germline mutations of *HRAS* in Costello syndrome are not identical, but overlap with that of somatic mutations found in cancer (Aoki et al. 2005). For example, one does not find the codon 12 glycine to valine substitution of *HRAS* in Costello syndrome. However other codon 12 variants such as glycine to alanine and glycine to serine have been identified in Costello syndrome (Aoki et al. 2005).

KRAS is a dominantly acting cancer gene located on chromosome 12p12.1 and is composed of 6 exons. Its predicted encoded protein has 188 amino acids. *KRAS* is mutated in approximately 20% of all cancers, particularly in colon, lung, and endometrial cancers (COSMIC). Gain-of-function mutations of *KRAS* have also been identified in the germline of individuals affected by Noonan syndrome, a dominant developmental disorder

characterized by short stature, facial dysmorphism, skeletal abnormabilities, cardiac defects, learning disabilities, and a predisposition to hematologic abnormalities (Schubbert et al. 2007). The germline mutations of *KRAS* introduce novel amino acid substitutions not found in cancer (Schubbert et al. 2006). Heterozygous germline variants in *KRAS* have also been found in individuals with cardio-facio-cutaneous (CFC) syndrome (Schubbert et al. 2007). CFC syndrome is characterized by distinctive facial appearance, heart defects, and mental retardation (OMIM 115150). Apparently, this syndrome can be caused by mutation of any of three families of signaling molecules in the RAS pathway: RAS, RAF or MEK (Schubbert et al. 2007).

NRAS is a dominantly acting cancer gene located on chromosome 1p13.2 and is composed of 2 exons. Its predicted encoded protein has 189 amino acids. *NRAS* is mutated in approximately 9% of all cancers, with a higher prevalence in melanoma, hepatocellular carcinomas, and myelocytic leukemias (COSMIC).

The *RAS* genes are mutated by missense base substitution mutations that most commonly affect either codon 12, 13, or 61 of the reading frame (Figures 1-7 to 1-9) (Weinberg 2007). Missense mutations at codon 19 (Akagi et al. 2007) and 146 (Edkins et al. 2006) of KRAS occur less frequently in cancer (Figure 1-8). Missense mutations at codon 18 of NRAS (Figure 1-9) have also been identified, and occur less frequently in cancer (Demunter et al. 2001). The RAS proteins are very similar in structure to each other. They are small G proteins exhibiting GTPase activity, and act as molecular switches

that cycle between inactive GDP-bound and active GTP-bound states (Mor and Philips 2006). The commonly occurring mutations at codon 12, 13, and 61 lock RAS in the GTP-bound, active state (Mor and Philips 2006). These dominantly acting mutations of the RAS proteins lead to the continued release of growth stimulatory signals in cells (Malumbres and Barbacid 2003). This can promote activation of many signaling pathways including the mitogen-activated protein kinase/ extracellular signal-regulated kinase (MAPK/ERK) pathway (Gire et al. 2000), phosphoinositol-3 kinase (PI3K) pathway, and the cellular Jun kinase (c-JNK) pathway. The activation of these downstream pathways results in uncontrolled cell growth and differentiation (Kolch et al. 2000).

The MAPK/ERK pathway is the best characterized of the signaling pathways that is regulated by RAS. MAPK/ERK pathway is activated when growth factors bind to receptor protein tyrosine kinases (Schlessinger 2000). The interaction of growth factors at the receptors leads to dimerization and cross-phosphorylation of tyrosine residues in their cytosolic domains (Mor and Philips 2006). The phosphorylation of tyrosine kinases recruits son of sevenless (SOS), a guanine nucleotide exchange factor (GEF) to the plasma membrane where it activates RAS (Schlessinger 2000). In turn, activated RAS recruits and activates RAF1, a serine/threonine kinase. Once active, RAF1 phosphorylates and activates MAPK/ERK kinase (MEK), a dual specificity tyrosine/threonine kinase. MEK in turn, phosphorylates and activates ERK1 and ERK2, serine/threonine kinases, transported into the

nucleus where they phosphorylate transcription factors (Mor and Philips 2006).

The mutant RAS proteins should be attractive targets for direct inhibition by small molecule inhibitors. However, to date, direct inhibition of activated RAS and hence its downstream effectors have not been effective in cancer therapy (Downward 2003). Farnesyltransferase inhibitors (FTIs), developed to inhibit the RAS proteins, have not been successful in clinical trials despite the fact that HRAS is post-translationally modified by farnesyltransferase. One problem is that KRAS and NRAS, which account for most RAS mutations in human tumors, can also be modified by geranylgeranyltransferase (GGT) (Downward 2003).

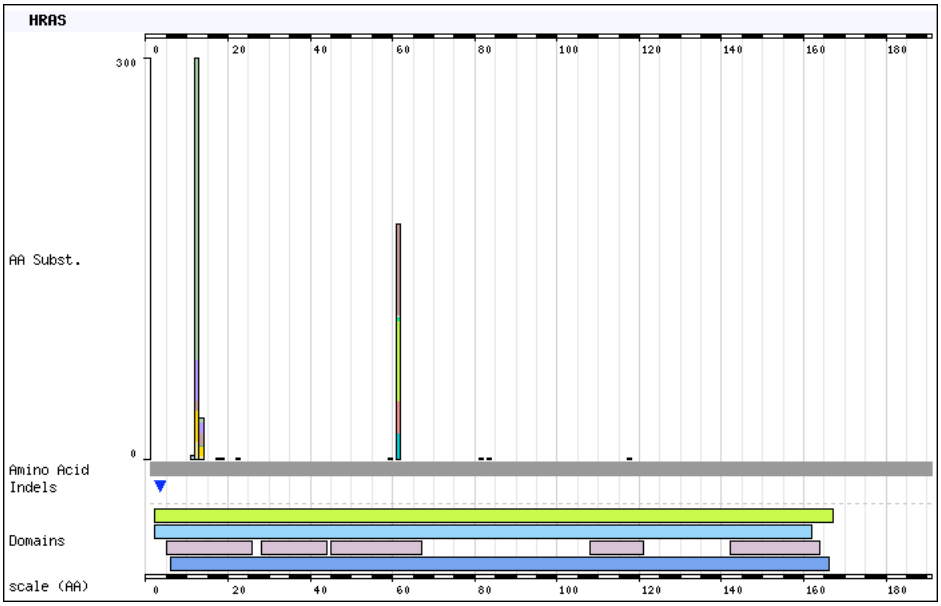

Figure 1-7: Distribution of somatic mutations in HRAS. A schematic of the HRAS protein displays missense amino acid substitutions occurring most frequently at codon 12 and 61. Adapted from Catalogue of Somatic

Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

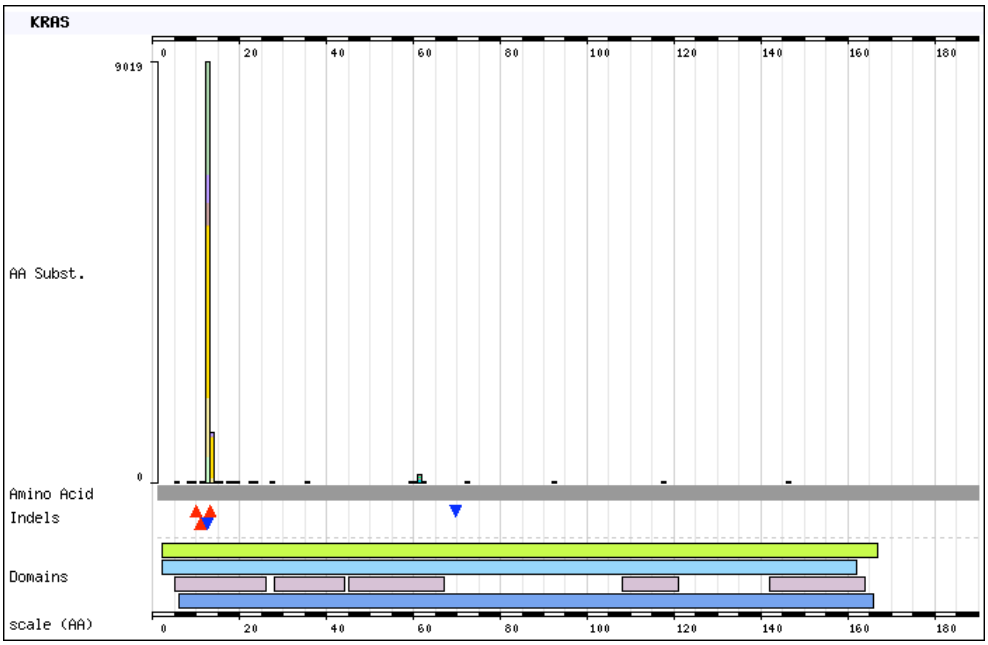

Figure 1-8: Distribution of somatic mutations in KRAS. A schematic of the KRAS protein displays missense amino acid substitutions occurring mostly at codon 12 and 13. To a lesser extent, amino acid substitutions occur at other codons such as 19, 61, and 146. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

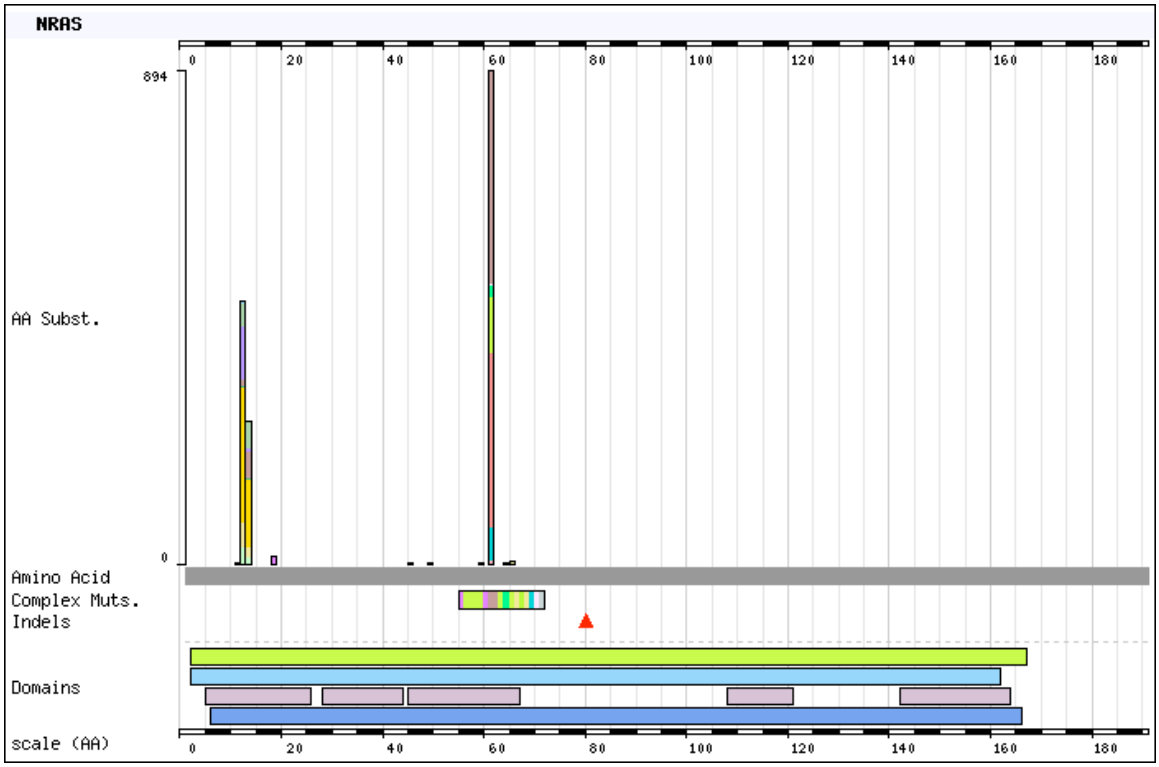

Figure 1-9: Distribution of somatic mutations in NRAS. A schematic of the NRAS protein displays missense amino acid substitutions occurring mostly at codon 61, and to a lesser extent at codons 12 and 13, and 18. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.7 *BRAF* : v-raf murine sarcoma viral oncogene homolog B1

BRAF is a dominantly acting cancer gene located on chromosome 7q34 and composed of 16 exons. Its predicted encoded protein has 766 amino acids and belongs to the *RAF* family of serine/threonine kinases (OMIM 164757). In mammalian genomes there are three highly conserved *RAF* genes, *ARAF*, *BRAF*, and *CRAF* whose protein products play distinct roles

(Dhomen and Marais 2007). *BRAF* was identified as the human homolog of the avian c-Rmil protooncogene (Eychene et al. 1992). High throughput sequencing of the exons and exon-intron boundaries of *BRAF* in a panel of human cancers of diverse origin, identified somatic mutations of *BRAF* (Davies et al. 2002). This finding implicated *BRAF* as an oncogene in human cancer. The somatic mutations of *BRAF* occur at a high frequency in melanomas (41%), thyroid cancers (35%), colon cancers (14%), and ovarian cancers (14%) (COSMIC). The somatic mutations are characterized by dominantly acting missense amino acid substitutions in conserved domains (Figure 1-10).

BRAF is also mutated in up to 80% of benign skin lesions, naevi, (Pollock et al. 2003) and in pre-malignant colon polyps (Rajagopalan et al. 2002). This finding implicates *BRAF* mutation in the initiation process of tumorigenesis. Notably, the V600E BRAF mutation although occurring in up to 70% of melanomas, does not seem to be UV-dependent (Dhomen and Marais 2007). Gain-of-function germline mutations of *BRAF* have also been found in individuals with cardio-facio-cutaneous (CFC) syndrome, a rare genetic disorder resulting in developmental defects with no cancer incidence (Dhomen and Marais 2007). The specific somatic *BRAF* mutations in cancer rarely occur in the germline disorders (Dhomen and Marais 2007). The germline mutations lead to elevated, but relatively modest, constitutive MEK signaling from BRAF, directly or through CRAF (Dhomen and Marais 2007).

There are over 40 different missense mutations in BRAF, involving 24 different codons. However, the predominant mutation in BRAF occurs on codon 600, converting valine to glutamate (V600E) (Davies et al. 2002). The resulting mutant BRAF protein was shown to have elevated kinase activity and transforming ability in NIH-3T3 cells (Davies et al. 2002). The codon 600 mutation accounts for more than 90% of the mutations in melanoma, thyroid, and colon cancer, but is relatively rare in non-small cell lung cancer (Davies et al. 2002, Rajagopalan et al. 2002, Fukushima et al. 2003, Kimura et al. 2003). The valine at codon 600 can also be mutated to other residues through tandem nucleotide changes that occur at much lower frequencies (Garnett and Marais 2004), (Wan et al. 2004). The activities of the resulting mutants are similar to that of V600E mutant.

The pattern of *BRAF* mutations in cancer reveals that *BRAF* mutations occur in many of the cancer types in which *RAS* is mutated. However coincident mutations are extremely rare (Davies et al. 2002, Rajagopalan et al. 2002). Therefore, *BRAF* mutations occur in cancers where there is strong selection for aberrant ERK signaling and mutations can occur at different levels in the pathway (Garnett and Marais 2004). The BRAF V600E mutation induces constitutive ERK signaling through hyperactivation of the RAS-MEK-ERK pathway, stimulating proliferation, survival and transformation (Dhomen and Marais 2007). An analysis of 22 BRAF mutants revealed that the majority of mutants exhibit elevated kinase activity by directly phosphorylating MEK. However, some BRAF mutants have reduced kinase activity toward MEK and instead activate CRAF to signal to MEK and ERK in cells (Wan et al. 2004).

The crystal structure of the BRAF kinase domains, wild type and V600E, in complex with a RAF inhibitor showed that the activation segment of BRAF is held in an inactive conformation by association with the P loop. Therefore, the clustering of mutations in the kinase domain and P loop of BRAF suggest that disruption of this interaction converts BRAF into its active state (Wan et al. 2004). A detailed discussion of the development of small molecule BRAF inhibitors will be included elsewhere in the introduction section of the thesis.

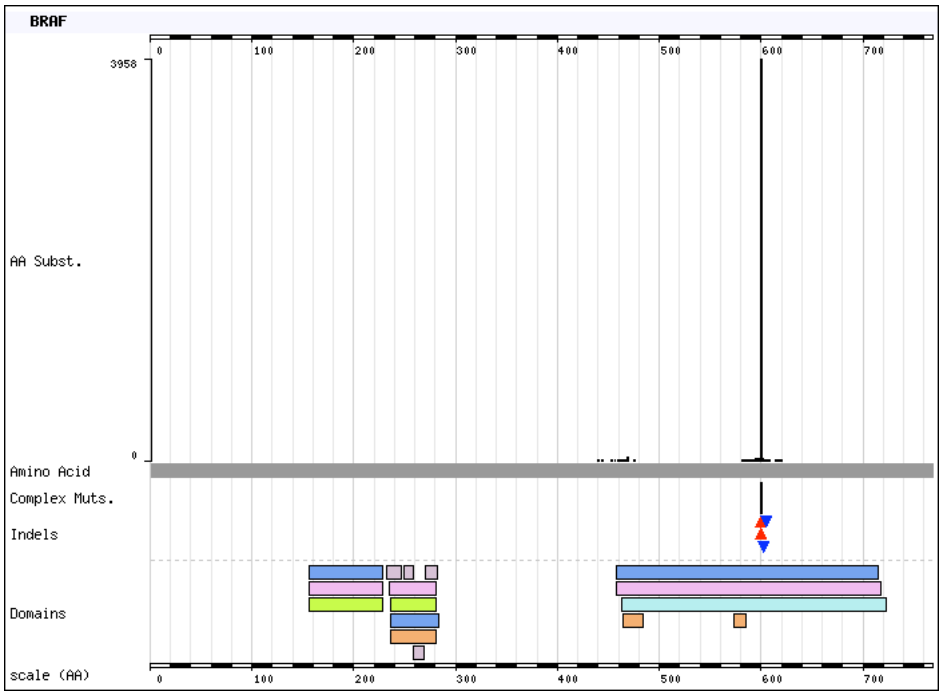

Figure 1-10: Distribution of somatic mutations in BRAF. A schematic of the BRAF protein displays amino acid substitutions, insertions and deletions occurring directly at and adjacent to codon 600. Adapted from Catalogue of

Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.8 *PIK3CA*: phosphoinositide-3-kinase, catalytic, alpha polypeptide

PIK3CA is a dominantly acting cancer gene, located on chromosome 3q26.3 and is composed of 20 exons. Its predicted encoded protein has 1068 amino acids and is a member of the Phosphatidylinositol 3-kinase (PI3K) lipid kinases, important regulators of cellular growth, transformation, adhesion, apoptosis, survival and motility (Vivanco and Sawyers et al. 2002). PI3Ks are heterodimeric lipid kinases composed of a catalytic and regulatory/adaptor subunit variants encoded by separate genes (Karakas et al. 2006). PI3K is composed of an 85-kD subunit that acts as an adaptor and an 110-kD (p110) catalytic subunit that couples to the 85-kD adaptor to activate protein tyrosine kinases (OMIM 171834). *PIK3CA* corresponds to the p110 catalytic subunit of *PI3K*. The kinase activity of PI3Ks was first reported with viral oncoproteins (Cantley et al. 1991). Mouse knockouts of both the regulatory and catalytic subunits of *PI3K* resulted in embryonic lethality, liver necrosis, and colorectal cancer (Katso et al. 2001).

Amplification of the *PI3K* locus resulted in elevated lipid kinase activity of the p110α catalytic subunit of PI3K (*PIK3CA*) in cancers, such as ovarian, breast, and gastric, implicating *PIK3CA* as an oncogene (Shayesteh et al. 1999, Ma et al. 2000, Katso et al. 2001). More recently, somatic missense mutations in *PIK3CA* have been identified by high-throughput sequencing of

eight genes in the *PI3K* family in several cancer types (Karakas et al. 2006). *PIK3CA* is mutated at a high frequency in endometrial cancer (36%), breast cancer (27%), and colon cancer (23%) (COSMIC). The dominantly acting somatic missense mutations are predominantly found in the helical and kinase domains of the PIK3CA subunit, with three hotspot mutations at codons 542, 545, and 1047 (Karakas et al. 2006), (Figure 1-11). Mutations in *PIK3CA* are associated with later stages of cancer progression (Karakas et al. 2006). The PI3K pathway is an important driver of cellular proliferation and survival, especially in cells responding to growth-factor-receptor engagement (Cully et al. 2006).

PTEN, a lipid phosphatase, negatively regulates the PI3K signaling pathway, functioning as a tumor suppressor (Stambolic et al. 1998). In the absence of *PTEN*, caused by deletion or inactivation, activation of *PI3K* effectors, such as protein kinase B (PKB/AKT), occur without stimulus, initiating tumorigenesis (Cully et al. 2006). The PI3K pathway can be activated by RAS recruitment of the p85 regulatory subunit of PI3K to the membrane by RAS. Subsequent activation of the catalytic subunit results in the generation of secondary messenger lipid phosphatidylinositol (3,4,5) triphosphate (PIP_3) (Cully et al. 2006). PIP_3 in turn recruits AKT to the membrane where it is phosphorylated and activated leading to evasion of apoptosis and increased proliferation (Cully et al. 2006). *In vivo* studies have demonstrated that the mutations in the *PIK3CA* gene are essential for tumor growth, therefore the gene is a clear target for the development of small molecule inhibitors (Samuels et al. 2005).

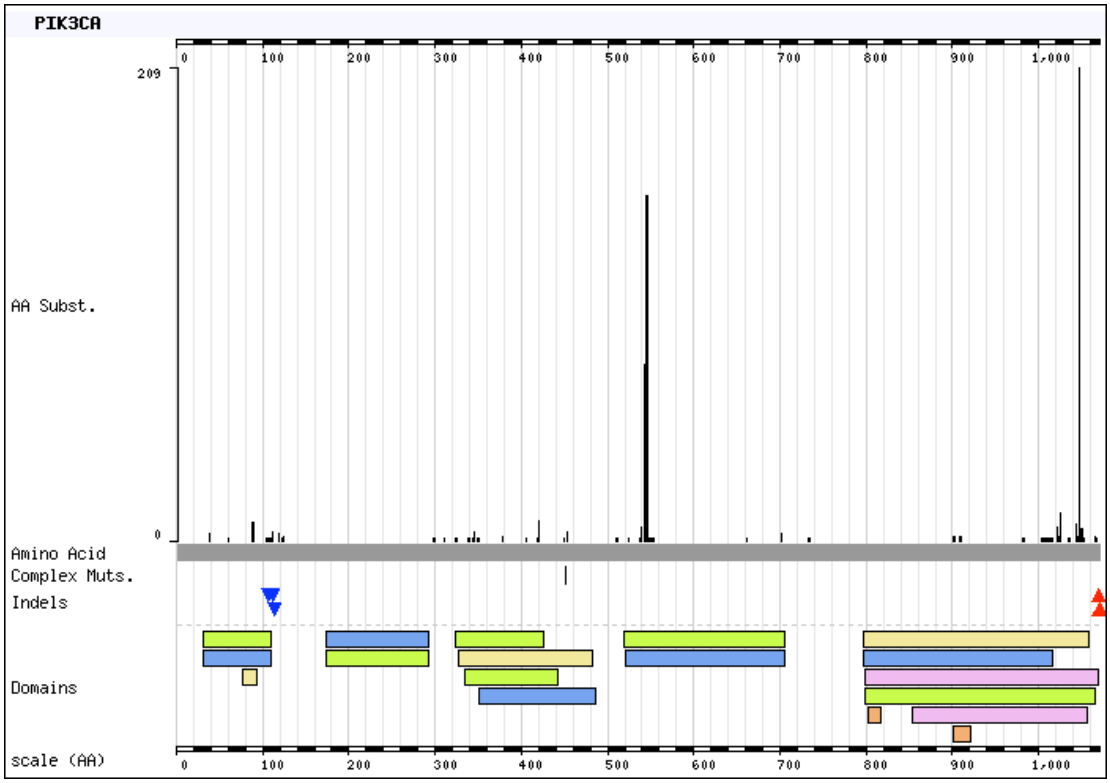

Figure 1-11: Distribution of somatic mutations in PIK3CA. A schematic of the PIK3CA protein displays missense amino acid substitutions at codon 542, 545 and 1047. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic

1.1.5.4.9 *PTEN* : phosphatase and tensin homolog

PTEN, a tumor suppressor gene, is located on chromosome 10q23 and is composed of 9 exons. Its predicted encoded protein has 403 amino acids and is a lipid phosphatase (OMIM 601728). *PTEN* was isolated by mapping homozygous deletions in human tumor cell lines and xenografts as it was the target of 10q22-25 deletions frequently involved in cancers of the prostate, breast, and glioblastoma (Lundgren et al. 1988, Gibas et al. 1984, Li et al.

1997). *PTEN* is frequently mutated in endometrial cancers (37%), glioblastomas (20%), melanomas (15%), and prostate cancers (13%) (COSMIC). The majority of somatic mutations of *PTEN* result in loss-of-function of the PTEN protein (Bonneau and Longy 2000). These inactivating mutations are characterized by nonsense, frame-shift, splicing mutations, and LOH resulting in truncation of the protein (Figure 1-12). Missense somatic mutations of *PTEN* cluster in exon 5, corresponding to the phosphatase core motif and are likely to alter the phosphatase activity of the protein (Bonneau and Longy 2000).

Germline *PTEN* mutations were detected in patients with Cowden syndrome (Nelen et al. 1997) and Bannayan-Riley-Ruvalcaba (BRRS) syndrome (OMIM 153480). Cowden syndrome is an autosomal dominant multiorgan hamartoma syndrome characterized by benign and malignant thyroid, breast, and colon cancer (OMIM 158350). BBRS is an autosomal dominant disorder characterized by macrocephaly, multiple lipomas, and angiomatosis (OMIM 153480). The germline mutations are scattered along the *PTEN* gene and include missense, nonsense, frame-shift, splice site mutations, and large deletions affecting one allele, resulting in inactivation of the protein (Bonneau and Longy 2000). Similar to somatic mutations, the majority of the germline missense mutations cluster in exon 5 of *PTEN* which encodes for the phosphatase core motif (Bonneau and Longy 2000).

PTEN is involved in cell cycle arrest and apoptosis (Li et al. 1998). Alterations in *PTEN* appear to be associated with late-stage disease in

several tumors and may occur through genetic inactivation or regulation of expression (El-Deiry 2003). Both *PTEN* and Phosphatidylinositol-3-kinase (*PI3K*) are important in suppressing apoptosis and in promoting the growth of cells (Weinberg 2007). PI3K phosphorylates phosphatidylinositol (PI) converting it to PIP_3. AKT binds to PIP_3, normal levels of which are regulated by PTEN. Subsequently, AKT becomes phosphorylated and activates a series of proteins that stimulate cell growth and inhibit apoptosis (Weinberg 2007). Therefore, dysregulation of the AKT pathway often happens by either activation of *PI3K* or inactivation of *PTEN* (Weinberg 2007). The missense, frame-shift and nonsense mutations of *PTEN* occur in the phosphatase/C2 domain required for its phosphatase enzymatic activity (Leslie and Downs 2004). These mutations inhibit the ability of PTEN to dephosphorylate PIP_3, thereby allowing uncontrolled downstream signaling from PI3K, resulting in increased cell proliferation and evasion of apoptosis (Weinberg 2007).

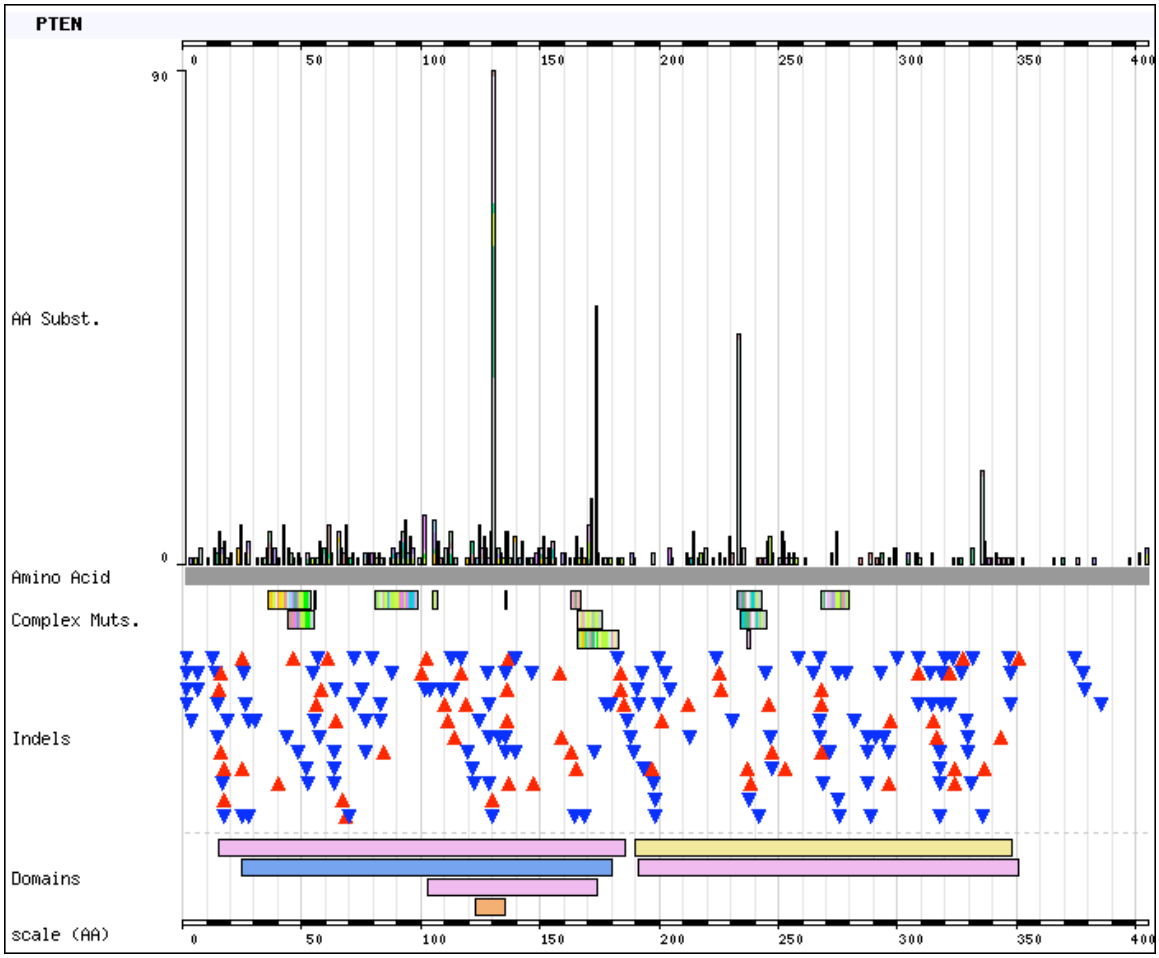

Figure 1-12: Distribution of somatic mutations in PTEN. A schematic of the PTEN protein displays missense amino acid substitutions and base substitutions resulting in nonsense mutations, and frame-shift mutations. The peaks of mutation occur in the catalytic domains of PTEN. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic

1.1.5.4.10 *ERBB2* : v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian)

ERBB2 is a dominantly acting cancer gene located on chromosome 17q21.1 and is composed of 26 exons. The predicted encoded protein has 1255 amino acids and is a receptor tyrosine kinase belonging to the epidermal growth factor receptor family. The *ERBB2* gene, originally called *NEU* was found to be repeatedly activated in neuroblastoma and glioblastoma cell lines derived from tumors of the BDIX strain of rat treated with ethylnitrosourea (Schubert et al. 1974). DNA isolated from four independently derived rat neuroblastoma and glioblastoma cell lines contained activated *ERBB2* detected by the transforming NIH-3T3 assay (Shih et al. 1981). *ERBB2* was first recognized as a distinct gene due to its association with the 185kD tumor antigen, p185, displayed on the surface of transfected cells (Padhy et al. 1982).

Complementary DNA (cDNA) clones were isolated from rat and human cancer cell lines transformed by *ERBB2* (Bargmann et al. 1986b, Yamamoto et al. 1986, Coussens et al. 1985). The DNA sequences of the rat and human clones were found to be colinear and 50% identical to the predicted amino acid sequence of the epidermal growth factor receptor (*EGFR)* gene (Bargmann et al. 1986b). The sequence similarity of *ERBB2* and *EGFR* implicated *ERBB2* as a transmembrane protein tyrosine kinase (Bargmann et al. 1986b). *ERBB2* was mapped on human chromosome 17q21 by in situ hybridization (Fukushige et al. 1986). Isolated genomic clones of normal and

transforming alleles of the rat *ERBB2* gene and their structural comparison revealed no evidence of gross rearrangements (Hung et al. 1986). Also, comparable levels of p185 expression were found in nontransformed cell lines with the normal allele and in transformed cell lines containing the mutant allele (Hung et al. 1986). These results pointed to the presence of subtle genetic alterations within the encoded protein p185 responsible for the activation of *ERBB2* (Bargmann et al. 1986a).

A comparison of an isolated transforming cDNA clone of the rat *ERBB2* gene to a cDNA clone of the normal allele revealed that a single base mutation appeared to confer potential for transformation of the *ERBB2* gene (Bargmann et al. 1986a). The mutation was observed in the transmembrane domain and caused a valine residue in the normal clone at codon 664 to be replaced by glutamic acid in the oncogenic clone (Bargmann et al. 1986a). The corresponding position in the human ERBB2 protein is codon 659 also encoding a valine residue. However, in order for the valine to be converted to a glutamic acid, a double point mutation would have to occur (Lemoine et al. 1990). The occurrence of such double point mutations are quite rare *in vivo* (Segatto et al. 1988). This prompted the screening of a panel of 100 breast cancers for such activating mutations at codon 659 of ERBB2 (Lemoine et al. 1990). No activating point mutations at codon 659 were identified in human breast cancers, even those cancers with amplified *ERBB2* (Lemoine et al. 1990). Therefore, over expression and amplification was not crucial for the activation of *ERBB2* gene in the chemically induced rat neuroblastomas

(Bargmann et al. 1986a). However, over expression of *ERBB2* may be essential for tumor formation in human cancers.

Over expression of normal *ERBB2* was found to be able to transform NIH-3T3 cells (Hudziak et al. 1987). Amplification and over expression of *ERBB2* has been observed in primary human tumors such as gastric (Fukushige et al. 1986), colon (D'Emilia et al. 1989), lung (Shi et al. 1992), salivary (Semba et al. 1985), ovarian (Hung et al. 1992), breast (Slamon et al. 1989, Lacroix et al. 1989), and bladder (Zhau et al. 1990). Amplification and overexpression of *ERBB2* occurs in approximately 30% of breast cancer (Slamon et al. 1989), and show clinical correlates with earlier relapse and shorter overall survival (Perren et al. 1991). The clinical correlate of *ERBB2* amplification and over expression and survival suggested the use of *ERBB2* levels as a prognostic marker in breast cancer (Perren et al. 1991). In some human breast primary tumors and cell lines, the over expression of *ERBB2* occurs in the absence of gene amplification (Berger et al. 1988, Slamon et al. 1989). This suggests that mechanisms other than amplification can lead to over expression of *ERBB2* in cancer. Indeed, post-transcriptional deregulation has been shown to contribute to over expression of *ERBB2* in breast cancer, in the absence of gene amplification (Child et al. 1999).

High throughput sequencing of the exons and exon-intron boundaries of *ERBB2* gene identified activating in-frame insertions in the kinase domain present in 10% of lung adenocarcinomas (Stephens et al. 2004), (Figure 1-13). Also, dominantly acting somatic point mutations in the kinase domain

have been identified in gastric, colorectal, and breast cancers (Lee et al. 2006a).

Trastuzumab, a humanized mouse monoclonal antibody was approved by the United States Food and Drug Administration (FDA) in 1998 for the treatment of amplified or over expressed ERBB2 protein in metastatic breast cancers (Roskoski 2004). *ERBB2* over expression in breast cancer cells, is identified by FISH allowing the selection of patients likely to respond to therapy (Roskoski 2004). This was the first genetic based drug or product approved for cancer therapy. The identification of in-frame insertion mutations of the *ERBB2* gene in lung cancers has implications for the use of Trastuzumab, or other anti-ERBB2 therapies as chemotherapeutics for a subset of lung cancers harboring *ERBB2* mutations (Futreal et al. 2005).

The mechanisms by which trastuzumab induces regression of *ERBB2* over expressing tumors are still being elucidated, however several molecular and cellular effects have been reported (Nahta et al. 2006). One of these, is the demonstration of trastuzumab's ability to reduce signaling mediated by ERBB2 through the PI3K and MAPK signaling cascades. Reduced downstream signaling through these pathways induces the cyclin-dependent kinase inhibitor p27kip1, which promotes cell-cycle arrest and apoptosis (Nahta and Esteva 2006). Trastuzumab has also been shown to rapidly dissociate Src kinase, a non-receptor tyrosine kinase, from ERBB2. This dissociation reduces Src activity leading to the dephosphorylation of PTEN and translocation to the plasma membrane, where PTEN becomes active

(Nagata et al. 2004). The activation of PTEN subsequently inhibits the activity of PI3K downstream effector AKT (Nagata et al. 2004).

The *EGF* receptor gene family has four members (Burgess et al. 2003). *ERBB2* is the second member of the *EGFR* family and has no known direct ligand (Roskoski 2004). Rather, ERBB2 forms heterodimers with other family members, and these heterodimers can bind growth factors (Graus-Porta et al. 1997). Another unique property of *ERBB2* is that its over expression can cause malignant transformation without the expression of a growth factor (Roskoski 2004). This suggests that *ERBB2* has a high level of ligand-independent (constitutive) activity, and its expression over a certain threshold can drive tumor growth (Yarden et al. 2001).

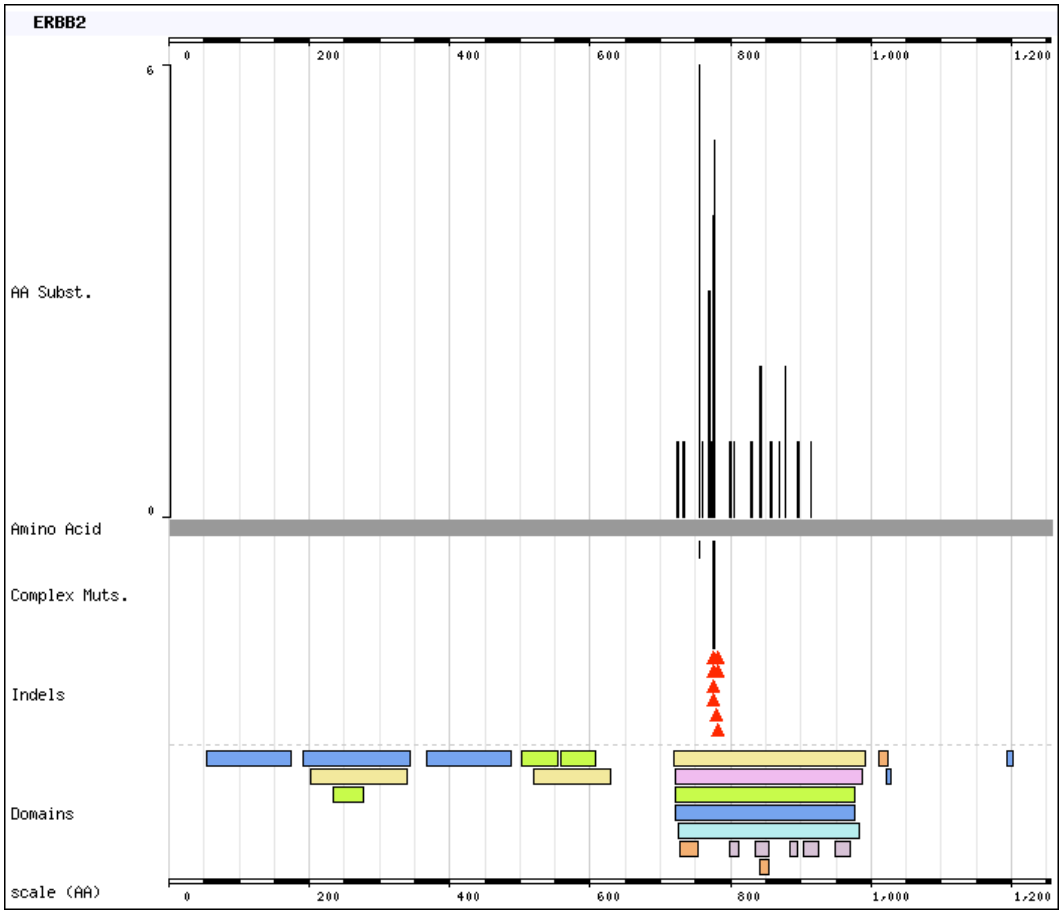

Figure 1-13: Distribution of somatic mutations in ERBB2. A schematic of the ERBB2 protein displays missense amino acid substitutions and insertion mutations in the kinase domain. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.11 *EGFR:* epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog, avian)

EGFR is a dominantly acting cancer gene located on chromosome 7p12.3-12.1 and is composed of 28 exons. The predicted encoded protein has 1210 amino acids and belongs to the epidermal growth factor receptor family of tyrosine kinases involved in signal transduction (Weinberg 2007). The human *EGFR* gene was identified by hybridization of sequences derived from the v-erb-B transforming protein of avian erythroblastosis virus (AEV) (Downward et al. 1984). *EGFR* was assigned to human chromosome 7 using human-mouse somatic cell hybrids and mapped by FISH to 7p12 (Davies et al. 1980). The AEV gene product in chickens had a truncated form of cellular EGFR and expressed a higher level of tyrosine kinase activity without its ligand (Kris et al. 1985).

Reports of *EGFR* gene amplification and over expression in human have been made in squamous cell carcinomas, glioblastomas, head and neck cancers, bladder, esophagus, breast, gastric, colon, endometrial, and lung cancers (Cowley et al. 1986, Libermann et al. 1984, Libermann et al. 1985, Henn et al. 1986, Nicholson et al. 2001b). By Southern blot analysis, the structural alterations of *EGFR* in two glioblastoma cell lines with amplified *EGFR*, were characterized by in-frame deletions of the extracellular domain (Yamazaki et al. 1988). In addition to amplifications there are coincident, in-frame deletions of the kinase domain found predominantly in lung adenocarcinomas (Figure 1-14). Point mutations in extracellular domain of EGFR are observed mainly in gliomas and in squamous cell lung carcinoma.

The over expression of *EGFR* in human cancer was associated with poor prognosis (Nicholson et al. 2001a, Sridhar et al. 2003). In those human cancer cell lines, EGFR also displayed elevated tyrosine kinase activity in the absence of its ligand (Kris et al. 1985). The fact that *EGFR* is over expressed in various human cancers, and its over expression is associated with poor prognosis made *EGFR* a good candidate for the development of targeted therapeutics (Sharma et al. 2007). Two EGFR-targeting-small –molecule inhibitors, gefitinib and erlotinib, received speedy approval from the US FDA in May 2003 and November 2004, respectively. Gefitinib and erlotinib were approved as treatment for patients with advanced chemotherapy-refractory non-small cell lung cancer (NSCLC) (Kris et al. 2003, Fukuoka et al. 2003). The mechanism of action of gefitinib and erlotinib is to compete with ATP binding at the catalytic site of the EGFR protein.

A multi-institutional phase II trial showed that 10% of patients with NSCLC responded dramatically to gefitinib and erlotinib, evidenced by shrinkage of the tumor mass (Cohen et al. 2004, Cohen et al. 2005). The response to these EGFR inhibitors was more evident in a subset of lung cancer cases characterized by non-smokers, women, East Asians, and patients with adenocarcinomas with bronchoalveolar histology (Pao et al. 2004a, Paez et al. 2004). Sequencing the exons and exon-intron boundaries of the *EGFR* gene showed that in most cases of marked response, patients harbored specific somatic mutations of the kinase domain (Paez et al. 2004, Lynch et al. 2004). These mutations were dominantly acting and

characterized mainly by heterozygous in-frame deletions of exon 19 and missense amino acid substitutions of exon 21 (Paez et al. 2004, Lynch et al. 2004).

Analysis by western blot showed that anti-apoptotic proteins, AKT and STAT, signaling were enhanced in cells with mutant EGFR. Upon inhibition of the EGFR protein signaling, downstream signaling of proteins involved in antiapoptosis, AKT and STAT, was markedly decreased (Sordella et al. 2004). Thus, inhibiting the signaling of the EGFR protein inhibits essential anti-apoptotic signals tranduced by the mutant receptor.

Additional studies have shown differences in the clinical outcomes associated with different mutations (Mitsudomi et al. 2005, Riely et al. 2006a). For example, NSCLCs that harbor exon 19 deletions of *EGFR* seem to respond better to gefitinib and erlotinib than tumors with missense mutations of exon 21, such as L858R (Riely et al. 2006a). Thus far, insertion mutations of exon 20 have not been shown to confer sensitivity to gefitinib and erlotinib *in vitro*, and have not been reported to occur in responsive cases (Sharma et al. 2007). Differential response of these EGFR mutants does not seem to correlate with the level of EGFR kinase activity (Greulich et al. 2005).

The predictive value of *EGFR* mutations and response to gefitinib and erlotinib may be limited to NSCLC. Recently EGFR inhibitors have demonstrated an inhibitory effect in hepatocellular and nasopharyngeal

carcinoma human cancer cell lines and animal models (Lee et al. 2006b). However, sequencing of the kinase domain, exons 18-21, of *EGFR* in a panel of 100 hepatocellular carcinomas and 100 nasopharyngeal carcinomas did not reveal the presence of somatic activating mutations in the kinase domain (Lee et al. 2006b). This study may point to alternative mechanisms important for the observed activity of small molecule EGFR inhibitors in these cancer types (Lee et al. 2006b).

However, resistance to gefitinib and erlotinib in NSCLC is linked to a specific secondary somatic EGFR mutation T790M (Kobayashi et al. 2005a, Pao et al. 2005a). This mutation is similar to the BCR-ABL tyrosine kinase T315I mutation acquired due to resistance to its selective inhibitor, imatinib (Kobayashi et al. 2005a). The T790M mutation of EGFR is postulated to weaken the interaction of the inhibitor with its target (Sharma et al. 2007). Gefitinib and erlotinib are reversible inhibitors of EGFR, and the use of irreversible EGFR inhibitor compounds *in vitro* has been shown to overcome resistance (Sharma et al. 2007). Interestingly, the T790M mutation has also been identified in the germline of familial cases of the broncho-alveolar carcinoma subtype of non-small cell lung cancer, implicating mutant *EGFR* in genetic susceptibility to lung cancer (Bell et al. 2005a).

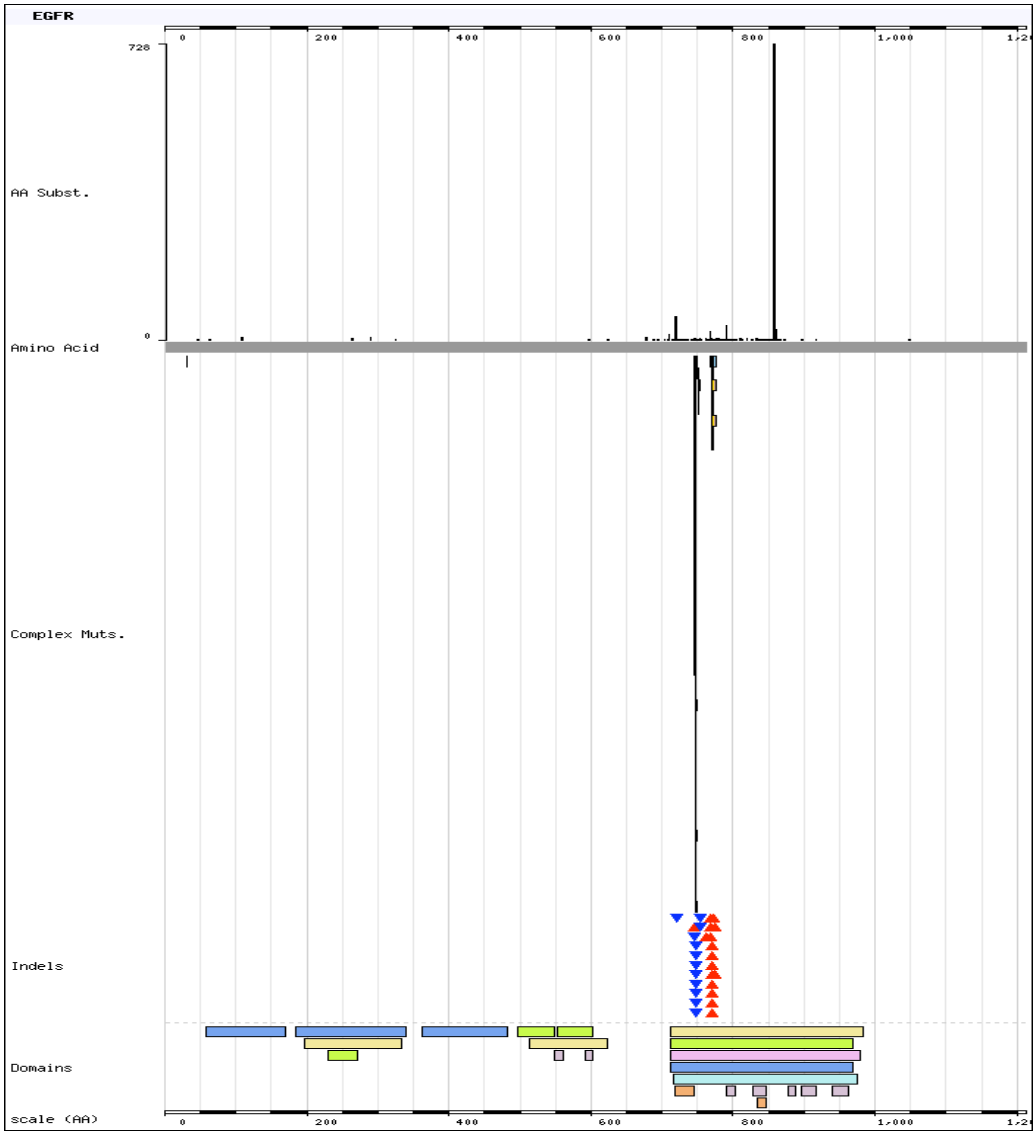

Figure 1-14: Distribution of somatic mutations in EGFR. A schematic of the EGFR protein displays amino acid substitutions, in-frame deletions and insertions occurring mainly in the kinase domain. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic

1.1.5.4.12 *KIT:* v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog

KIT is a dominantly acting cancer gene located on chromosome 4q11-q12 and is composed of 21 exons. Its predicted encoded protein has 976 amino acids and is a transmembrane tyrosine kinase required for normal hematopoiesis and is structurally homologous to *PDGFRA* (OMIM 164920). The provirus, Hardy-Zuckerman 4 feline sarcoma virus, was molecularly cloned and a segment showed homology to mammalian genomic DNA and termed v-Kit (OMIM 164920). Its human homolog, *KIT*, was assigned to chromosome 4 using human-mouse somatic cell hybrids (Barker et al. 1985). *KIT* was mapped by FISH to 4q11-12 (Mattei et al. 1987).

Mutations of the *KIT* gene in mouse are associated with white spots (W mutant), sterility and anemia (Geissler et al. 1988). The anemia was attributable to the failure of stem cell populations to migrate and/or proliferate effectively during development (Geissler et al. 1988). Deletions of *KIT* and *PDGFRA*, which reside on the same locus, were found in a patient with piebaldism (Spritz et al. 1992). The patient was hemizygous for the *KIT* and *PDGFRA* deletions (Spritz et al. 1992). Piebaldism is an autosomal dominant disorder characterized by mental retardation and integumentary pigment changes (OMIM 172800). In mice, aganglionic megacolon is associated with the piebald trait (Bielschowsky and Schofield 1962). Hirschprung disease has been observed in some human cases of piebaldism (Mahakrishnan and Srinivasan 1980).

Mice with *KIT* mutations were shown to lack a network of interstitial cells of Cajal (ICC) (Huizinga et al. 1995). The ICC is associated with the external muscle layer of the gut and is proposed to stimulate phasic contractions. Using antibodies against the intracellular domain of KIT in mice showed that wild type and heterozygous mice had high levels of KIT expression in the muscle layers of the gut (Huizinga et al. 1995). However, mice homozygous for the *KIT* mutation had little to no KIT expression in the muscle layers of the gut. This finding prompted the hypothesis that the gut abnormalities and megacolon observed in patients with piebaldism may reflect an identical function of the KIT signaling pathway.

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal neoplasms in the human digestive tract (OMIM 606764). GISTs were considered to be leiomyosarcomas because they resembled smooth muscle histologically (OMIM 606764). However, GISTs are gastrointestinal sarcomas without muscle or Schwann-cell markers, and have a lower rate of response to doxorubicin-based regimens compared to other leiomyosarcomas (OMIM 606764). Almost all GISTs express *KIT* and CD34, both of which are expressed on hematopoietic progenitor cells. Approximately 70% of GISTs develop in the stomach, 20% in the small intestine, and 10% in the esophagus, colon, and rectum (OMIM 606764). GISTs are now thought to be derived from ICC or their precursors (Joensuu 2006a).

The mutation status of *KIT* was investigated in panel of mesenchymal tumors and somatic mutations of exon 11 were identified in 60% of GISTs (Hirota et al. 1998). *KIT* is somatically mutated most frequently in gastrointestinal sarcomas (37%) and myeloid leukemias (28%) (COSMIC). The somatic mutations are gain-of-function mutations and are dominantly acting. The gain-of-function mutations are characterized by in-frame deletions and missense amino acid substitutions (Hirota et al. 1998). The *KIT* somatic mutations are located in the region between the transmembrane and kinase domains (Figure 1-15). All of the mutant KIT proteins are constitutively activated without the KIT ligand, stem cell factor (SCF) (Hirota et al. 1998). Mutant *KIT* cDNA was shown to transform NIH-3T3 cells, suggesting that mutations contribute to tumor development (Hirota et al. 1998).

In familial cases of GIST, germline deletion mutations in the kinase domain of the *KIT* and *PDGFRA* gene have also been identified (Nishida et al. 1998). The clinical manifestation of GIST varies in patients with germline *KIT* mutations. For example, germline mutations in the juxtamembrane domain have been associated with mastocytosis, hyperpigmentation, in addition to generalized hyperplasia of the GIST progenitor ICC (Carballo et al. 2005). In contrast, mastocytosis and hyperpigmentation appear absent in the case of familial GISTs with germline KIT mutations in the kinase domain (Li et al. 2005). Familial and sporadic GIST appear biologically similar (Fletcher and Rubin 2007).

Mutations in *KIT* result in constitutive tyrosine kinase activity, uncontrolled cell proliferation, and stimulation of downstream signaling pathways (Weinberg 2007). Imatinib, an inhibitor of the tyrosine kinase activity of BCR-ABL in chronic myeloid leukemia, was demonstrated to be effective in treating GISTs (Joensuu et al. 2001). In 2002, the US FDA approved imatinib for the treatment of advanced metastatic GISTs (Savage and Antman 2002). There is a differential response to imatinib of patients with GIST that is based on the location of *KIT* mutations. Almost all the juxtamembrane KIT mutations in GISTs are highly sensitive to imatinib, and patients have better than 80% clinical response (Heinrich et al. 2003). However, secondary and some intrinsic KIT mutations in the kinase domain are responsible for resistance to imatinib (Heinrich et al. 2006).

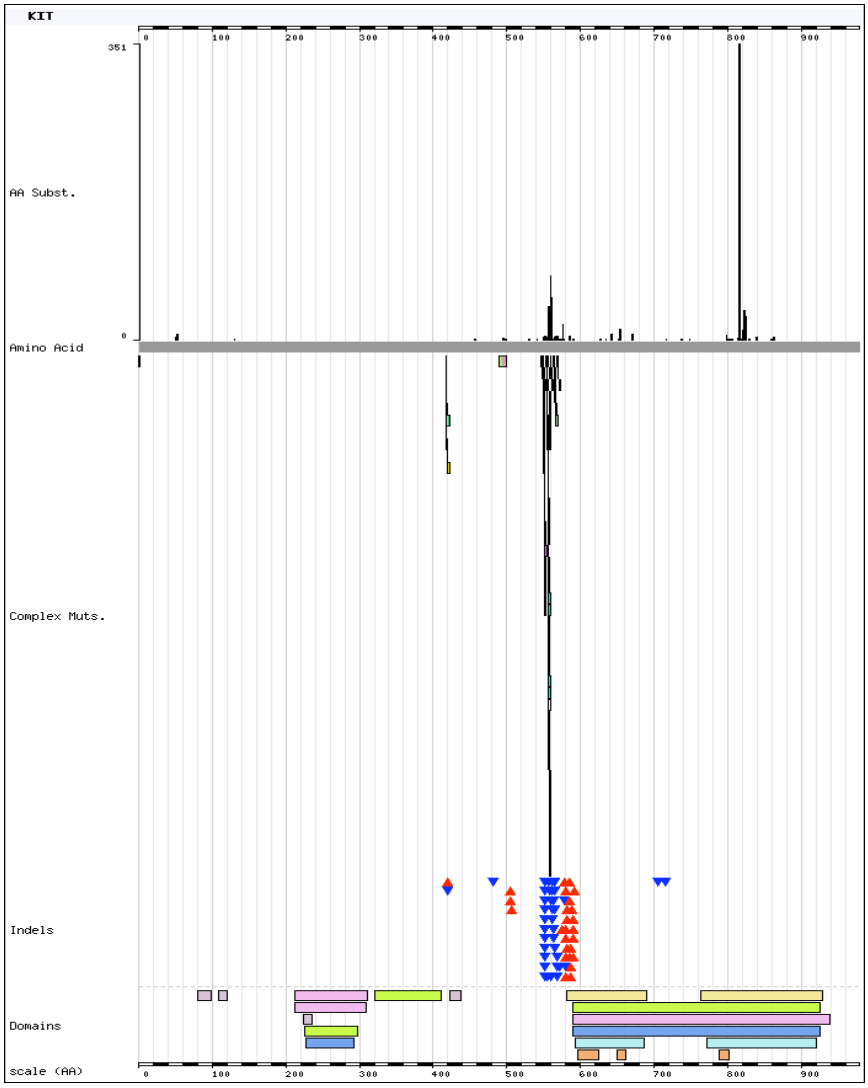

Figure 1-15: Distribution of somatic mutations in KIT. A schematic of the KIT protein displays missense amino acid substitutions, in-frame insertions and deletions. These somatic mutations occur mainly in the juxtamembrane and kinase domains. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic

1.1.5.4.13 *PDGFRA:* platelet-derived growth factor receptor, alpha polypeptide

PDGFRA is a dominantly acting cancer gene located on chromosome 4q12-13 and is composed of 23 exons. Its predicted encoded protein has 1089 amino acids and belongs to the platelet-derived growth factor receptor family. *PDGFRA* was identified when a 6.4-kilobase cDNA was found to coexpress with the already known platelet-derived growth factor receptor, beta (*PDGFRB)* in normal human tissues (Matsui et al. 1989). When transfected into COS-1 cells, the protein expression of the 6.4-kilobase transcript was detected by antiserum raised against the predicted peptide (Matsui et al. 1989). Also, a characteristic pattern of binding of the platelet-derived growth factors (*PDGF*) isoforms was observed and was different from that of *PDGFRB* (Matsui et al. 1989). This new *PDGFR* was called *PDGFRA*. *PDGFRA,* was localized to 4q, the same region as *KIT* by FISH (Matsui et al. 1989).

The 4q region of the human chromosome is homologous with mouse chromosome 5. It was hypothesized that the mouse *PDGFRA* gene would also map to the same region as the *KIT* gene (Stephenson et al. 1991). If so, *PDGFRA* in the mouse would be related to other mutant loci in the region that affect mammalian development (Stephenson et al. 1991). It was shown that *PDGFRA* in the mouse is closely linked to *KIT* (W locus), and analysis of interspecific F_1 hybrids show that the heterozygote patch mutants carry a deletion in the genomic sequence associated with *PDGFRA* (Stephenson et al. 1991). Intercrossing of heterozygote patch mutants yielded 25% offspring

homozygous patch mutants. These homozygous patch mutants lacked the *PDGFRA* genomic sequences and undetected mRNA expression (Stephenson et al. 1991). The analysis of *KIT* gene in these patch homozygotes revealed that the deletion did not affect KIT coding sequence (Stephenson et al. 1991).

Since *KIT* and *PDGFRA* share the same locus, and mouse mutants of these genes shared similar phenotypes, it was postulated that dysregulated forms of *KIT* and *PDGFRA* would confer similar phenotypes in human disease. Most GISTs have activating mutations in the *KIT* receptor tyrosine kinase (Heinrich et al. 2003). Most patients with GISTs respond to imatinib, a KIT tyrosine kinase inhibitor. A mutation screen of the *PDGFRA* gene in a panel of GISTs revealed that 35% of GISTs lacking *KIT* mutations had dominantly acting, intragenic activating mutations in *PDGFRA* (Heinrich et al. 2003, Hirota et al. 2003). The tumors expressing either *KIT* or *PDGFRA* were indistinguishable with respect to the activation of downstream signaling molecules and cytogenetic changes associated with tumor progression (Heinrich et al. 2003). It was therefore concluded that *KIT* and *PDGFRA* mutations appeared to be alternative mutually exclusive oncogenic mechanisms in GISTs (Heinrich et al. 2003).

Although there are similarities between *PDGFRA* mutant and *KIT* mutant GISTs at the molecular level, there are a number of clinicopathologic differences between these tumors (Corless et al. 2005). Compared to *KIT* mutant GIST, *PDGFRA* mutant GISTs are weak or negative for KIT protein

expression as measured by immunohistochemistry (Debiec-Rychter et al. 2004), and arise almost exclusively in the stomach. *KIT* mutant GISTs however, occur at a variety of sites along the GI tract (Wardelmann et al. 2004). Also, the gene expression profiles of *PDGFRA* mutant tumors clusters separately from *KIT* mutant tumors (Subramanian et al. 2004). Therefore, *PDGFRA* mutant GISTs are distinct from *KIT* mutant GISTs.

The somatic activating mutations of *PDGFRA* render the protein constitutively active. This constitutive phosphorylation by PDGFRA activates the MAPK/ERK and STAT3 signaling pathways. The mutations are dominantly acting and are characterized by missense amino acid substitutions, in-frame insertion and deletion/substitution mutations (Figure 1-16). The majority of mutations cluster in the activation loop of PDGFRA. The most common PDGFRA mutant in GISTs, D842V, occurs in more than 60% of cases and shows significant resistance to imatinib *in vitro* (Corless et al. 2005). The D842V mutation of PDGFRA is homologous to the D816V mutation of KIT, which is known as being resistant to imatinib *in vitro* (Zermati et al. 2003). Therefore, the therapeutic benefit of imatinib may be limited in the majority of *PDGRA* mutant GISTs, but may be of use to the minority without the resistance mutation.

PDGFRA is also implicated in disease by formation of fusion genes. Chronic myeloid leukemia (CML) is characterized by the presence of the *BCR-ABL* fusion gene, resulting from a chromosomal translocation involving the q arms of chromosome 9 and 22. This resulting chromosome is called the

Philadelphia (Ph) chromosome. In some patients with CML a fusion gene involving *BCR* and *PDGFRA* (*BCR-PDGFRA*) has been identified (Baxter et al. 2002). Also, idiopathic hypereosinophilic syndrome is often caused by an interstitial deletion on chromosome 4q12 resulting in fusion of *PDGFRA* and *FIP1L1*, a neighboring gene (Cools et al. 2003). Hypereosinophilic syndrome is a rare hematologic disorder characterized by sustained over production of eosinophils in the bone marrow, eosinophilia, tissue infiltration, and organ damage (OMIM 607685). The *FIP1L1-PDGFRA* fusion gene was found to be a constitutively activated tyrosine kinase able to transform hematopoietic cells. The activity of this fusion gene product is inhibited by imatinib. Resistance to imatinib in acquired through missense amino acid substitutions in the kinase domain of *PDGFRA* (Cools et al. 2003).

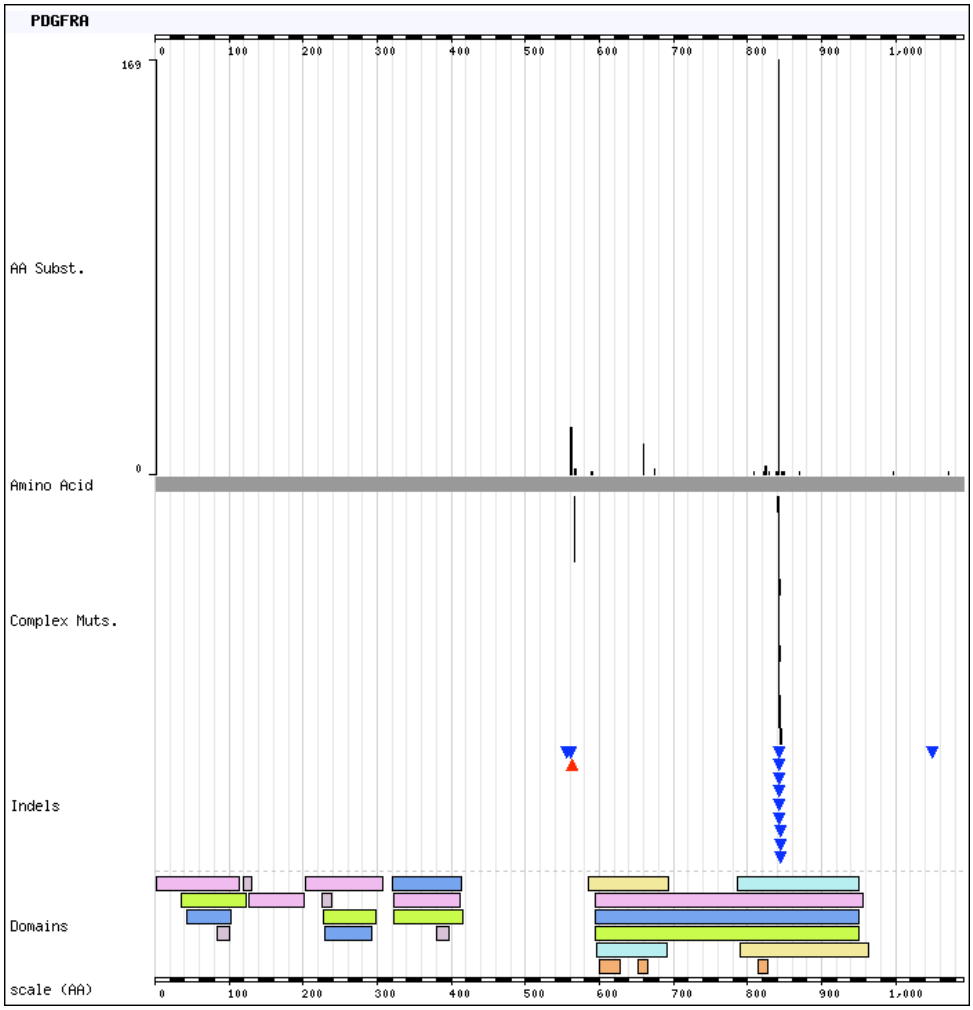

Figure 1-16: Distribution of somatic mutations in PDGFRA. A schematic of the PDGFRA protein displays missense amino acid substitutions, in-frame insertion and deletion mutations. Most of these mutations cluster in the kinase domain. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.14 *RET:* ret proto-oncogene

RET is a dominantly acting cancer gene located on chromosome 10q11.2 and is composed of 20 exons. Its predicted encoded protein has 1114 amino acids and is a receptor tyrosine kinase cell surface molecule that

transduces signals for cell growth and differentiation (OMIM 164761). *RET* was identified as a transforming gene by transfection of NIH-3T3 cells with DNA from a human lymphoma (Takahashi et al. 1985). The transforming ability of *RET* appeared to be due to cytogenetic rearrangement (Takahashi et al. 1985, Grieco et al. 1990). *RET* was assigned to chromosome 10q11.2 by FISH (Ishizaka et al. 1989).

Germline mutations of *RET* were identified in patients with multiple endocrine neoplasia, type IIA (MEN2A) (Mulligan et al. 1993), and type IIB (MEN2B) and medullary thyroid carcinoma (MTC) (Donis-Keller et al. 1993). Multiple endocrine neoplasia type 2 (MEN2) is an autosomal dominant syndrome that occurs as one of two entities: MEN2A or MEN2B. The disorder is characterized by pheochromocytoma and medullary thyroid carcinoma (OMIM 171400). A syndrome related to MEN2 is familial MTC (FMTC). FMTC differs from MEN2 in that the manifestation of the disease is of thyroid carcinoma only without other primary tumors (OMIM 155240).

Genetic linkage analysis of families affected by Hirschsprung disease (HSCR) revealed that a locus on chromosome 10 was tightly linked to the disease. Hirschsprung disease is an autosomal dominant congenital disorder characterized by absence of enteric ganglia along a variable length of the intestine (OMIM 142623). There are two forms of HSCR: short segment HSCR and long segment HSCR. A mutation screen of the locus identified mutations of *RET* to account for 50% of familial cases of HSCR. Most of these mutations in *RET* were associated with the more severe long segment

HSCR (Romeo et al. 1994). These mutations were missense amino acid substitutions in the kinase domain of *RET* (Romeo et al. 1994, Edery et al. 1994).

These missense mutations of *RET* in HSCR represent loss-of-function mutations that impair the kinase activity (Pasini et al. 1995). The effects of the *RET* mutations were assessed using cDNA constructs of different HSCR *RET* mutations of the extracellular domain (Iwashita et al. 1996). The mutations of the extracellular domain were found to inhibit the transport of the RET protein to the plasma membrane (Iwashita et al. 1996). Therefore, it was shown that cell surface expression of RET is required for ganglion migration and differentiation.

Unlike Hirschsprung disease, missense *RET* mutations in MEN2A, MEN2B, and FMTC are activating and render the RET protein constitutively activated. In the majority of MEN2A cases, the *RET* mutation consists of loss of one of the cysteine residues in the extracellular domain (Santoro et al. 1995). However, in the majority of MEN2B cases, *RET* is mutated by the substitution of methionine at codon 918 with a threonine (M918T) (Santoro et al. 1995). FMTC can be caused by point mutations of the *RET* gene corresponding to the extracellular or intracellular domain (Ponder et al. 1999), (Figure 1-17). The missense M918T *RET* mutation, interestingly, also occurs in sporadic cases of MTC (Hofstra et al. 1994).

Papillary thyroid carcinoma is the most common thyroid malignancy in humans. Somatic rearrangement of the *RET* gene generating the chimeric *RET*-papillary thyroid carcinoma (RET-*PTC*) oncogene is one of the molecular lesions associated with papillary carcinoma (Salvatore et al. 2000). It was demonstrated that *RET-PTC* oncogene leads to ligand-independent tyrosine kinase activity and this constitutive signaling is required to maintain oncogenic cell proliferation (Salvatore et al. 2000). The oncogenic effects of *RET-PTC* require signaling along the MAPK/ERK pathway. Alterations of this pathway in thyroid cancer can occur at different levels as a result of somatic missense amino acid substitutions or rearrangements involving *RET*, *RAS*, and *BRAF* (Ciampi and Nikiforov 2007).

RET is a cell membrane receptor tyrosine kinase. The ligands of the RET receptor are growth factors belonging to the glial cell line derived neurotrophic factor family. Binding of the ligand to RET causes receptor dimerization, autophosphorylation of tyrosine residues in the intracellular domain, and activation of the signaling cascade (Ciampi and Nikiforov 2007). Several tyrosine kinase inhibitors have been shown to inhibit the oncogenic RET-PTC signaling (Ciampi and Nikiforov 2007). Inhibitors of different kinases along the MAPK pathway have shown therapeutic effect in thyroid cells *in vitro*, and are currently being tested in clinical trials (Ciampi and Nikiforov 2007).

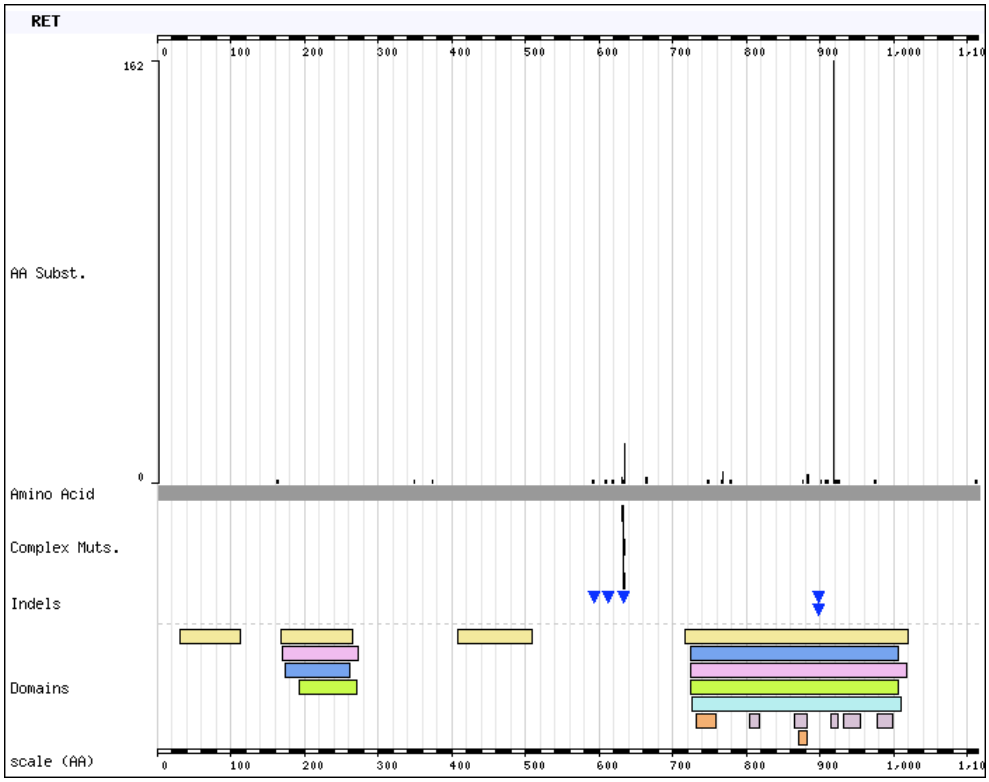

Figure 1-17: Distribution of somatic mutations in RET. A schematic of the RET protein displays missense amino acid substitutions peaks in the kinase domain and in-frame deletions. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.15 *MET* : met proto-oncogene (hepatocyte growth factor receptor)

MET is a dominantly acting cancer gene located on chromosome 7q31 and is composed of 21 exons. Its predicted encoded protein has 1390 amino acids and is a receptor tyrosine kinase for hepatocyte growth factor. *MET*, a transforming gene in NIH-3T3 cells, was cloned from a chemically transformed human osteosarcoma-derived cell line. By direct hybridization

MET was shown to be distinct from all other previously known oncogenes (Cooper et al. 1984), and was mapped to chromosome 7q (Cooper et al. 1984).

The transforming activity of *MET* was due to a DNA rearrangement where sequences from the translocated promoter region (TPR) locus on chromosome 1 were fused to sequences from the *MET* locus on chromosome 7 (TPR-MET). This rearrangement was later identified in some patients with gastric carcinoma (Yu et al. 2000). *MET* was sequenced and shown to be related to the protein tyrosine kinase growth factor receptors, particularly the insulin receptor and *ABL* (Dean et al. 1985). Further analysis of the *MET* sequence proved it was a cell-surface receptor (Park et al. 1987) and its ligand was identified as the hepatocyte growth factor (Bottaro et al. 1991).

Hereditary papillary renal carcinoma is an autosomal dominant disorder characterized by the development of multiple, bilateral papillary renal tumors (OMIM 605074). By genetic linkage analysis of affected families, an interval on chromosme 7q containing the *MET* gene was mapped (Schmidt et al. 1997). Germline amino acid substitution mutations in the conserved kinase domain of MET were identified in patients with hereditary papillary renal carcinoma. Activating somatic mutations of *MET* were also identified in a subset of sporadic cases of papillary renal carcinomas (Schmidt et al. 1997). These dominantly acting somatic mutations confer constitutive activation of the MET protein leading to increased cellular motility in papillary renal carcinomas (Jeffers et al. 1997), (Figure 1-18). Somatic mutations of *MET*

occur in head and neck cancers (15%), liver cancers (6%), small cell and non-small cell lung cancers (5%), and renal cancers (4%) (COSMIC).

MET is also found to be frequently amplified in 95% of sporadic papillary renal carcinoma, as identified by FISH and comparative genomic hybridization (CGH) (Zhuang et al. 1998). This amplification event was shown to be non-random and affected the mutant *MET* allele more than the wild type *MET* allele, thus implicating the event in tumorigenesis (Zhuang et al. 1998). *MET* is also found over expressed in human cancers and is amplified during the transition from primary tumor to metastatis (Giordano et al. 1997). The over expression of *MET* has been shown to correlate with poor prognosis (Birchmeier et al. 2003). Overexpression and amplification occurs in bladder cancers (Cheng et al. 2002), nasopharyngeal cancers (Qian et al. 2002), oral squamous cell cancers (Morello et al. 2001) and gastric cancers (Park et al. 2000).

Activation of MET signaling results from the binding of hepatocyte growth factor (HGF) to the receptor. Activated MET recruits scaffolding proteins such as growth-factor-receptor-bound protein 2(GRB2)-associated binder 1 (GAB1), which activate SH2-domain containing protein tyrosine phosphatase 2 (SHP2), RAS, and MAPK/ERK (Birchmeier et al. 2003). This causes changes in the gene expression of cell-cycle regulators, extracellular-matrix proteinases, and cytoskeletal functions that control migration, invasion and proliferation (Birchmeier et al. 2003).

Dysregulation of MET and its downstream signaling pathways is a feature of many human cancers. The MAPK pathway and the PI3K pathway are downstream effectors of MET signaling (Birchmeier et al. 2003). It has been demonstrated that over expression of *MET* or its ligand, *HGF*, leads to tumorigenesis and metastatis in athymic nude mice (Rong et al. 1994). Downregulation of *MET* or *HGF* expression in human cancer cell lines decreases their tumorigenic potential (Abounader et al. 2002).

Therefore, therapeutic targeting of MET may be beneficial in the treatment of human cancers harboring mutant forms of *MET*. Some highly selective synthetic inhibitors of MET have been developed and tested in various model systems (Sattler and Salgia, 2007). Amplification of *MET* (Smolen et al. 2006), activating mutation and overexpression of *MET* (Ma et al. 2005) seem to identify a subset of patients sensitive to selective tyrosine kinase inhibitors.

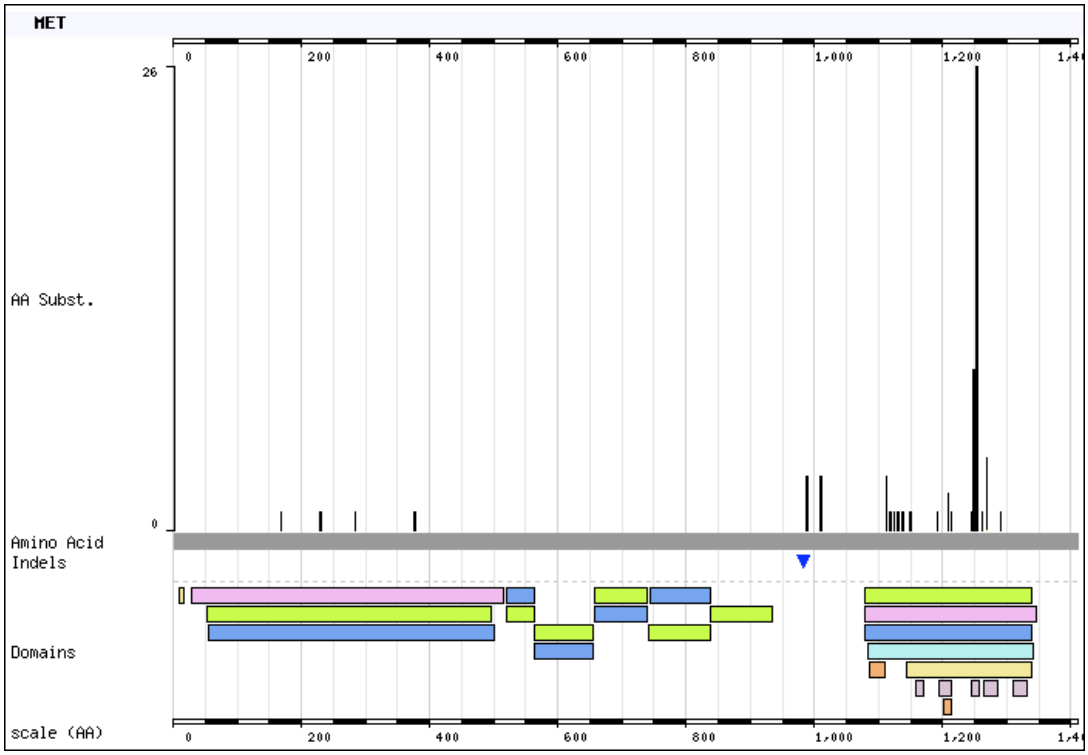

Figure 1-18: Distribution of somatic mutations in MET. A schematic of the MET protein displays mainly amino acid substitutions clustered in the kinase domain. Infrequently, in-frame deletions do occur in MET. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.16 *FLT3* : fms-related tyrosine kinase 3

FLT3 is a dominantly acting cancer gene located on chromosome 13q12 and is composed of 24 exons. Its predicted encoded protein has 993 amino acids and is a member of the growth factor receptor tyrosine kinase family that includes *KIT* and *PDGFRA*. Murine *FLT3* was cloned first from a stem cell fraction of fetal liver (Rosnet et al. 1991a). The human *FLT3* was cloned from a CD34+ hematpoietic stem cell-enriched library (Small et al.

1994). The FLT3 encoded protein sequence was similar to that of KIT and PDGFRA. *FLT3* was observed as highly expressed and restricted to CD34+ cells of human blood and marrow (Small et al. 1994). CD34+ cells are enriched for stem/ progenitor cells. Antibodies raised against FLT3 inhibited hematopoietic colony formation. These data suggested that *FLT3* may function as a growth factor receptor on hematopoietic stem and/or progenitor cells (Small et al. 1994). Using FISH, *FLT3* was mapped to human chromosome 13q12 and mouse chromosome 5 (Rosnet et al. 1991b). In mouse *KIT*, *PDGFRA* and *FLT3* are in close physical linkage (Rosnet et al. 1993).

Mouse knockouts of *FLT3* are born healthy and have normal peripheral blood counts. However, these mice have reduced numbers of early B-cell precursors in bone marrow (Mackarehtschian et al. 1995). Similarly, mice lacking the FLT3 ligand (FL) have significantly reduced leukocytes, myeloid and lymphoid progenitors of the bone marrow (McKenna et al. 2000). However, the phenotype is alleviated by the administration of exogenous FL. As well, the over expression of FL in transgenic mice leads to the development of leukemia (Hawley et al. 1998).

In human leukemia cells, the expression of *FLT3* is observed in most patients with acute myeloid leukemia (AML) and precursor-B acute lymphoblastic leukemia (ALL) (Birg et al. 1992, Carow et al. 1996). *FLT3* is not expressed in chronic myeloid leukemia (CML) patients in the chronic phase, but is present during disease transformation (Naoe and Kiyoi 2004).

High levels of *FLT3* expression in AML are associated with the auto-activation of FLT3 and poor prognosis (Ozeki et al. 2004).

Somatic mutations of *FLT3* were identified in the investigation into the elevated expression of *FLT3* in some patients with AML (Nakao et al. 1996). The *FLT3* transcript contained an elongated juxtamembrane (JM) region detected in PCR products of the transcript in some patients with AML (Nakao et al. 1996). *FLT3* was sequenced from samples of AML and somatic mutations were identified. The somatic mutations were dominantly acting and characterized by in-frame internal tandem duplication of sequences within the juxtamembrane (JM) domain, sometimes accompanied by insertion of additional nucleotides (Naoe and Kiyoi 2004). The presence of the in-frame tandem duplication mutations defined a high risk group of AML patients (Abu-Duhier et al. 2000, Kottaridis et al. 2001). Internal tandem duplications (ITD) of *FLT3* occur in approximately 21% of AML (COSMIC). These mutations confer growth factor- independence of FLT3 and induce constitutive activation of downstream signaling molecules such as SHP2, MAPK, and AKT (Kiyoi et al. 2002).

Activating amino acid substitution mutations of the kinase domain of FLT3 are also found in AML (Yamamoto et al. 2001). The most prevalent codon affected in *FLT3* is codon 835 (Figure 1-19), analogous to the *KIT* mutation (Yamamoto et al. 2001). The kinase domain mutations and those affecting the juxtamembrane domain are mutually exclusive. Unlike the ITD of the JM, the kinase domain substitution mutations are also found in ALL

(Yamamoto et al. 2001). The kinase domain amino acid substitution mutations also cause constitutive tyrosine phosphorylation of FLT3.

FLT3 mutations are the most frequent genetic alteration in AML and are associated with poor prognosis. Therefore inhibition of FLT3 may be of therapeutic benefit in that subset of patients. A number of tyrosine kinase inhibitors have shown potency in inhibiting the kinase activity of FLT3 and some are in clinical trials (Kiyoi and Naoe 2006).

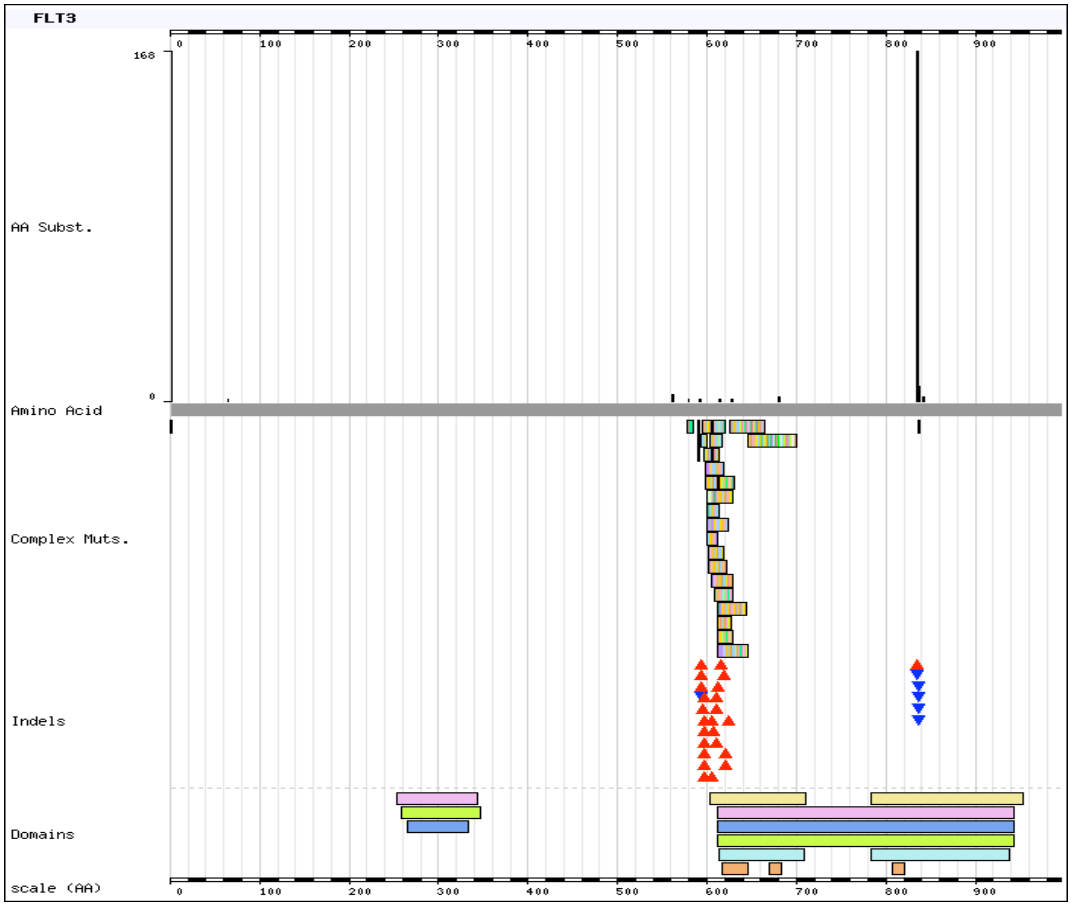

Figure 1-19: Distribution of somatic mutations in FLT3. A schematic of the FLT3 protein displays missense amino acid substitutions, in-frame insertion and deletion mutations. These mutations cluster in the juxtamembrane and

kinase domains. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.17 *MAP2K4*: mitogen-activated protein kinase kinase 4

MAP2K4 is a tumor suppressor gene located on chromosome 17p11.2 and composed of 11 exons. Its predicted encoded protein has 399 amino acids and is a mitogen-activated protein kinase. *MAP2K4* was first cloned in mouse (Sanchez et al. 1994). The human *MAP2K4* was later cloned (Derijard et al. 1995).

MAP2K4 is expressed widely in murine and human tissues and is more abundant in brain and skeletal muscle (Sanchez et al. 1994, Derijard et al. 1995). As studied by *in-situ* hybridization, mouse *MAP2K4* expression is detected preferentially in the central nervous system, liver, and thymus during early stages of mouse development (Carboni et al. 1997). The expression of *MAP2K4* in the liver and thymus gradually decreases during embryogenesis. In mice, *MAP2K4* is essential for viability and embryonic development, because mice null for *MAP2K4* die before embryonic day 14 (Nishina et al. 1997). Abnormal hepatogenesis and liver formation was also observed in *MAP2K4* null mouse embryos. This phenotype is similar to that reported for c-Jun NH2-terminal kinases (*JNK*) null mouse embryos, suggesting a role for both *MAP2K4* and *JNK* in liver formation during organogenensis (Nishina et al. 1997).

Mitogen activated protein kinases (MAPK) function in signal transduction pathways involved in controlling cellular processes in many organisms (Kim et al. 2002). MAPK are phosphorylated and activated in response to various extracellular stimuli (White et al. 1996). Mouse embryos null for *MAP2K4* showed defects in the downstream phosphorylation of stress-activated protein kinase (SAPK) (Ganiatsas et al. 1998). This finding implicated MAP2K4 as a signal transducer in the SAPK pathway, and a direct upstream activator of JNK. The phosphorylation and activation of JNK by MAP2K4 leads to the phosphorylation and activation of transcription factors such as activating transcription factor 2 (ATF2) (Cuenda et al. 1998). The SAPK pathway is activated in response to cellular stress such as irradiation, heat shock, DNA damage, and inflammatory cytokines. Activation of the SAPK pathway leads to apoptosis.

By positional cloning of a locus on chromosome 17 frequently affected by LOH in cancer, homozygous deletions of the *MAP2K4* gene were identified in one pancreatic and one lung cancer cell line (Teng et al. 1997). Of a panel of 213 human cancer cell lines screened for mutations in *MAP2K4* by sequencing, 7 harbored alterations in *MAP2K4* (Teng et al. 1997). These somatic alterations included homozygous deletions, nonsense and missense amino acid substitutions (Figure 1-20). *MAP2K4* is somatically mutated in approximately 2% of all cancers (COSMIC). Another study corroborated the identification of homozygous deletions of *MAP2K4* in pancreatic cancer, biliary adenocarcinomas, and breast cancer (Su et al. 1998). Missense mutations of the phosphorylation sites within the kinase domain of MAP2K4

have been shown to abolish its kinase activity (Yan et al. 1994). These findings implicate *MAP2K4* as a candidate tumor suppressor gene in human cancer (Su et al. 1998, Su et al. 2002).

Contradictory reports have been made of the function of *MAP2K4* in gastric cancer. MAP2K4 protein expression in gastric cancers has been shown to be associated with shorter relapse free survival compared to gastric cancers without MAP2K4 expression, implicating it as an oncogene (Wu et al. 2000). However, recently, it has been reported that lack of MAP2K4 expression in gastric cancers after surgical resection is predictive of poor survival, implicating it as a tumor suppressor gene (Cunningham et al. 2006).

Recent evidence points toward *MAP2K4* being a tumor suppressor gene. Sequence analysis of *MAP2K4* in several tumor types revealed the presence of truncating and missense mutations (Greenman et al. 2007).

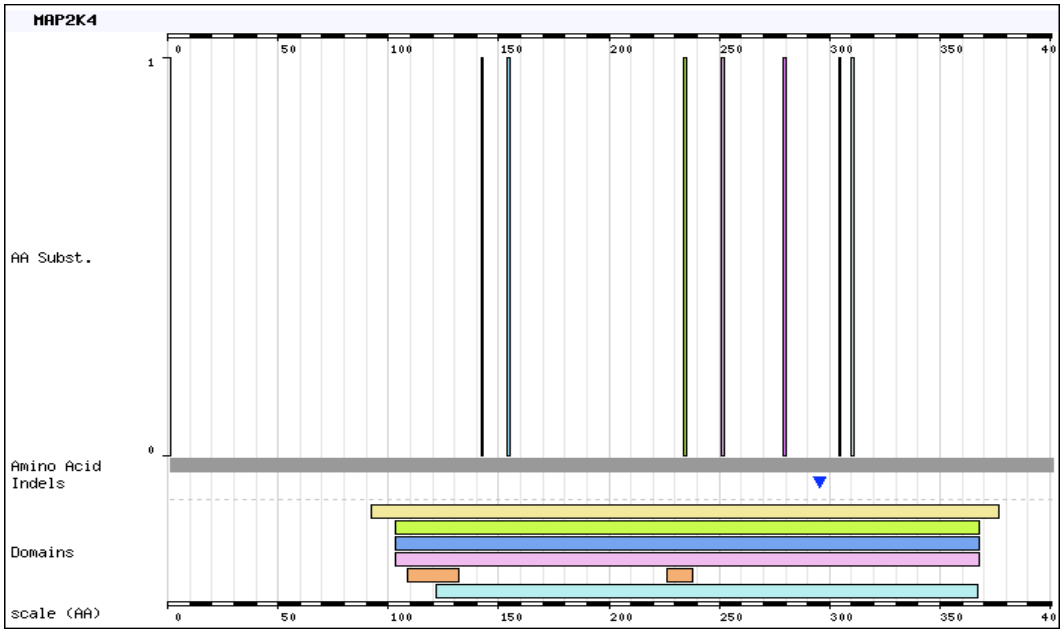

Figure 1-20: Distribution of somatic mutations in MAP2K4. A schematic of the MAP2K4 protein displays predominantly, missense amino acid substitutions and an insertion mutation. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.18 *STK11* : serine/threonine kinase 11

STK11 is a tumor suppressor gene located on chromosome 19p13.3 and is composed of 9 exons. Its predicted encoded protein has 433 amino acids and is a member of the serine/threonine kinase family involved in regulation of cell polarity. A locus on chromosome 19p was found frequently heterozygously deleted in cases of Peutz-Jeghers syndrome (PJS). PJS is an autosomal dominant disorder characterized by melanocytic macules of the lips, buccal mucosa, and digits, multiple gastrointestinal hamartomatous polyps, and increased risk of various cancers (OMIM 175200). Positional

cloning of the locus identified *STK11* as the causative gene in PJS (Hemminki et al. 1997). The genomic sequence of *STK11* was found to be similar to a human serine/threonine kinase, previously named *LKB1*. *LKB1* was then renamed *STK11*. Mutation screening of *STK11* in cases of PJS revealed the presence of deletions, splice site mutations and nonsense mutations affecting one allele of STK11 (Jenne et al. 1998). Mutation of the second allele was presumed to follow the Knudson two-hit model and caused the manifestations of PJS. This finding made PJS the first cancer susceptibility syndrome due to inactivating mutations in a protein kinase (Hemminki et al. 1998).

Mutations of *STK11* in germline PJS and sporadic tumors are characterized by deletions, nonsense mutations, splice site mutations and missense mutations (Figure 1-21). All of these mutations result in loss-of-function of kinase activity of STK11 (Ylikorkala et al. 1999). Somatic mutations of *STK11* occur in lung cancers (10%), melanomas (10%), and pancreatic cancers (4%) (COSMIC).

In C. elegans, *STK11* mutants displayed disrupted asymmetries established normally during early embryogenesis (Kemphues et al. 1988). As well, in drosophila, *STK11* mutants disrupted the apical-basal epithelial polarity, suggesting a role of *STK11* in cell polarization. It was postulated that human *STK11* may be a functional homolog of *STK11* in drosophila. Therefore loss of *STK11*, and subsequent loss in cell polarity in humans may contribute to tumorigenesis (Martin and St Johnston 2003).

Homozygous inactivation of *STK11* in mouse leads to mid-gestation embryonic lethality and defects in the vasculature (Ylikorkala et al. 2001). However, inactivation of one allele of *STK11* in mouse led to the formation of gastrointestinal polyps similar to those found in FJP patients (Miyoshi et al. 2002). A similar finding of inactivation of one allele of *STK11* in FJP patients with gastrointestinal polyps (Hernan et al. 2004). Therefore, haploinsufficiency of *STK11* appears to be sufficient for the formation of polyps.

Further analysis of the molecular mechanisms characterizing *STK11* heterozygous polyposis revealed that cyclooxygenase-2 (COX2) is highly upregulated in mouse polyps coincident with activation of ERK1 and ERK2 (Rossi et al. 2002). COX2 was also highly upregulated in most of a series of human FJP polyps. This study identified COX2 as a potential target for chemoprevention in FJP patients (Rossi et al. 2002). There have been small pilot clinical trials evaluating the use of COX2 inhibitors, such as celecoxib, in FJP patients. The results show that a subset of patients respond to treatment with COX2 inhibitors with reduced polyposis (Udd et al. 2004).

Conditional knockout of *STK11* in mouse liver resulted in hyperglycemia and increased gluconeogenic and lipogenic gene expression (Shaw et al. 2005). It was subsequently discovered that STK11 signaling modulates the therapeutic effects of metformin, a drug FDA approved for use in the treatment of type 2 diabetes (Shaw et al. 2005). Retrospective

epidemiological analysis of metformin use has found an association with decreased cancer rates among diabetics (Evans et al. 2005). *in vitro*, the activity of phenformin, an analogue of metformin, in the activation of the AMP-activated protein kinase (AMPK) pathway needs a functional STK11 protein (Lizcano et al. 2004). Therefore, the use of metformin as chemoprevention in FJP patients may be of benefit as long as there is one wild type *STK11* allele (Katajisto et al. 2007).

Using yeast two-hybrid analysis, STK11 was found to exist in mammalian cells in a complex with two other proteins: STE20-related adaptor (STRAD) and mouse protein 25 (MO25) (Baas et al. 2003). STK11:STRAD:MO25 complexes have been shown to be present in a similar stoichiometry in mammalian cells (Alessi et al. 2006). In this complex STK11 becomes activated and phosphorylates at least 14 other related kinases belonging to the AMP-activated protein kinase (AMPK) subfamily, including AMPK which regulates cellular energy levels. STK11 is constitutively active in mammalian cells and under conditions of low cellular energy or stress, the binding of AMP to AMPK permits its phosphorylation and activation by STK11 (Alessi et al. 2006). The STK11-AMPK pathway functions as a cellular-energy sensing checkpoint, permissive to the growth and proliferation of cells only when there is sufficient availability of fuel (Alessi et al. 2006). Therefore, activation of AMPK pathway by antidiabetic drugs such as metformin, may reduce cancer progression.

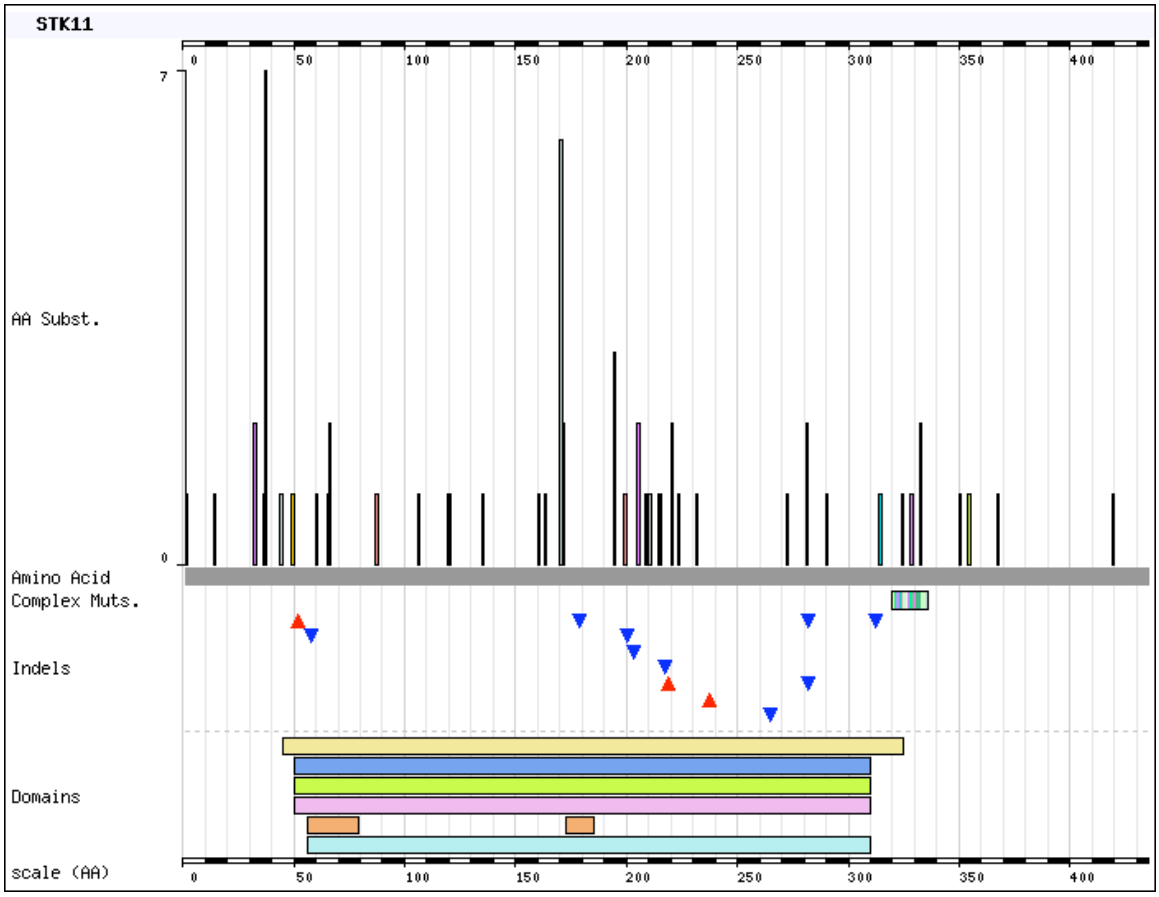

Figure 1-21: Distribution of somatic mutations in STK11. A schematic of the STK11 protein displays missense amino acid substitutions, frame-shift, deletion and nonsense mutations of the catalytic domain. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.19 *VHL* : von Hippel-Lindau tumor suppressor

VHL is a tumor suppressor gene located on chromosome 3p26-25 and is composed of 3 exons. Its predicted encoded protein has 213 amino acids and plays a critical role in oxygen-regulated signal transduction. *VHL* was mapped by genetic linkage analysis of von Hippel-Lindau (VHL) disease cases to chromosome 3p (Seizinger et al. 1988). *VHL* was isolated by

positional cloning (Latif et al. 1993). VHL disease is an autosomal dominant hereditary cancer syndrome characterized by a predisposition to developing tumors in a number of organs, including kidney, eye, pancreas, with renal clear cell carcinoma (RCC) a major cause of morbidity and mortality in VHL disease patients (OMIM 608537, Richards et al. 1998). VHL disease patients inherit an inactivated allele of *VHL*. The second allele is somatically inactivated during tumor development. The germline mutations of *VHL* are characterized by homozygous deletions, missense amino acid substitutions, nonsense and frame-shift mutations (Stolle et al. 1998). The mutations are widely distributed throughout the coding sequence (Zbar et al. 1996).

VHL disease can be classified into two types, based on the presence (type 2) or absence (type 1) of pheochromocytoma (Kim and Kaelin 2004). Type 2 VHL families almost always harbor missense mutations in *VHL* in contrast to type 1 families. Type 1 families frequently harbor truncation or deletion mutations in *VHL* (Zbar et al. 1996). VHL disease is one of several hereditary cancer syndromes associated with an increased risk of renal cell carcinoma. Somatic mutations of *VHL* are associated with renal clear cell carcinoma (RCC). Papillary renal carcinoma is associated with gain-of-function mutations of MET or loss of function mutations of fumarate hydratase (FH) (Kim and Kaelin 2004).

Somatic inactivating mutations of the *VHL* gene occur frequently in sporadic renal cell carcinoma (42%) and hemangioblastomas (22%) (COSMIC). These mutations are characterized either by truncation of VHL

protein by frame-shift or nonsense mutations or hypermethylation of the *VHL* promoter.

Tumors linked to *VHL* inactivation are highly vascular and can over produce angiogenic factors such as vascular endothelial growth factor (*VEGF*) and erythropoeitin (*Epo*) (Kim and Kaelin 2004). VEGF and Epo are hypoxia-inducible proteins. Cells lacking *VHL* constitutively over produce VEGF and Epo. Restoration of *VHL* suppresses the production of VEGF and Epo in the presence of oxygen (Krieg et al. 2000). *VEGF* and *Epo* are regulated by a transcription factor called hypoxia-inducible factor (*HIF*). There are two types of *HIF*, α and β. Formation of HIFα heterodimers is restricted to hypoxic conditions. HIFα heterodimers are unstable in the presence of oxygen. VHL and HIFα can bind to one another, and cells lacking VHL fail to degrade HIFα in the presence of oxygen (Maxwell et al. 1999). This provided evidence for the role of VHL as a ubiquitin ligase involved in the direct polyubiquitylation of HIFα subunits in the presence of oxygen (Kamura et al. 2000, Cockman et al. 2000).

Mutations of *VHL* result in the accumulation of HIF and leads to the increased transcription of growth stimulatory genes and angiogenic genes such as *VEGF*, *EGFR*, *TGF*, and *PDGF* (Linehan and Zbar 2004). The over production of angiogenic factors provides a rationale for the development of antibodies against VEGF or small molecule inhibitors of VEGF or HIFα. Sorafenib, a small molecule inhibitor of VEGF was approved by the FDA in 2006 for the treatment of RCC.

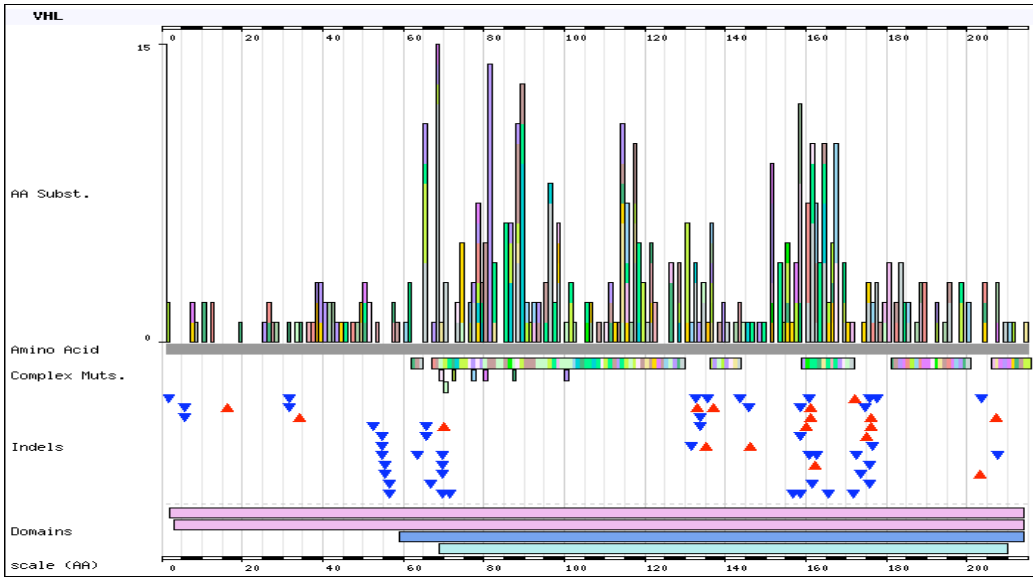

Figure 1-22: Distribution of somatic mutations in VHL. A schematic of the VHL protein displays nonsense, frame-shift and deletion mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.20 *CDKN2A* : cyclin-dependent kinase inhibitor 2A (melanoma, p16, inhibits CDK4)

CDKN2A is a tumor suppressor gene located on chromosome 9p21 and is composed of 3 exons. Its predicted encoded protein has 155 amino acids and is involved in cell cycle regulation. The 9p21 chromosomal region is frequently rearranged and deleted in dysplastic nevi (Cowan et al. 1988). The 9p21 locus was identified by genetic linkage analysis of cases of familial malignant melanoma (Petty et al. 1993). *CDKN2A* was first identified as a tumor suppressor gene and mapped to 9p21 when the chromosomal region

was also found to be deleted in more than 50% of human melanoma cell lines (Kamb et al. 1994).

CDKN2A gives rise to two distinct transcripts from different promoters with common exons 2 and 3 (Stone et al. 1995). The transcripts are called p16(INK4A), also known as the alpha transcript and beta transcript, p14(ARF). Mice null for INK4A have been shown to be prone to development of metastatic melanoma (Krimpenfort et al. 2001). However, mice null for ARF do not spontaneously develop tumors (Latres et al. 2000). These findings support data that mutations of INK4A are more common in human cancer than mutations of ARF (Sherr 2001). INK4A is involved in the regulation of RB1 and ARF is involved in the regulation of TP53.

Germline mutations of *CDKN2A* are characterized by deletions, missense, nonsense mutations, and frame-shift mutations in families with malignant melanoma. Somatic mutations of *CDKN2A* are characterized by similar modes with the exception of missense mutations (Figure 1-23). Somatic mutations are predominantly characterized by large homozygous deletions resulting in a absence of the CDKN2A gene product. Somatic mutations of *CDKN2A* occur in approximately 15% of all cancers and most frequently in pancreatic cancers (33%), biliary tract cancers (27%) and melanomas (22%) (COSMIC).

The G1 phase of the cell cycle is controlled by cyclin-dependent kinases (CDK4 and CDK6) and cyclins (D1, D2, D3). Cyclins and cyclin-

dependent kinases form complexes to phosphorylate retinoblastoma (RB1) and stimulate entry into the S phase. The cyclin/CDK complexes are inhibited by CDK inhibitor proteins such as CDKN2A. Loss-of-function mutations of CDKN2A impair its ability to inhibit the catalytic activity of the cyclin/CDK complexes and lead to uncontrolled cell growth (Ranade et al. 1995).

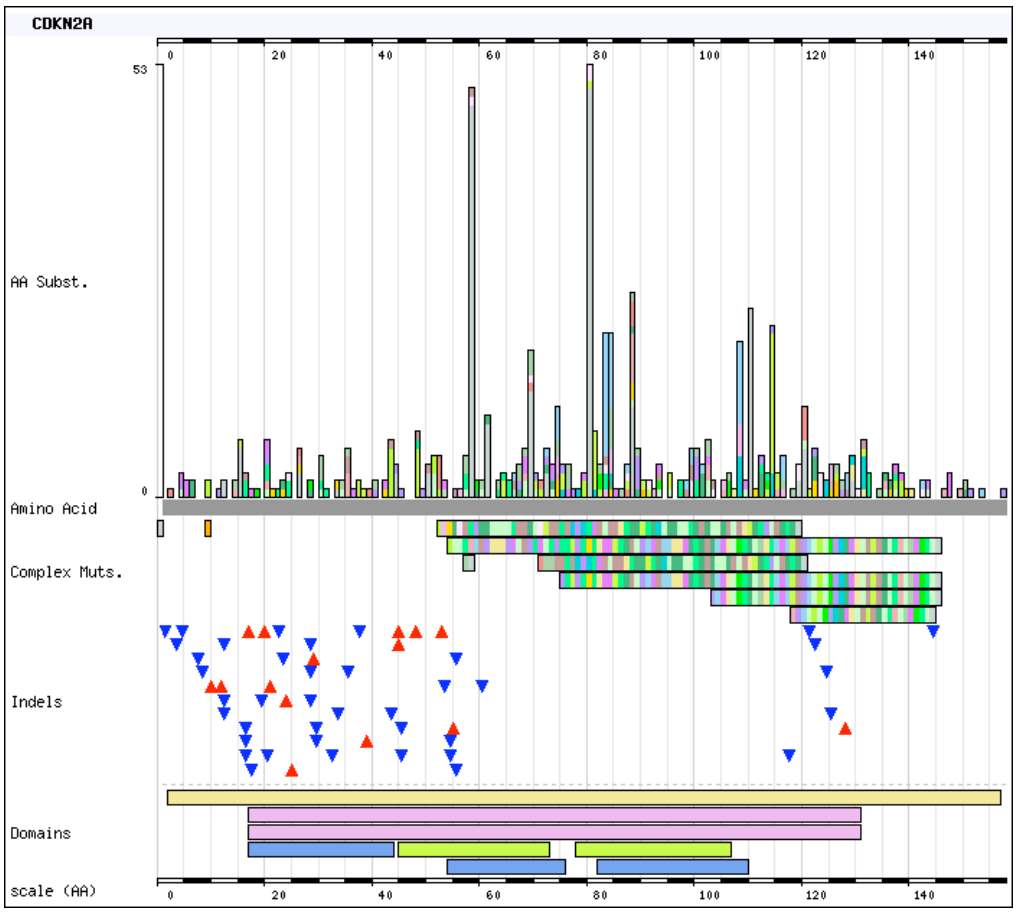

Figure 1-23: Distribution of somatic mutations in CDKN2A. A schematic of the CDKN2A protein displays nonsense, frame-shift, complex and deletion mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.21 *BRCA1* : breast cancer 1, early onset

BRCA1 is a tumor suppressor gene located on chromosome 17q21 and composed of 23 exons. Its predicted encoded protein has 1863 amino acids and is implicated in DNA double-strand break repair and cell-cycle checkpoint control. *BRCA1* was mapped to chromosome 17q by genetic linkage analysis of families with hereditary breast cancer (Hall et al. 1990). Hereditary breast cancer is characterized by young age at diagnosis, frequent bilateral disease, and frequent occurrence of disease among men (OMIM 114480). *BRCA1* was isolated by positional cloning and shown to be mutated by deletions, frame-shift mutations, missense and nonsense mutations in the germline of patients with hereditary breast-ovarian cancer syndrome (Miki et al. 1994, Smith et al. 1992, Kelsell et al. 1993).

A mutational screen in primary breast and ovarian cancers revealed the presence of germline not somatic *BRCA1* mutations (Futreal et al. 1994). All the BRCA1 mutations were identified in early-onset type breast and ovarian cancers. Therefore, *BRCA1* was thought not important in the development of most sporadic breast and ovarian cancers (Futreal et al. 1994). Unlike many other cancer genes which are involved in both familial and sporadic cancer, *BRCA1* mutations seem restricted to germline breast and ovarian cancer predisposition. A genetic linkage analysis of familial male breast cancer cases revealed that familial male breast cancer is not linked to the *BRCA1* locus (Stratton et al. 1994). *BRCA1* somatic mutations are rare (Figure 1-24) but can occur in sporadic breast, ovarian, and bladder cancers

(COSMIC). There is also evidence of hypermethylation of the *BRCA1* promoter in 10 to 15% of sporadic breast carcinomas (Esteller et al. 2001).

Mice null for *BRCA1* die in mid-gestation and have severe developmental abnormalities (Gowen et al. 1996). *In vitro* and *in vivo* biochemical assays have been used to demonstrate that BRCA1 may be important for cellular responses to DNA damage (Zhong et al. 1999). Specifically, BRCA1 inhibits the nucleolytic activity of the RAD50-MRE11-p95 complex implicated in numerous aspects of double-strand break repair (Paull et al. 2001). As well, reduction of the *BRCA1* transcript by RNA interference enhances cellular proliferation in normal human mammary epithelial cells (Furuta et al. 2005). Interestingly, mutant *BRCA1* sensitizes cells to the inhibition of PARP, a DNA repair enzyme, resulting in chromosomal instability, cell cycle arrest, and apoptosis (Farmer et al. 2005).

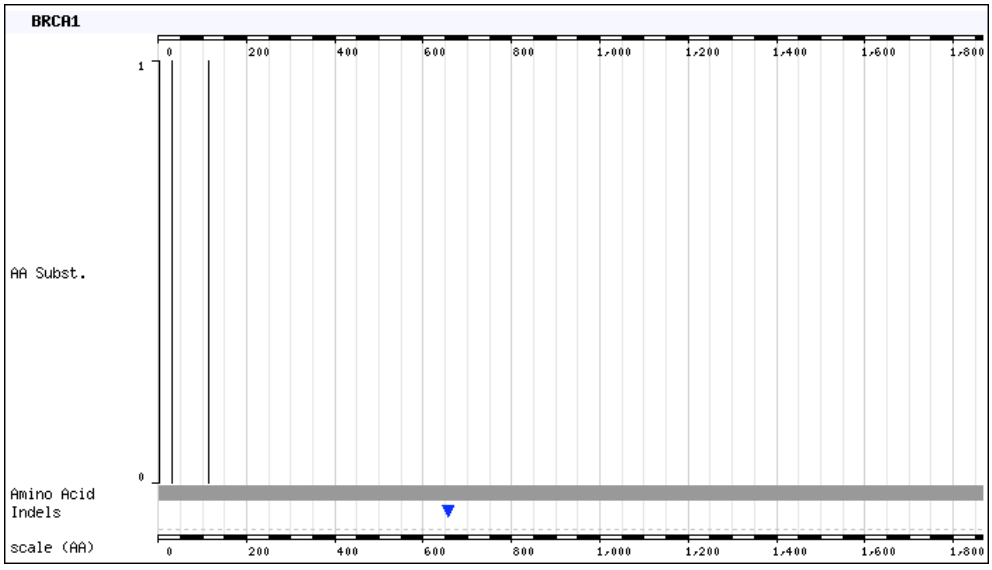

Figure 1-24: Distribution of somatic mutations in BRCA1. A schematic of the BRCA1 protein displays a single missense amino acid substitution in one

bladder cancer cell line and frame-shift mutation in one ovarian cancer cell line. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.22 *BRCA2* : breast cancer 2, early onset

BRCA2 is a tumor suppressor gene located on chromosome 13q12.3 and composed of 27 exons. Its predicted encoded protein has 3418 amino acids and is involved in DNA double-strand break repair. *BRCA2* was mapped by genetic linkage analysis of high-risk breast cancer families unlinked to the *BRCA1* locus (Wooster et al. 1994). Similar to *BRCA1*, the *BRCA2* locus seemed to confer a high risk of breast cancer. However, although there is an elevated risk of ovarian cancer, it is lower than for BRCA1 (Wooster et al. 1994). *BRCA2* was fine mapped and ultimately identified through the presence of truncating germline mutations (Wooster et al. 1995). These loss-of-function mutations are characterized by deletions, frame-shift and nonsense mutations. Occasional missense mutations of *BRCA2* have also been reported, but the significance of many is cleear. Fanconi anemia is an autosomal recessive disorder affecting the bone marrow and is associated with cardiac, renal, and limb malformations and changes in dermal pigmentation (OMIM 227650). Fanconi anemia can be caused by mutation of any of twelve Fanconi Anemia complementation group genes (OMIM 227650). One of the Fanconi anemia complementation groups, FANCD1, is due to homozygous or compound heterozygote biallelic *BRCA2* mutations (Howlett et al. 2002, Alter et al. 2007).

Somatic mutations of *BRCA2* are rare (Figure 1-25). A mutational screen of sporadic primary breast and ovarian cancers revealed that of the 2 mutations of *BRCA2* identified, one was a germline deletion and the other a somatic missense amino acid substitution of unknown significance (Lancaster et al. 1996, Teng et al. 1996). Mutations of *BRCA2* are infrequent in sporadic cancers of all types including breast and are associated with fewer cases of breast cancer than *BRCA1* (Krainer et al. 1997).

Mice null for *BRCA2* also die in mid-gestation with severe developmental abnormalities, similar to mice null for *BRCA1* (Suzuki et al. 1997). This suggests that *BRCA1* and *BRCA2* may have cooperative roles during embryogenesis (Suzuki et al. 1997). *In vivo*, BRCA2 has been shown to co-localize with FANCD2, implicating BRCA2 in the repair of replication-associated-double-strand-breaks (Hussain et al. 2004).

RAD51 is a key component of the mechanism DNA damage repair by homologous recombination. BRCA2 is also known to interact directly with RAD51 in a nuclear complex (Sharan et al. 1997). When DNA is damaged, BRCA1, BRCA2 and RAD51 localize to the damaged region. The BRCA2-RAD51 complex holds RAD51 in an inactive state. In the absence of BRCA2, RAD51 does not localize to areas of DNA damage (Narod and Foulkes 2004). Truncating mutations of *BRCA2* have been shown to inhibit the nuclear localization of BRCA2 protein. Nuclear localization of BRCA2 depends on two nuclear localization signals in the final 156 residues of BRCA2 (Spain et al. 1999). Truncated *BRCA2* protein localizes in the cytoplasm and therefore

cannot perform its DNA repair function in the nucleus (Spain et al. 1999). As with *BRCA1* mutations, mutant *BRCA2* sensitizes cells to PARP inhibition (Farmer et al. 2005). PARP is a DNA repair enzyme and when inhibited in the context of *BRCA* mutation, leads to apoptosis (Farmer et al. 2005).

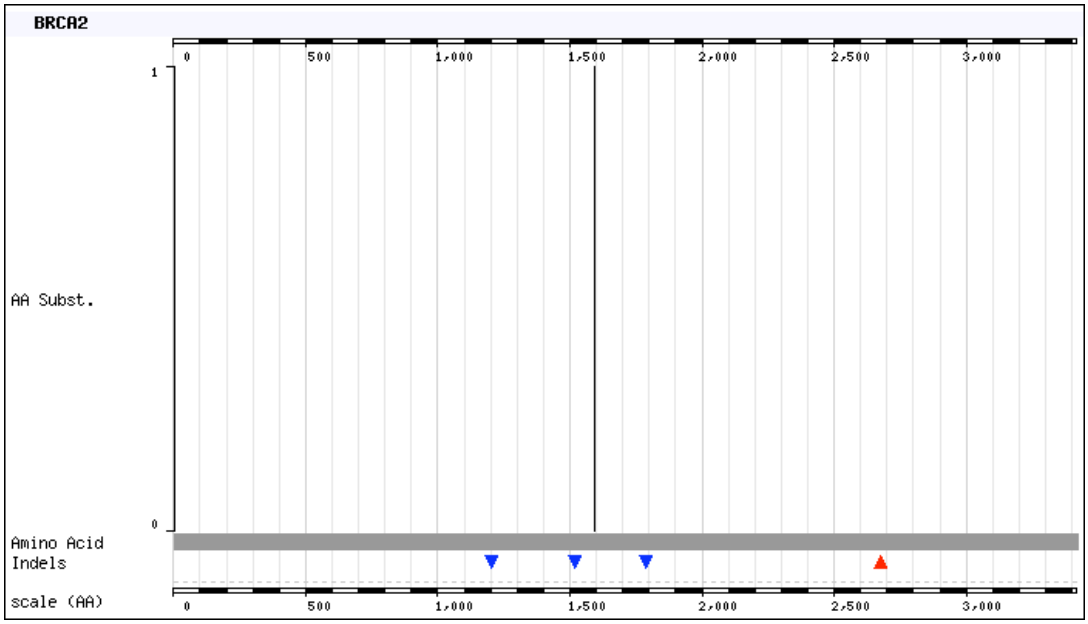

Figure 1-25: Distribution of somatic mutations in BRCA2. A schematic of the BRCA2 protein displays nonsense and frame-shift mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.23 *RB1* : retinoblastoma 1

RB1 is a tumor suppressor gene located on chromosome 13q14.2 and is composed of 27 exons. Its predicted encoded protein has 928 amino acids and is involved in cell cycle regulation. Retinoblastoma is an embryonic malignant cancer of retinal origin that presents early in childhood and is often

bilateral (OMIM 180200). The gene responsible for retinoblastoma, *RB1*, was mapped by recurring observations made by karyotype analysis of a deletion of chromosome 13q in patients with retinoblastoma (Grace et al. 1971, Wilson et al. 1973). The critical segment common to all the deletions was identified as 13q14 by Giemsa banding (Francke et al. 1976). *RB1* was identified by positional cloning and was the first tumor suppressor gene identified in human cancer.

RB1 is inactivated in hereditary and sporadic cases of retinoblastoma (Murphree and Benedict 1984). Germline mutations of *RB1* are detected by cytogenetic analysis of peripheral blood lymphocytes and show deletions, rearrangements and nonsense mutations (Lohmann et al. 1996). The germline mutations of *RB1* affect one allele and the second allele is somatically mutated during progression of the disease. Carriers of *RB1* mutations have an increased risk of acquiring secondary cancers which include osteosarcomas, melanomas and brain cancers (Lohmann 1999).

Somatic mutations of *RB1* occur in sporadic cases of retinoblastoma and other cancers (Horowitz et al. 1990). The somatic mutations of *RB1* are characterized by homozygous deletions, LOH, nonsense mutations, splice site mutations (Figure 1-26) and hypermethylation of the *RB1* promoter (Lohmann 1999). Somatic mutations of *RB1* occur frequently in cancers of diverse origin such as retinoblastoma (51%), bladder (33%), endometrial (29%), lung (12%), prostate (11%), breast (11%), gastric (10%), melanoma (10%) and central nervous system (10%) (COSMIC).

Mice deficient for *RB1* exhibit widespread development defects characterized by impaired neurogenesis and hematopoiesis and die mid-way through gestation (Lee et al. 1992). That suggests *RB1* is required for embryonic development and viability. RB1 plays a role in regulating cell cycle progression by a phosphorylation/dephosphorylation mechanism during cell proliferation and differentiation (Chen et al. 1989).

In a normal cell, RB1 controls the transition from G1 to S phase in the cell cycle. During the G1 phase of the cell cycle, RB1 binds transcription factors such as E2F in its unphosphorylated state (Hiebert et al. 1992). This binding causes a repression of E2F-mediated transcription. During late G1 and M phase, RB1 is phosphorylated by cyclin-dependent kinases (CDKs) (Taya et al. 1997). Upon RB1 phosphorylation, E2F is released and promotes the expression of genes required for cell division (Hsieh et al. 1999, Laurie et al. 2006). Therefore, inactivating mutations of *RB1* lead to uncontrolled cellular proliferation.

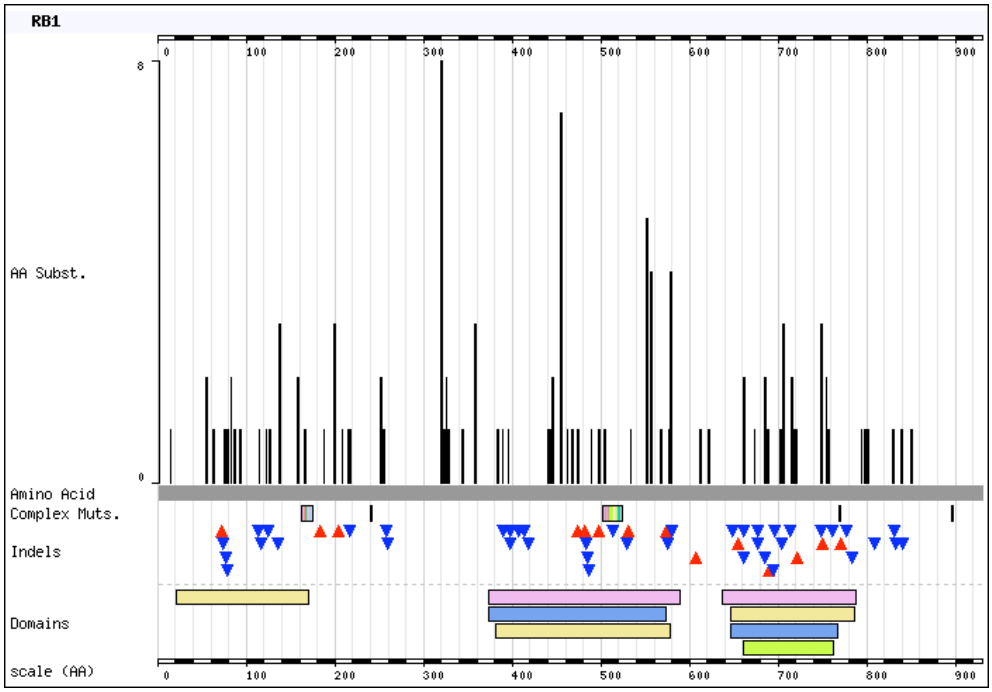

Figure 1-26: Distribution of somatic mutations in RB1. A schematic of the RB1 protein displays nonsense, frame-shift and complex mutations. Adapted from Catalogue of Somatic Mutations in Cancer (COSMIC) http://www.sanger.ac.uk/perl/genetics/CGP/cosmic.

1.1.5.4.24 *TP53* : tumor protein p53 (Li-Fraumeni syndrome)

TP53 is a tumor suppressor gene located on chromosome 17p13.1 and is composed of 10 exons. Its predicted encoded protein has 393 amino acids and plays a role in regulation of the cell cycle. *TP53* was identified as a tumor suppressor gene because the chromosomal region 17p was frequently deleted in colorectal cancers. Positional cloning localized the common region of deletion in a panel of colorectal cancers (Baker et al. 1989). This region was also frequently altered in lung cancer and the common region of deletion also contained the *TP53* gene (Takahashi et al. 1989). A survey of other cancer types for the presence of *TP53* alteration revealed that indeed *TP53* is

mutated in diverse human tumor types (Nigro et al. 1989). These somatic mutations of *TP53* are inactivating and are characterized by homozygous deletions, splice site mutations, deletion of one allele and missense, nonsense, and frame-shift mutations of the remaining allele.

Li-Fraumeni syndrome is an autosomal dominant disorder characterized by early onset of a variety of tumors, multiple tumors within an individual, and multiple affected family members (OMIM 151623). Patients with Li-Fraumeni syndrome are at an increased risk of sarcomas, breast cancers and brain cancers (Li and Fraumeni 1982). Because tumor suppressor genes, such as *RB1* had previously been identified to be associated with familial neoplasms, it prompted the search for a causative gene for Li-Fraumeni syndrome. Inactivating mutations of *TP53* were identified in sporadic cases of osteosarcomas, leukemias, colon, lung and brain cancers. These tumor types were also found in families with Li-Fraumeni. Therefore *TP53* was studied in families with Li-Fraumeni. Germline mutations of *TP53* were identified in almost all Li-Fraumeni families (Malkin et al. 1990). Germline mutations of other genes such as cell-cycle checkpoint kinase 2 (*CHK2*) have been associated with Li-Fraumeni syndrome (Bell et al. 1999). The germline mutations of *TP53* are characterized by deletions, missense, frame-shift, nonsense, and splice site mutations.

The missense *TP53* mutations lead to the synthesis of a stable, but inactive, protein that accumulates in the nucleus (Dowell et al. 1994). The

frame-shift, deletion and nonsense mutations of *TP53* lead to a truncated protein product. Most of the *TP53* mutations are missense and cluster in the conserved DNA binding domain (DBD) and oligomerization domain of the TP53 protein (Soussi and Beroud 2001), (Figure 1-27). The DBD and oligomerization domain of *TP53* corresponds to exons 5 through exon 8 of the gene. In contrast, most of the nonsense and frame-shift mutations occur outside of the DBD (Soussi and Beroud 2001).

TP53 plays an essential role in the regulation of the cell cycle, specifically in the transition from G0 to G1 phase (Vogelstein et al. 2000). TP53 inactivation is the most common genetic change in human cancers (Levine et al. 1991). TP53 is thought to bind as a homotetramer to a p53 binding site on DNA and activate expression of adjacent genes that inhibit growth and invasion. This binding is mediated by an oligomerization domain in the carboxyl terminus of the protein. However, missense mutations of the DBD and oligomerization domain, deletion of one or both alleles produce an altered TP53 protein incapable of binding DNA, thereby resulting in impaired transactivation of growth inhibitory genes (Vogelstein and Kinzler, 1992). In the absence of *TP53* mutations, TP53 can be inactivated by *MDM2*. MDM2 acts to stabilize TP53 by binding to its N terminus. *MDM2* is found amplified and overexpressed in some human sarcomas (Oliner et al. 1992). Disruption of the TP53-MDM2 interaction may be a viable therapeutic strategy for the treatment of cancers with TP53 or MDM2 alterations (Toledo and Wahl 2006).

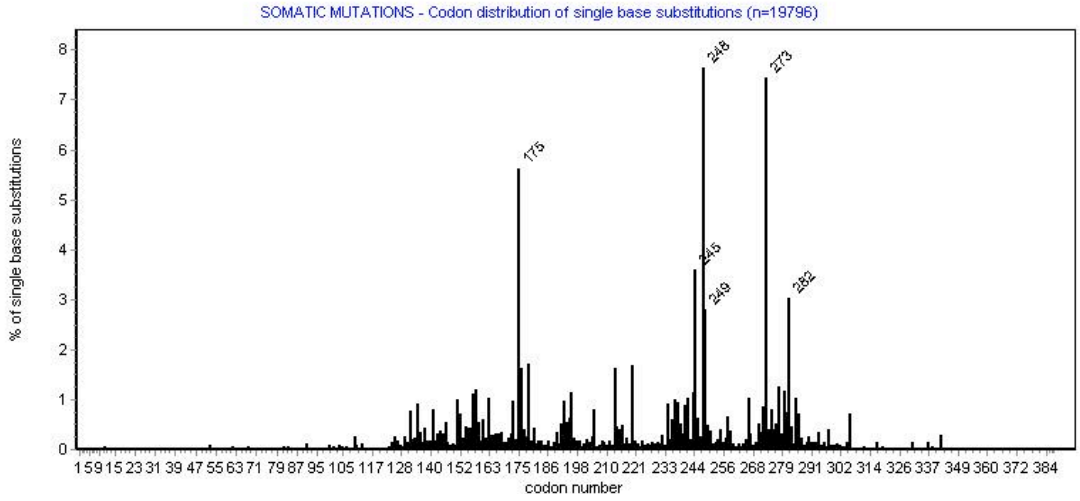

Figure 1-27: Distribution of somatic mutations in TP53. A schematic of the TP53 protein displays the codon distribution of missense substitutions in TP53. TP53 is also inactivated by nonsense, frame-shift, and complex mutations. Peaks of mutation correspond to the DNA binding domain and oligomerization domain. Adapted from International Agency for Research on Cancer (IARC)- http://www-p53.iarc.fr/index.

1.2 Pharmacogenomics

Pharmacogenomics is the study of the genetics of inter-individual response to drugs and aims at molecular subsetting of patients for more effective therapy (Weinstein and Pommier 2003). The field of pharmacogenomics is especially important in cancer chemotherapy where most clinically used drugs have a narrow therapeutic window, that is the difference between the dose required to achieve the desired therapeutic effect and that causing toxicity, is small (Weinshilboum 2004). Therefore, knowledge of genetic variations, inherited or acquired, that may predict differential response to cancer chemotherapy is key to individualized therapy (Figure 1-28).

Figure 1-28: Pharmacogenomics aims to identify patients at risk for toxicity or efficacy to therapy prior to medication selection. Adapted from Marsh & McLeod (2006).

Genetic variations in drug effect are classified into two groups: those due to either pharmacokinetic or pharmacodynamic factors. The

pharmacokinetic factors that influence drug effect involve the drug's absorption, distribution, metabolism and excretion (Lindpaintner 2002). In contrast, the pharmacodynamic factors that influence drug effect involve the transport of the drug into the cell and the interaction of the drug (ligand) and its target(s) (Lindpaintner 2002).

1.2.1 Germline variants as predictors of drug response

Many of the first examples from pharmacogenomics in cancer chemotherapy center on understanding the inherited (germline) inter-individual differences involved in drug metabolism. I have highlighted three classic examples (Table 1-1).

Examples of germline variants reported to modulate drug action		
Gene product	**Drug**	**Drug action (minor allele)**
TPMT	6- Mercaptopurine	Enhanced toxicity
UGT1A1	Irinotecan	Enhanced toxicity
CYP2D6	Tamoxifen	Decreased efficacy

Table 1-1: Clinically important germline variants reported to affect response to cancer chemotherapeutics. TPMT, thiopurine methyltransferase; UGT1A1, uridine 5'-diphosphate-glucuronosyl-transferase; CYP2D6, cytochrome P450 2D6.

Thiopurine drugs, such as 6-mercaptopurine, are purine antimetabolites used clinically to treat leukemias and are metabolized by the enzyme thiopurine methyltransferase (TPMT) (Weinshilboum et al. 1980). It

was discovered that patients with low levels of TPMT activity, due to a coding single nucleotide polymorphism, had elevated concentrations of mercaptopurine metabolites. The elevated levels of mercaptopurine metabolites led to an increased risk of myelosuppression. There is now an FDA approved genetic test (DNA-based) to determine which patients would be likely to experience severe myelosuppression prior to administration of mercaptopurine (Goetz et al. 2004).

Irinotecan is a topoisomerase I inhibitor used in combination with 5-fluorouracil (5-FU) as first line therapy for the treatment of metastatic colorectal cancer (Mathijssen et al. 2003). Irinotecan is metabolized to its active metabolite, SN-38, 1000 times more potent than the parent drug (Kaneda et al. 1990). SN-38 is inactivated by a polymorphic hepatic enzyme uridine diphosphate glucuronosyltransferase 1A1 (UGT1A1) (Iyer et al. 1998). It has been observed that human liver samples harboring a dinucleotide repeat sequence (seven TA repeats) in the UGT1A1 promoter have reduced metabolism of SN-38 (Iyer et al. 1999). Patients with the seven TA repeats of UGT1A1 have an increased risk for severe neutropenia and death than patients without the polymorphism (Ando et al. 2000). Recently, the FDA approved a genetic test to identify patients with the seven TA repeat polymorphism of *UGT1A1* prior to administration of standard doses of Irinotecan (O'Dwyer and Catalano et al. 2006).

Tamoxifen is an anti-estrogenic drug used for the treatment of estrogen receptor positive breast cancer (Goetz et al. 2007). Although the estrogen

receptor is a marker for predicted response to tamoxifen, not all women with estrogen positive breast cancer benefit from tamoxifen (Thurlimann et al. 2005). Tamoxifen is metabolized to the potent anti-estrogen, endoxifen, by the cytochrome P450 2D6 enzyme (CYP2D6) (Goetz et al. 2007). It has been reported that breast cancer patients taking tamoxifen with low levels of CYP2D6 activity, have significantly shorter time to disease recurrence than patients with higher CYP2D6 enzyme activity (Goetz et al. 2007). Importantly, CYP2D6 metabolism has been shown to be an independent predictor of breast cancer outcome in post-menopausal women receiving tamoxifen for treatment of early breast cancer (Goetz et al. 2007).

1.2.2 Somatic variants as predictors of drug response

There are examples where acquired mutations of the tumor DNA are predictive of response to cancer chemotherapy. The elucidation of the signal-transduction networks that drive neoplastic transformation has led to rationally designed cancer therapeutics that target specific molecular events (Sebolt-Leopold and English 2006). These targeted therapeutics, unlike traditional cancer chemotherapeutics, do not have narrow therapeutic indices. Many of the currently known drugs in this class are protein kinase inhibitors. Genes that encode protein kinases are often dysregulated and constitutively activated in cancer. Kinase inhibitors therefore reduce the activity of the activated protein kinases, reducing the cellular oncogenic drive and inducing tumour regression. I have highlighted three classic examples where somatic mutations of tumor DNA predict response to kinase inhibitors (Table 1-2).

Examples of somatic mutations reported to modulate drug action		
Gene product	**Drug**	**Drug action (mutant protein)**
BCR-ABL, KIT, PDGFRA	Imatinib	Tumor regression
ERBB2	Trastuzumab	Tumor regression
EGFR	Gefitinib, Erlotinib	Tumor regression

Table 1-2. Clinically approved kinase inhibitors and predicted response based on somatic mutations of cancer genes. BCR-ABL, break point cluster region-ABL; KIT, v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog; PDGFRA, platelet-derived growth factor receptor, alpha polypeptide; ERBB2, v-erb-b2 erythroblastic leukemia viral oncogene homolog 2, neuro/glioblastoma derived oncogene homolog (avian); EGFR, epidermal growth factor receptor.

1.2.2.1 BCR-ABL, KIT, PDGFRA kinase inhibitors

The BCR-ABL protein tyrosine kinase is the fusion of the BCR (Breakpoint cluster region) and non-receptor protein tyrosine kinase ABL that result from the reciprocal chromosomal translocation t(9;22) producing a shortened chromosome 22, called the Philadelphia (Ph) chromosome. This resultant chromosome has constitutive tyrosine kinase activity (Manley et al. 2002). The BCR-ABL protein is associated predominantly with chronic myeloid leukemia but also with acute lymphoblastic leukemia (Manley et al. 2002). Imatinib was the first small molecule kinase inhibitor approved as treatment for chronic myeloid leukemia via targeting of the BCR-ABL oncoprotein. Imatinib works by binding to an inactive conformation of the BCR-ABL protein kinase (Roskoski et a 2003; Capdeville et al. 2002).

Imatinib also has specificity for the PDGFRA and KIT protein kinases, and is used in treatment of malignancies associated with dysregulated forms of those proteins (Capdeville et al. 2002).

Treatment of CML patients with imatinib leads to complete cytogenetic and hematological remission, however imatinib fails to deplete leukemic stem cells that harbor the BCR-ABL fusion protein (Michor et al. 2005). Therefore, some patients develop resistance to imatinib, particularly in the advanced phases of CML and Ph-positive ALL (Weisberg et al. 2007). Mechanisms of imatinib resistance involve BCR-ABL amplification and over expression of mRNA and protein (le Coutre et al. 2000, Weisberg and Griffin et al. 2000). However the most common mechanism of resistance is the acquisition of point mutations in the kinase domain of the *ABL* gene (Gorre et al. 2001). The first point mutation identified as associated with imatinib resistance was T315I (Gorre et al. 2001). Crystal structure of an analogue of imatinib bound to the ABL kinase domain revealed that the T315 residue was crucial for the interaction between imatinib and ABL (Schindler et al. 2000). It was also found that a different mutation at the T315 residue (T315V) conferred constitutive kinase activity to ABL and was less sensitive to imatinib compared with wild type ABL (Corbin et al. 2002). The T315I point mutation impairs imatinib binding thereby reducing its tyrosine kinase inhibition of ABL.

More than 50 different point mutations in ABL associated with imatinib resistance have been reported (Weisberg et al. 2007). However most are rare and six amino acid residues (Gly250, Tyr253, Glu255, Thr315, Met351,

and Phe359) thus far account for 60 to 70% of imatinib-resistant mutations (Weisberg et al. 2007). Two of the more frequently detected ABL mutants, Y253F and E255K, have been shown to have high *in vitro* transforming potential. The *in vitro* finding is consistent with clinical findings that show P-loop mutations such as Y253F and E255K are associated with a greater likelihood of progression to blast crisis and shorter overall survival in imatinib treated patients (Soverini et al. 2005). Similarly the T315I mutation, generally found in patients with advanced CML, has a worse overall survival compared with other ABL mutations in patients on imatinib therapy (Nicolini et al. 2006).

Drug-resistant BCR-ABL point mutations can be found in imatinib-naïve CML or can be acquired during imatinib treatment. Acquired imatinib-resistance involves the re-emergence of BCR-ABL tyrosine kinase activity. This suggests that the mutant BCR-ABL protein is still a viable target for inhibition by small molecule inhibitors (Barthe et al. 2002, Branford et al. 2002). To this end, alternative therapies have been designed to overcome resistance to imatinib. One of these drugs, nilotinib, is approximately 30-fold more potent than imatinib as an ABL inhibitor (Weisberg et al. 2007). Phase I and II trials of nilotinib in patients with imatinib-resistant CML in all phases of disease and patients with Ph-positive ALL demonstrate significant clinical response to nilotinib (Kantarjian et al. 2006). Nilotinib also inhibits KIT and PDGFRB protein kinases (Manley et al. 2005). Another drug designed to overcome imatinib-resistance is dasatinib. Dasatinib is a potent inhibitor of BCR-ABL, Src-family kinases, KIT and PDGFR (Das et al. 2006, Melnick et al. 2006). In contrast to imatinib and nilotinib, dasatinib binds to the active

conformation of the ABL kinase (Tokarski et al. 2006). Based on data from phase I and II trials of dasatinib in patients with imatinib-resistant CML and Ph-positive ALL patients, it was recently approved in the US and Europe for the treatment of adults in all phases of CML with imatinib-resistance or intolerance (Cortes et al. 2007). Dasatinib was also approved for the treatment of patients with Ph-positive ALL with imatinib-resistance or intolerance (Talpaz et al. 2006).

Most gastrointestinal stromal tumors (GISTs) harbor oncogenic KIT or PDGFRA receptor tyrosine kinase mutations (Fletcher and Rubin 2007). The KIT or PDGFRA gain-of-function mutations are early events in GIST oncogenesis (Corless et al. 2002). Imatinib, a potent inhibitor of KIT signaling, has recently become first-line treatment of metastatic GIST following *in vitro* studies suggesting a therapeutic potential for imatinib in a human GIST cell line (Tuveson et al. 2001). Prior to treatment with imatinib, surgical resection of primary localized GIST was the only chance for cure (Gold and DeMatteo 2006). GISTs proved refractory to standard chemotherapy and radiation. Recurrence of disease was certain with a predicted 5-year survival of 30% (DeMatteo et al. 2000). Prospective trials of imatinib in metastatic GIST have shown that approximately 80% of patients will respond to imatinib and will have stable disease (Demetri et al. 2002, Verweij et al. 2003). In addition, 70% of metastatic GIST patients will have at least a 2-year disease-free survival and 50% will be free of disease progression (Verweij et al. 2004).

KIT mutations occur in up to 90% of GISTs and clinical response to imatinib is dependent on the presence of specific KIT mutations (Heinrich et al. 2003). *KIT* exon 11 mutations are found in 75% of GISTs and result in the abrogation of the juxtamembrane autoinhibition of the KIT kinase (Tarn et al. 2005). Patients with exon 11 *KIT* mutations have a higher response rate to imatinib treatment and longer time to treatment failure compared with other *KIT* or *PDGFRA* mutations (Joensuu 2006a). Patients without detectable *KIT* or *PDGFRA* mutation respond less frequently to imatinib treatment compared with exon 11 *KIT* mutants (Heinrich et al. 2003). However up to 39% of those patients without *KIT* or *PDGFRA* mutations do respond to imatinib (Heinrich and Corless et al. 2005). As well, patients with the rare exon 13 *KIT* mutation or exon 17 *PDGFRA* mutation may respond to imatinib (Joensuu 2006a). These data suggest that imatinib treatment ought to be considered for all GIST patients, regardless of *KIT* or *PDGFRA* mutation status. The only exception may be patients with primary imatinib-resistant mutation of PDGFRA (D842V) (Joensuu 2006a).

The majority of metastatic GIST patients will develop resistance to imatinib. The most common resistance mechanisms involve the acquisition of secondary exon 13, 14 or 17 *KIT* mutations that prohibit imatinib binding (Joensuu 2006a). Some of these secondary mutations are intrinsically imatinib-resistant, such as the frequently occurring V654A substitution (Heinrich et al. 2006). However, other mutations such as those involving the N822 residue are intrinsically imatinib-sensitive but are associated with clinical

imatinib-resistance when coincident with an exon 11 *KIT* mutation (Heinrich et al. 2006).

Sunitinib, an inhibitor of KIT, PDGFRA, FLT3, and VEGFR2, has recently been approved for the treatment of imatinib-resistant GIST and patients unable to tolerate treatment with imatinib (Joensuu 2006b). In a randomized phase III trial of sunitib in patients who had progressed on imatinib therapy, sunitinib was found to prolong median time to tumor progression compared with placebo (Demetri et al. 2006). Sunitinib provides a temporary benefit for imatinib-resistant GIST patients, therefore more therapeutic options are needed. Preclinical studies in GIST cell lines have shown that treatment with Heat shock protein 90 (HSP90) inhibitors resulted in degradation of the KIT oncoprotein and may therefore be of benefit in imatinib-resistant GIST (Bauer et al. 2006).

1.2.2.2 ERBB2 kinase inhibitors

ERBB2 protein tyrosine kinase amplification and over expression occurs in approximately 30% of metastatic breast cancer (Slamon et al. 1989) and shows clinical correlates with earlier relapse and shorter overall survival (Perren et al. 1991). Trastuzumab is a monoclonal antibody approved for the treatment of amplified or over expressed ERBB2 in metastatic breast cancer (Roskoski 2004) and was the first genetic based drug approved for cancer therapy. Trastuzumab is active as a single agent and in combination with chemotherapy in ERBB2 over expressing metastatic breast cancer. However, the response rates to trastuzumab monotherapy

range from 12% to 34% (Nahta and Esteva 2006). Primary resistance to trastuzumab monotherapy occurs in approximately 66% to 88% of ERBB2 over expressing metastatic breast tumors (Baselga et al. 1999, Vogel et al. 2002). Trastuzumab with adjuvant chemotherapy (paclitaxel or docetaxel) significantly improved disease-free and overall survival in patients with early stage ERBB2 over expressing breast cancers compared with trastuzumab monotherapy (Seidman et al. 2001, Slamon et al. 2001, Esteva et al. 2002).

ERBB2 is also overexpressed, to a lesser degree, in lung cancers, specifically adenocarcinomas and large-cell carcinomas and is predictive of poorer outcomes (Azzoli et al 2002, Shi et al. 1992, Brabender et al. 2001). Intragenic mutations have also been found in the conserved kinase domain of the *ERBB2* gene in some lung cancers (Stephens et al. 2004, Shigematsu et al. 2005b). The ERBB2 mutations seem to occur exclusively in non-small cell lung cancer (NSCLC) of adenocarcinoma histology and are more common in female patients and never smokers (Shigematsu et al. 2005b). So far, there has not been demonstrated benefit of trastuzumab monotherapy or in combination with cancer chemotherapeutics for the treatment of NSCLC with over expressed ERBB2 (Gatzemeier et al. 2004, Langer et al. 2004, Zinner et al. 2004, Clamon et al. 2005).

Invariably, the majority of patients who achieve an initial response to trastuzumab-based regimens develop resistance within one year (Nahta and Esteva 2006). The mechanisms of resistance (primary or acquired) have not been fully elucidated, however there are numerous proposed mechanisms.

Elucidating the molecular mechanisms underlying primary or acquired trastuzumab resistance is critical to improving the survival of metastatic breast cancer patients whose tumors over express ERBB2 (Nahta et al. 2006).

Resistance to trastuzumab has been associated with increased expression of the membrane-associated glycoprotein MUC4 (Price-Schiavi et al. 2002). MUC4 was shown to bind and sterically hinder ERBB2 from binding to trastuzumab (Price-Schiavi et al. 2002). In a trastuzumab-resistant cell line with ERBB2 amplification demonstrating primary resistance to trastuzumab, protein levels of MUC4 were shown to be inversely correlated with trastuzumab binding capacity (Nagy et al. 2005). Knockdown of MUC4 RNA increased the sensitivity of the resistant line to trastuzumab (Nagy et al. 2005). This study also found that binding of MUC4 to ERBB2 disrupted the interaction between ERBB2 and ERRB3 and EGFR (Nagy et al. 2005).

Compensatory signaling from other EGFR family members can disrupt the inhibitory effect of trastuzumab. Trastuzumab binds the domain IV of ERBB2 and domain II is involved in heterodimerization with EGFR and ERBB3 (Motoyama et al. 2002). Also increased signaling from other receptor types such as insulin growth factor I receptor (IGF-IR) have been shown to reduce trastuzumab-mediated growth arrest (Lu et al. 2001). IGF-IR interacts with ERBB2 in trastuzumab resistant cells but not in sensitive cells (Nahta et al. 2005). Inhibition of IGF-IR by antibody blockade or tyrosine kinase inhibition increased trastuzumab sensitivity *in vitro* (Nahta et al. 2005).

Altered downstream signaling from ERRB2 has been shown to confer primary resistance to trastuzumab. ERBB2 signaling activates the PI3K signaling pathway. Constitutive PI3K/AKT activity has been shown to inhibit trastuzumab-mediated cell-cycle arrest and apoptosis (Yakes et al. 2002). An ERBB2 over expressing breast cancer cell line, BT474, resistant to trastuzumab had elevated levels of phosphorylated AKT compared to the parent line (Chan et al. 2005). The resistant cells were sensitive to a small molecule inhibitor of PI3K (Chan et al. 2005). Patients with PTEN-deficient, ERBB2 over expressing breast tumors have a poorer response to trastuzumab-based therapy (Nagata et al. 2004). Subsequently, it was shown that in PTEN-deficient cells, PI3K inhibitors rescued trastuzumab resistance *in vitro* and *in vivo* (Nagata et al. 2004). Therefore, PTEN loss may serve as a predictor of trastuzumab resistance and that PI3K inhibitors may be potential therapies in PTEN-null trastuzumab-resistant tumors (Nahta and Esteva et al. 2006).

Novel therapeutic strategies are being employed to overcome resistance to trastuzumab. Pertuzumab, is a monoclonal ERBB2 antibody that represents a new class of drugs called dimerization inhibitors (Nahta and Esteva 2006). Pertuzumab can block signaling by other EGFR family receptors, as well as inhibit signaling in cells expressing normal ERBB2 levels. Pertuzumab sterically blocks dimerization of ERBB2 with EGFR and ERBB3, inhibiting signaling from ERBB2/EGFR and ERBB2/ERBB3 heterodimers (Agus et al. 2002). Pertuzumab is also able to disrupt interaction between ERBB2 and IGF-IR in trastuzumab-resistant cells (Nahta

et al. 2005). Trastuzumab and pertuzumab bind to different epitopes in the extracellular domain of ERBB2 (Cho et al. 2003, Franklin et al. 2004). The combination of trastuzumab and pertuzumab produced synergistic apoptosis in ERBB2 over expressing trastuzumab-naïve breast cancer cells (Nahta et al. 2004), without any significant effect on the viability of trastuzumab-resistant breast cancer cells (Nahta et al. 2005).

Another alternative therapeutic agent against trastuzumab-resistant tumors is lapatinib, a dual tyrosine kinase inhibitor targeted against both EGFR and ERBB2. Binding of lapatinib to EGFR and ERBB2 is reversible but its dissociation is much slower allowing for prolonged down regulation of receptor tyrosine phosphorylation (Nahta and Esteva 2006). ERBB2 status and not EGFR status is a determinant of lapatinib activity (Nahta and Esteva 2006). It has been shown that combination of lapatinib with trastuzumab enhanced apoptosis of ERBB2 over expressing breast cancer cells (Xia et al. 2005). Resistance to lapatinib seems to be mediated by increased signaling from the estrogen receptor in estrogen receptor-positive ERBB2 over expressing breast cancers (Xia et al. 2006). This suggests that targeting of both the estrogen receptor and ERBB2 may be beneficial in that subset of cancer patients (Xia et al. 2006). A recent phase III trial of trastuzumab-resistant ERBB2 over expressing breast cancer patients demonstrated that combination of lapatinib and capecitabine resulted in longer median progression-free survival compared with capecitabine alone (Geyer et al. 2006).

1.2.2.3 EGFR kinase inhibitors

EGFR protein tyrosine kinase over expression has been implicated in numerous cancer types (Sridhar et al. 2003). Gefitinib and erolotinib are EGFR tyrosine kinase inhibitors marketed as single drug therapy for chemotherapy-refractory advanced non-small cell lung cancer (NSCLC) (Kris et al 2003, Fukuoka et al. 2003). Reports have demonstrated that somatic mutations in the conserved kinase domain of the *EGFR* gene are associated with sensitivity to Gefitnib (Lynch et al. 2004; Paez et al. 2004) and Erlotinib (Pao et al. 2004a).

Somatic mutations of EGFR kinase domain are most common in NSCLC. Rarely mutations of EGFR are found in head and neck cancers, cholangiosarcomas, colon cancers, ovarian cancers, esophageal cancers, and pancreatic cancers (Guo et al. 2006, Gwak et al. 2005, Lee et al. 2005, Nagahara et al. 2005, Schilder et al. 2005, Kwak et al. 2006). In lung cancers EGFR kinase domain mutations are more common in adenocarcinomas, East Asians, women, and never smokers (Shigematsu et al. 2005a). Mutations outside of the kinase domain are rare in NSCLC (Lynch et al. 2004, Paez et al. 2004, Pao et al. 2004a). However, mutations of the extracellular domain of EGFR are common in gliomas (Mellinghoff et al. 2005) and squamous cell lung cancers (Ji et al. 2006).

In NSCLC EGFR mutations are commonly associated with amplification (Kaye et al. 2005). Studies have shown that patients with EGFR amplification were more likely to respond to gefitinib or erlotinib and

had longer median time to disease progression and overall survival compared to patients with normal EGFR copy number (Riely et al. 2006b). However, it remains to be established whether amplification of wild type EGFR contributes to lung cancer development and response to gefitinib or erlotinib (Riely et al. 2006b).

Several mutations of EGFR have been reported. Thus far there are five mutations known to confer sensitivity to EGFR tyrosine kinase inhibitors. The drug-sensitive mutations are: point mutations in exon 18 (G719A or G719C), point mutations of exon 21 (L858R and L861Q), and in-frame deletions of exon 19 that eliminates four amino acids (LREA) (Riely et al. 2006a). The most common of these drug-sensitive mutations are the exon 19 in-frame deletion and exon 21 missense amino acid substitution (L858R) accounting for up to 90% of EGFR mutations in NSCLC (Riely et al. 2006a).

In retrospective studies, the association between the presence of EGFR mutation and sensitivity to gefitinib and erlotinib is quite consistent showing 75% response rate for patients with EGFR mutations compared to 10% response rate for patients with wild type EGFR (Bell et al. 2005b, Tsao et al. 2005, Han et al. 2005, Mitsudomi et al. 2005, Uramoto et al. 2006). Prospective trials have confirmed the association between EGFR mutation and sensitivity to gefitinib and erlotinib. These studies showed that 78% of patients with somatic exon 19 deletion or exon 21 L858R mutation had radiographic responses to gefitinib and erlotinib (Inoue et al. 2006, Sunaga et al. 2007, Sutani et al. 2006).

Large phase III retrospective trials have been conducted where NSCLC patients were randomized to receive either standard cytotoxic chemotherapy alone or standard chemotherapy in combination with gefitinib or erlotinib (Eberhard et al. 2005, Bell et al. 2005b). These studies have reported that patients with EGFR mutations have prolonged survival compared with patients with wild type EGFR treated with gefitinib or erlotinib. Interestingly, these studies have also found that the prolonged survival may occur in the absence of treatment with gefitinib, erlotinib, surgery, or standard cancer chemotherapy (Riely et al. 2006b). In the standard cytotoxic chemotherapy alone treatment arm, patients with EGFR mutation had prolonged progression-free and overall survival compared with patients with wild type EGFR (Eberhard et al. 2005, Bell et al. 2005b).

The mechanism of increased sensitivity of EGFR mutants to Gefitinib and Erlotinib is still unknown. However, it has been shown that the exon 19 in-frame deletions and exon 21 L858R missense amino acid substitution confer ligand-independent activation and prolonged kinase activity after ligand stimulation (Lynch et al. 2004, Paez et al. 2004). It has also been shown that the exon 21 L858R mutant form of EGFR has approximately 20-fold higher catalytic efficiency than that of the wild type EGFR (Zhang and Chang et al. 2007).

There are reported differences in the clinical course between patients with exon 19 in-frame deletions and patients with exon 21 L858R missense

substitution. One study has shown that NSCLC patients with the L858R mutation treated with surgery alone, have a prolonged overall survival compared with NSCLC patients with exon 19 deletions (Shigematsu et al. 2005a). Another study has shown that after treatment with gefitinib or erlotinib, NSCLC patients with EGFR exon 19 deletions have a longer overall survival compared with patients harboring the exon 21 L858R mutation (Riely et al. 2006a). The molecular basis for such findings is yet unknown.

Despite a dramatic initial response to gefitinib and erlotinib, NSCLC patients with EGFR mutations rarely achieve a complete response and resistance to treatment develops. There are thus far, three EGFR kinase domain mutations associated with drug resistance: an exon 19 point mutation (D761Y), an exon 20 point mutation (T790M), and an exon 20 insertion (D770_N771insNPG) (Riely et al. 2006b). The most common of these drug-resistant mutations is the T790M reported to occur in about 50% of tumors after disease progression (Kobayashi et al. 2005a, Pao et al. 2005a). The T790M mutation has been predicted to block the binding of gefitinib or erlotinib to the kinase ATP binding pocket. This mutation is analogous to the acquired drug-resistance to imatinib seen in GIST and CML (Riely et al. 2006b). Interestingly, the T790M mutation has been seen in the germline and tumor DNA of family members with hereditary bronchioloalveolar carcinoma (Bell et al. 2005a). One of the family members with the T790M mutation did not respond to treatment with gefitinib (Bell et al. 2005a). *In vitro* data has suggested that irreversible EGFR inhibitors may have activity in patients with acquired resistance to gefitinib or erlotinib (Kobayashi et al. 2005b, Kwak et

al. 2005). A phase II trials of HKI-272, an irreversible EGFR kinase inhibitor, is ongoing to determine the efficacy in patients who have progressed after initial treatment with gefitinib or erlotinib (Riely et al. 2006b).

In addition to mutations of EGFR other molecular parameters involved in the EGFR signaling cascade are associated with activity of gefitinib or erlotinib. For example, it has been shown, by cDNA microarray analysis, that increased expression of TGF-α, a ligand for EGFR, is associated with poor response to gefitinib (Kakiuchi et al. 2004). Also, increased expression of heregulin, a ligand for ERBB3, is associated with insensitivity to gefitinib (Zhou et al. 2006). It has been reported that increased copy number of ERBB2 in the presence of EGFR mutation, is associated with response to gefitinib (Cappuzzo et al. 2005). However, NSCLC patients with ERBB2 mutations do not respond to gefitinib or erlotinib (Wang et al. 2006). Downstream of EGFR, it has been observed that NSCLC patients with KRAS mutations are resistant to gefitinib or erlotinib (Pao et al. 2005b). AKT is phosphorylated on EGFR activation transmitting signals for cell survival (Sordella et al. 2004). It has been reported that increased phosphorylation of AKT is predictive of response to gefitinib or erlotinib (Cappuzzo et al. 2004, Pao et al. 2004b). A novel drug-resistant gefitinib mutation has recently been reported, that does not involve mutation of EGFR. It was reported that gefitinib resistant clones from an *EGFR* mutant lung cancer cell line displayed amplification of MET oncogene and maintained activation of ERBB3/PI3K/AKT signaling in the presence of gefitinib (Engelman et al. 2007). Following the initial observation in cell lines, a panel of 18 gefitinib or

erlotinib resistant primary lung tumors were assessed for MET amplification. MET amplification was found to occur in 22% of those tumors (Engelman et al. 2007).

1.3 BRAF as a potential drug target in melanoma

The studies described above indicate that understanding cancer genetics is key to the further development of targeted therapies. BRAF is a serine/threonine kinase recently found to be dysregulated in up to 80% of melanomas (Davies et al. 2002). I have highlighted below the major efforts underway to identify targeted therapies against mutant BRAF for the treatment of melanoma.

Metastatic melanoma is a difficult disease to treat. The current standard of care is the use of cytotoxic agent, dacarbazine, hydroxyurea and interleukin (IL)-2. However randomized controlled trials have shown there is no significant overall survival advantage with the use of any single drug or combination of drugs (Kalinsky and Haluska 2007).

As reviewed earlier, BRAF is mutated in up to 80% of melanoma and the predominant mutation is the V600E amino acid substitution in the kinase domain. The mutation confers constitutive and elevated kinase activity to the BRAF protein leading to increased phosphorylation of downstream effectors MEK and ERK. The V600E BRAF mutation also occurs frequently in benign melanocytic nevi, implicating the presence of BRAF mutation as an initiating event the development of melanoma. In terms of prognostic value of the

V600E BRAF mutation, one study has shown that patients with the V600E BRAF mutation have longer survival (Kumar et al. 2003), whereas another study reports the opposite (Houben et al. 2004). Therefore the role of the V600E BRAF mutation in melanoma progression is yet to be elucidated.

Somatic mutations of NRAS are found in approximately 15% of melanomas and are rarely coincident with V600E BRAF mutation (Davies et al. 2002, Thomas et al. 2007). This suggests redundancy of the RAS-RAF-MEK-ERK pathway. It has been demonstrated that dysregulation of the PI3K-AKT pathway is involved in melanoma progression. RAS activates the PI3K-AKT pathway and RAS mutation has been shown to be mutually exclusive with mutation of either BRAF or PTEN (Tsao et al. 2004). It has been shown that most melanomas with PTEN inactivation also harbored concurrent BRAF mutation. It has also been shown that expression of phosphorylated AKT increases significantly with melanoma progression and invasion and inversely correlated with survival (Dai et al. 2005). The data support the notion of a cooperative role of both the RAS-RAF-MEK-ERK and PI3K-AKT pathways in melanoma progression (Tsao et al. 2004).

Following the elucidation of the dysregulated pathways leading to the development of melanoma recent efforts have focused on developing targeted therapeutics against mutant BRAF protein. Below, I review the major therapeutic strategies currently being investigated for treatment of BRAF mutant melanoma.

1.3.1 BRAF inhibitors

Sorafenib is a small molecule kinase inhibitor initially identified as a potent inhibitor of CRAF (Flaherty 2007). During phase II trials, hypertension was observed in patients taking sorafenib. It was then hypothesized that sorafenib may be an inhibitor of VEGF (Flaherty 2007). Sorafenib was subsequently shown to have specificity for wild type BRAF, V600E BRAF mutant, VEGFR-2, VEGFR-3, and PDGFR-β (Wilhelm et al. 2004). Preclinical studies of sorafenib showed it was able to block ERK activation and cell proliferation in V600E BRAF mutant melanoma cell lines and delay tumor growth *in vivo* (Karasarides et al. 2004). Phase I trials showed sorafenib was well tolerated as monotherapy or in combination with chemotherapy (Kalinsky and Haluska 2007).

Sorafenib was recently approved for treatment of metastatic renal cell carcinoma. In phase III trials, sorafenib monotherapy was shown to offer progression-free survival benefit to patients with advanced renal cell carcinoma due to its inhibition of VEGFR-2 (Escudier et al. 2007). However clinical benefit of sorafenib montherapy in metatstatic melanoma has not been established. In fact, in a phase II randomized discontinuation trial of sorafenib no relationship was noted between the presence of V600E BRAF mutation and disease stability (Eisen et al. 2006). Sorafenib is currently being tested in combination with other chemotherapeutic agents.

There are mechanistic reasons for the lack of benefit in treatment of BRAF mutant melanoma with sorafenib. In addition to the ability of sorafenib

to inhibit wild type BRAF and V600E BRAF, it also has specificity for CRAF (Karasarides et al. 2004). In fact sorafenib has up to 20 times more specificity for CRAF than for V600E mutant BRAF (Flaherty 2007). Drugs with higher affinity for BRAF may be more active in BRAF mutant melanoma and are currently being developed and some are being evaluated in phase I trials (Kalinsky and Haluska 2007).

1.3.2 MEK inhibitors

Mutations in the kinase domain of BRAF render the protein constitutively active (Davies et al. 2002). This constitutive activity of the BRAF kinase increases the phosphorylation of its downstream effector, MEK, increasing growth signals (Davies et al. 2002). It has been shown that cells with mutant BRAF display *in vitro* sensitivity to MEK inhibition compared with wild type BRAF and RAS mutant cells (Solit et al. 2006). In mutant BRAF xenografts, MEK inhibition completely inhibited tumor growth whereas RAS mutant tumors had a partial response (Solit et al. 2006). Currently a phase II multicenter open-label randomized trial is underway comparing the efficacy of a MEK inhibitor versus temozolomide in advanced melanoma patients (Kalinsky and Haluska 2007).

1.3.3 Heat shock protein (HSP) 90 inhibitors

HSP90 is a molecular chaperone involved in maintaining the conformation, stability, activity and cellular localization of many oncogenic proteins (Powers and Workman 2006). These proteins include ERBB2, AKT, CRAF, and mutant TP53 (Powers and Workman 2006). Therefore inhibiting

the action of HSP90 has the potential to simultaneously abrogate multiple signaling pathways implicated in tumorigenesis. Recently, it has been shown that wild type ARAF and CRAF proteins require HSP90 for stability, whereas wildtype BRAF does not (Grbovic et al. 2006). Instead V600E BRAF mutant requires HSP90 for its stability and function (da Rocha Dias et al. 2005, Grbovic et al. 2006).

Preclinical studies showed that indeed treatment of V600E BRAF mutant melanoma cells with an HSP90 inhibitor, 17-allylamino-17-demethoxygeldanamycin (17-AAG), resulted in degradation of mutant BRAF, inhibition of MAPK activity, induction of apoptosis and *in vivo* antitumor activity (Grbovic et al. 2006). However, there was no relationship between BRAF mutation status and sensitivity to 17-AAG in melanoma cell lines (Powers and Workman 2006). Phase I trials have demonstrated that 17-AAG can be safely administered. Phase II trials are currently being carried out to evaluate the benefit of administering 17-AAG in combination with cancer chemotherapeutics (Powers and Workman 2006).

1.3.4 Highthroughput screen for oncogenic BRAF inhibitors

Recently a highthroughput screen of approximately 64,000 compounds was carried out to identify inhibitors of V600E BRAF kinase activity and the cascade signaling to ELK1, a direct substrate of ERK1/2 (Newbatt et al. 2006). The largest cluster of hits was identified to occur with 3-(3'-hydroxybenzylamino) 5 pyridines (Newbatt et al. 2006), with IC50 measures ranging from 0.5-37.9 μM in the V600E BRAF kinase assay and 0.8 μM to

more than 50 μM in the ELK1 cascade signaling assay. Therefore, those compounds may serve as a starting point for further drug discovery and development of V600E BRAF inhibitors.

1.4 NCI-60 cell lines

As highlighted in the above examples, molecular profiling of tumors in combination with drug response profiles will aid in building predictors of response to chemotherapy. A panel of 60 human cancer cell lines (NCI-60) for which extensive molecular profiles and drug response profiles have been collected, serves as a tool for interrogating relationships between molecular profiles of cancer and drug response.

The NCI-60 cell lines consist of 60 human cancer cell lines assembled by the US National Cancer Institute for anticancer drug discovery and represents nine tissue of origin types: breast, colon, central nervous system, renal, lung, melanoma, ovarian, prostate, and hematogenous. The NCI-60 was assembled in the late 1980s with the aim of changing the emphasis of drug discovery from leukemia to human solid tumors (Shoemaker 2006). Since then, more than 100,000 compounds have been screened for anticancer activity in the NCI-60. The resulting data have proved rich in information about the mechanisms of action and resistance of those compounds (Paull et al 1989, Weinstein et al. 1992, Weinstein et al. 1997).

In addition to the extensive drug response profiles in the NCI-60, the lines have been extensively profiled on the DNA, RNA, and protein levels

(Weinstein 2004). On the DNA level, DNA copy number changes have been assessed by array-based comparative genomic hybridization (aCGH) (Garraway et al. 2005, Bussey et al. 2006) and chromosomal aberrations have been catalogued by spectral karyotyping (Roshke et al. 2003). Prior to the initiation of this thesis, at the DNA sequence level, mutations of four known cancer genes had been analyzed: *TP53* (O'Connor et al. 1997), *KRAS, NRAS, HRAS* (Koo et al. 1996). At the transcript level, RNA expression has been studied on various array-based platforms (Scherf et al. 2000, Ross et al. 2000, Szakacs et al. 2004). As well, protein expression has been analyzed by two-dimensional gel electrophoresis and by reverse-phase lysate arrays (Myers et al. 1997, Nishizuka et al. 2003).

1.4.1 Major scientific outcomes from the NCI-60 anticancer drug screen

The various data on the NCI-60 have been integrated and analyzed resulting in several lead compounds with possible therapeutic implications. This effort continues to be a major focus in the discovery of potential drug targets in cancer as well as identifying molecular profiles that may predict response to available chemotherapeutic agents. I will highlight three examples of useful outcomes from the NCI-60 cell line screen: i) identification of MDR-1 inverse compounds, ii) L-asparaginase for treatment of ovarian cancer, iii) MEK inhibitors for the treatment of BRAF mutant cancer.

1.4.1.1 MDR-1 inverse compounds for treatment of drug-resistant cancers

Multidrug resistance (MDR-1) gene belongs to the family of ATP-binding cassette (ABC) transporters. MDR-1 and its protein product (Pgp) are overexpressed in tumors as a mechanism of developing drug resistance. As a component of the molecular profiling efforts on the NCI-60, the expression of the complete family of 48 ABC transporters, including MDR1, was characterized in the NCI-60 (Szakacs et al. 2004). The expression profiles were correlated with the drug activity of compounds tested in the NCI-60 anticancer drug screen and a compound, thiosemicarbazone, was identified as a drug lead for targeting cancer cells overexpressing MDR1 (Ludwig et al. 2006). This finding has implications for secondary treatment of drug resistant tumors with overexpression of MDR-1 (Ludwig et al. 2006).

1.4.1.2 L-asparaginase for the treatment of ovarian cancers

Asparagine synthetase (ASNS) is ubiquitously expressed in most mammalian cells and is responsible for the biosynthesis of L-asparagine (ASN) (Jousse et al. 2004). In contrast to normal cells, leukemic cells are ASN auxotrophs that fail to express ASNS (Peng et al. 2001). Therefore leukemic cells cannot synthesize ASN and rely on exogenous forms. In those cells, deprivation of exogenous ASN by a drug, L-asparaginase induces cell death. This has formed the basis for the clinical use of L-asparaginase in the treatment of childhood ALL (Broome et al. 1963; Boyse et al. 1967).

Recently a correlation was identified between ASNS expression and copy number and response to L-asparaginase in ovarian cancer cells (Bussey et al. 2006). Ovarian cancer cells expressing lower levels of ASNS were more sensitive to knockdown of ASNS plus treatment with L-asparaginase than ovarian cancer cells expressing higher levels of ASNS (Lorenzi et al. 2006). This finding has initiated retrospective clinical trials using ASNS expression as a biomarker for L-asparaginase activity (Lorenzi et al. 2006).

1.4.1.3 MEK inhibitors for treatment of BRAF mutant cancers

As previously discussed, BRAF mutation leads to constitutive activity of the BRAF kinase and increases the phosphorylation of MEK (Davies et al. 2002). As part of the molecular profiling of the NCI-60, mutations of BRAF were genotyped in the NCI-60 cell lines (Garraway et al. 2005). It was found that BRAF mutant cells of the NCI-60 were more sensitive to growth inhibition by compounds known to inhibit MEK activity than BRAF wild type cells (Solit et al. 2006). The *in vitro* selectivity was shown to translate into *in vivo* xenograft efficacy, which supported the further development of MEK inhibitors for treatment of melanoma (Shoemaker 2006).

1.5 Introduction to the thesis project

The NCI-60 has been characterized pharmacologically and at the molecular level more extensively than any other set of cell lines. However, prior to this thesis, there had not been a large-scale systematic sequence analysis of genes causally implicated in oncogenesis in the NCI-60. We

hypothesized that mutations in cancer genes in the NCI-60 are associated with differential drug sensitivity. Therefore this thesis reports on the following:

i) Sequence analysis of twenty-four known cancer genes in the NCI-60;

ii) Statistical analysis of relationship between mutations in cancer genes and drug activity;

iii) Experimental validation of statistically significant associations between cancer gene mutation and drug activity.