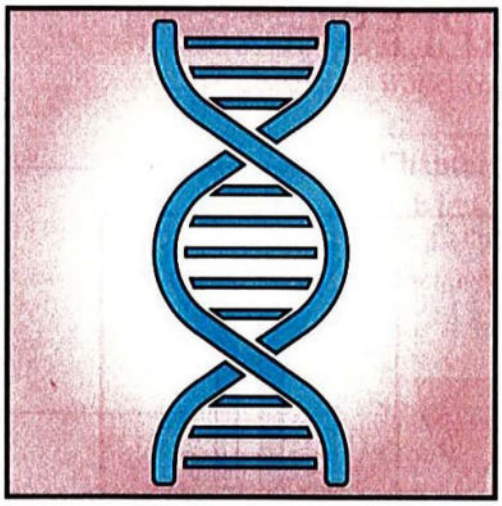
Contents

1 Cancer Genetics

Malte Renz
Allison W. Kurian



Cancer is considered a disease of the genome. It arises from a series of alterations in the genetic material that results in unrestrained cell proliferation and invasion of territories reserved for other cells. Even under physiologic conditions, the genome is constantly undergoing genetic alterations, and the vast majority of these mutations are being repaired. **If a mutation persists, however, the location of the mutation within the genome determines its effect.** Some alterations in the genetic material have more impact than others and can affect oncogenes or tumor suppressor genes, thus promoting or suppressing cancer development.

Historic Perspective

Over the centuries, various theories on cancer and its etiology have been developed and revised. **Only in the 19th century was it recognized that cancer consists of an accumulation of cells** (1)—as opposed to originating from an imbalance of body fluids, for example—and that those cancer cells could spread and metastasize throughout the organism (2). **In the early 20th century, Theodor Boveri**[1] (3) **hypothesized that cellular processes gave rise to cancer and attributed cancer development to chromosomal abnormalities that were acquired during cell division.** He described certain chromosomes that were amplified during cancer development and promoted cell division. Boveri (4) called those chromosomes "**oncogenes.**" **Later, it was recognized that single genes rather than entire chromosomes promoted cancer development.** Furthermore, **the mutagenic effect of certain chemicals, x-rays, cigarette smoke, and other exposures on the genetic material were described** (5). **Evidence for the clonal origin of cancer arising from a single genetically aberrant cell and its stepwise progression via premalignant forms was presented in 1990** by Bert Vogelstein in a multistep model for colorectal cancer development (5).

To investigate underlying genomic mechanisms of cancer development in detail, a human genome-wide analysis was proposed in 1986 and realized in the United States from 1990 to 2003 as the **Human Genome Project** (6). Subsequently, between 2006 and 2012, **The Cancer Genome Atlas (TCGA)** collected data on 12 different types of cancers with the goal of comparing the complete cancer genomes with the normal human genome and identifying characteristic and common genetic abnormalities in human cancers (7,8).

Genetic Changes

Sporadic

The majority of genetic alterations happen sporadically during life. Various environmental factors may increase the frequency of their occurrence, that is, the mutational frequency. **The vast majority are repaired by different DNA repair mechanisms.** As a result, the mutational rate is about 1 mutation in 1 billion base pairs (9,10).

Hereditary

Some of the genetic alterations may affect the germlines and are thus hereditary. There are only a few known hereditary genetic alterations that give rise to gynecologic malignancies. **Hereditary breast and ovarian cancer (HBOC) syndrome**

[1]Theodor Boveri (1862–1915), German biologist.

and Lynch syndrome are the major syndromes relevant for gynecologic malignancies, both of which are caused by genetic alterations of high penetrance.

Oncogenes

Oncogenes are genes whose gain in function has the potential to transform a normal cell into a tumor cell. Mutant germline oncogene alleles typically follow an autosomal dominant inheritance pattern.

Originally, the term "oncogene" was used by Theodor Boveri in the beginning of the 20th century to describe specific cell division promoting chromosomes in plants. In 1969, Huebner and Todaro (11) at the United States National Cancer Institute (NCI) rediscovered the term "oncogene" for specific cancer-promoting genes (i.e., not entire chromosomes) in the context of viral carcinogenesis. **Experiments with the Rous sarcoma virus** (12) **resulted, in 1970, in the discovery of *v-src* (viral-sarcoma) gene in the chicken retroviral Rous sarcoma virus (RSV)**, which later turned out to be the first confirmed oncogene (13,14). In 1976, further experiments showed that the *src* gene of Rous sarcoma virus (v-*src*) was a transduced allele of a cellular gene, the **c-*src***, which was incorporated up by the virus during recombination into the retroviral life cycle (15). These experiments demonstrated that **oncogenes are activated proto-oncogenes**, which are present in many organisms and thereby changed the view on cancer development: **cancer was no longer considered to be caused by foreign substances but rather by changes in a cell's own genetic material.**

Proto-oncogenes naturally occur in any cell type and regulate cell growth and differentiation. If proto-oncogenes gain in function, they turn into oncogenes. Gain in function can occur by different mechanisms: (i) **point mutations** may result in higher protein activity or loss of regulatory function and thereby overexpression; (ii) **increase in protein or mRNA stability** may prolong its half-life; (iii) **gene duplication** may result in higher expression; (iv) **chromosomal translocation** to new nearby regulatory domains may cause overexpression; (v) **gene fusions** may result in overexpression of the fusion protein. Furthermore, protein expression of oncogenes can be regulated by microRNA; (vi) **mutations in microRNA,** known as oncomirs, can lead to oncogene overexpression and thereby gain in function (16).

Types of Oncogenes

Cell Membrane Receptors

Examples of cell membrane receptors that may function as oncogenes include the receptor tyrosine kinases (RTKs). Receptor tyrosine kinases translate signals from the extracellular milieu to the inside of the cell by binding polypeptide growth factors that cannot cross the plasma membrane. Upon ligand binding, these RTK monomers dimerize. The process of dimerization brings intracellular kinase domains of the receptor proteins in proximity and permits transphosphorylation and subsequent activation of intracellular signaling pathways. Thereby, extracellular signals are propagated through the plasma membrane into the inside of a cell. **In the following, the most cancer relevant receptor tyrosine kinases are described.**

ErbB (its name is derived from the avian **er**ythro**b**lastosis oncogene **B**, whose human equivalent is EGFR or ErbB1). This receptor family includes four receptor tyrosine kinases (ErbB1–4 = Her 1–4). **The most clinically relevant members of the ErbB receptor family are ErbB1 and ErbB2. ErbB1 = EGFR** (epidermal growth factor receptor), is its first and best-described member. **ErbB2 = Her2neu** was derived from a rodent glioblastoma cell line (**neu**ral tumor). All family members are structurally related, single transmembrane receptors with an intracellular tyrosine kinase domain and an extracellular ligand binding domain. Eleven potential growth factor ligands have been identified thus far, which differentially bind to different members of this receptor family and include epidermal growth factor (**EGF**), transforming growth factor-α (**TGF-α**), heparin-binding EGF-like growth factor (**HB-EGF**), amphiregulin, betacellulin, epigen, epiregulin, and neuregulin 1–4.

Upon activation with a growth factor ligand, the ErbB protein receptors form homodimers and heterodimers and possibly higher-order oligomers. Notably, Her2neu has no known ligand but is the preferred dimerization partner for the other ErbB family members. Ligand binding changes the conformation of the extracellular domain from a *tethered* to an *untethered* form allowing for dimerization of the receptor proteins. Ectodomain dimerization results in close approximation of two cytoplasmic domains such that transphosphorylation can occur. **Activation of the tyrosine kinase domain leads in turn to the activation of a range of downstream signaling cascades**, most importantly the **mi**togen-**a**ctivated **p**rotein **k**inase (**MAPK**) (Fig. 1.1) and **p**hospho**i**nositide **3** kinase (**PI-3K**) (Fig. 1.2) pathways but also **p**hospho**l**ipase **C**-γ (**PLCγ**), extracellular signal-**r**egulated **k**inase **1-2** (**ERK1-2**) and others.

PDGFR (Platelet-Derived Growth Factor Receptor)

The ligands PDGF-A–D form homo- or heterodimers and bind to the monomeric platelet-derived growth factor receptor (PDGFR)-α and β with different affinity. The extracellular part of the receptor consists of five immunoglobulin-like domains with the ligand binding sites locating to the first three immunoglobulin-like domains. Upon ligand binding, these receptor isoforms dimerize, transphosphorylate and "unlock" their tyrosine kinase activity by autophosphorylation. PDGFRs predominantly stimulate the PI-3K pathway (Fig. 1.2).

VEGFR (Vascular Endothelial Growth Factor Receptor)

Five different ligands, that is, VEGF-A–E, bind differentially to the three receptors vascular endothelial growth factor receptor (VEGFR) 1–3. These receptors may be membrane-bound (mbVEGFR) or soluble (sVEGFR). The membrane-bound receptor system is mainly expressed in endothelial cells. As such, **VEGFR 1 and 2 play important roles in vasculogenesis during early development and angiogenesis later in life as well as vascular permeability, while VEGFR 3 enhances lymphangiogenesis.** Furthermore, **this receptor system functions as an immune modulator with effects on monocytes**

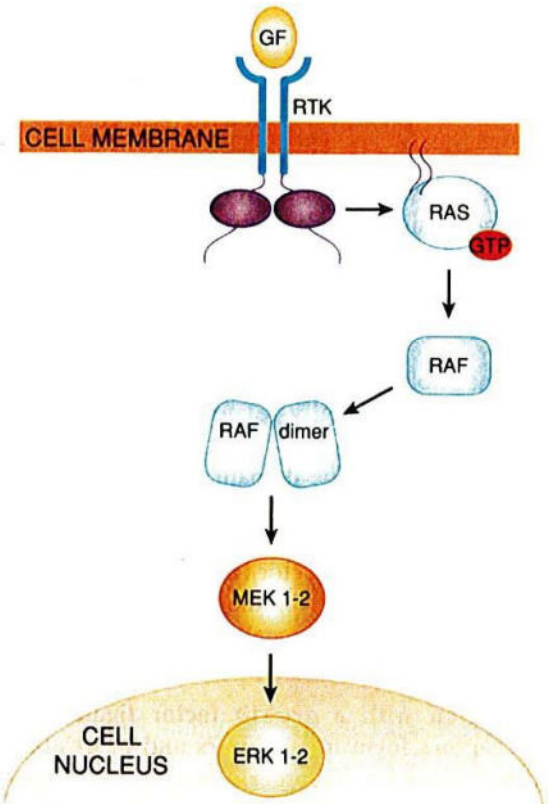


Figure 1.1 Schematic illustration of the mitogen-activated protein kinase (MAPK) signaling cascade. Binding of a number of different growth factors (GFs) to receptor tyrosine kinases (RTKs) results in receptor dimerization, autophosphorylation and activation of Ras (Rat sarcoma), which in turn activates rapidly activated fibrosarcoma (Raf). Dimerized activated Raf signals to MAPK/ERK kinase (MEK1-2). MEK1-2 activates the transcription activator extracellular signal-regulated kinase (ERK1-2), which translocates into the cell nucleus. Through the convergence of various growth factor ligands and intracellular divergence of the activating signal, the MAPK signaling cascade is involved in the complex regulation of gene expression, cell cycle progression, translation, mRNA stability, cell adhesion and migration.

and macrophages (VEGFR 1), **effector and regulator T cells** (VEGFR 2) (17), **dendritic cells** (VEGFR 1) (18,19) **and myeloid-derived suppressor cells** (MDSCs). Effects on sensory (VEGFR 1) and motor neurons (VEGFR 2) have been described as well. The VEGF receptors consist of seven immunoglobulin-like extracellular domains, a single transmembrane domain, and an intracellular domain with a split tyrosine kinase domain. VEGFRs stimulate the MAPK pathway (Fig. 1.1).

RET (Rearranged During Transfection)

The three isoforms of the transmembrane receptor rearranged during transfection (RET) are RET51, RET43, and RET19. They bind the glial cell line derived neurotrophic factor (GDNF) family ligands (GFLs). In order to activate RET, these ligands need to form a complex with a glycosyl phosphatidyl-inositol (GPI)-anchored co-receptor, the GFR-α (GDNF receptor). Only then can RET be activated to dimerize and transphosphorylate its kinase domain. In general, the signaling cascades activated by RET include MAPK (Fig. 1.1), PI-3K (Fig. 1.2), c-Jun N-terminal kinases (JNKs), and PLCγ-dependent pathways. The different activated signaling cascades seem to be cell type specific.

Cytoplasmic Oncogenes

Ras (Rat Sarcoma)

Rat sarcoma (Ras) proteins are a family of small GTPases involved in signal transduction. Mutations in the three human Ras proteins, *HRas, KRas, NRas,* are the most common alterations in human cancers and found in 20–30% of cancers. HRas and KRas were first identified as viral oncogenes in Harvey rat sarcoma and Kirsten rat sarcoma viruses and later recognized as oncogenes in human cancer cells (20–22). NRas was initially identified in neuroblastoma cells (23,24). Ras proteins consist of two domains, the larger G domain that binds and hydrolyzes guanosine nucleotides and a short C-terminal membrane targeting domain. Membrane binding of Ras proteins may be facilitated by prenylation and palmitoylation. In the GTP-bound state, the Ras proteins are active and have a high affinity for their numerous effector proteins. After hydrolyzation of GTP to GDP, the Ras proteins become inactive. Ras' intrinsic hydrolyzation activity is slow and inefficient. GTPase activating proteins (GAPs) enhance the intrinsic GTPase activity, accelerate hydrolyzation of GTP to GDP, and thereby inactivate Ras. Guanine nucleotide exchange factors (GEFs), on the other hand, catalyze push and pull reactions, which release GDP from Ras. Because GTP exists in a cell in about 10-fold excess, GTP is more likely to bind to Ras than GDP. Thereby, GEFs facilitate Ras activation. Constitutively active Ras proteins have been described, which show mutations that prevent GTP hydrolysis, and thereby lock Ras permanently in the "on" state. Ras activates several downstream pathways including the MAPK (Fig. 1.1) and PI-3K (Fig. 1.2) signaling pathway.

Raf (Rapidly Activated Fibrosarcoma)

The name of this serine/threonine-kinase family is derived from a structurally related mouse sarcoma retroviral oncogene that induces fibrosarcoma growth. A-Raf, B-Raf, and C-Raf are the family members. The rapidly activated fibrosarcoma (Raf) kinases participate in the mitogen activated protein kinase (MAPK) cascade and activate downstream MEK (MAPK/ERK kinase) proteins. They are composed of three conserved domains: a Ras-GTP–binding self-regulatory, that is, autoinhibitory domain, a serine-rich hinge domain, and a catalytic protein kinase domain. The kinase domain is inhibited by both autoinhibition by its own regulatory domain and the lack of phosphorylation of the hinge region. Binding of Ras-GTP to the self-regulatory domain, releases the catalytic domain and the Raf proteins become active serine/threonine kinases. In their active form, Raf proteins form dimers.

Src (Sarcoma)

Cellular-sarcoma (C-src) was the first discovered proto-oncogene. Sarcoma (src) is a non-RTK, which is ubiquitously expressed in all cell types. It consists of six domains: the so-called unique region, the src homology 2, 3, 4 domains (**SH2, SH3, SH4**), the catalytic domain, and a short regulatory tail. Its activation results in dephosphorylation of Y257, destabilization

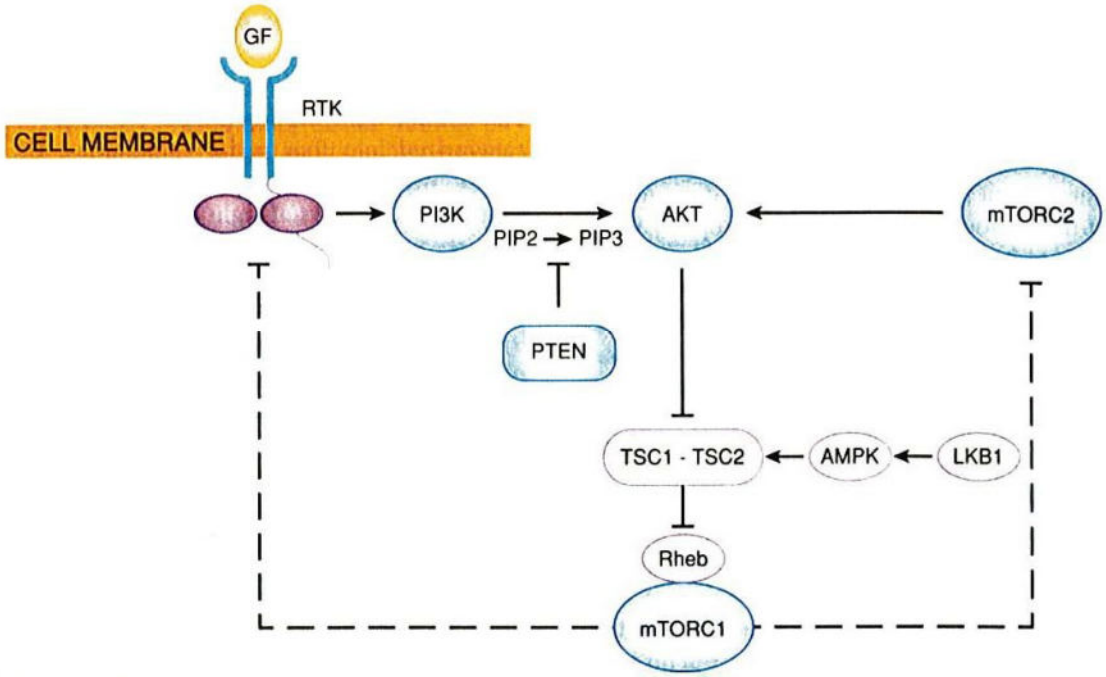


Figure 1.2 Schematic illustration of the phosphoinositide 3-kinase or phosphatidylinositol 3-kinase (PI-3K) signaling cascade. PI-3K is divided into four classes. Most is known about class I, which is further subclassified into class IA and IB. The schematic shown here summarizes some of the functions of class IA whose role in cancer is best described. The PI-3K pathway is activated by various growth factor (**GF**) receptor tyrosine kinases (**RTKs**) or G protein–coupled receptors (not shown). Class I PI-3K proteins are recruited to the plasma membrane by adapter proteins and are composed of a catalytic subunit and a regulatory subunit (not shown). Gain in function of the catalytic subunit, for example, the phosphatidylinositol 3-kinase, catalytic subunit α (***PIK3CA***), and loss in function of the regulatory subunit, for example, the phosphatidylinositol 3-kinase regulatory subunit α (***PIK3R1***), can result in overactivity of the PI-3K pathway. Activated class I PI-3K phosphorylates phosphatidylinositol 4,5-bisphoshate (**PIP2**) to phosphatidylinositol 3,4,5-trisphosphate (**PIP3**). The phosphatase and tensin homolog (***PTEN***) dephosphorylates PIP3 to PIP2 and thereby counteracts PI-3K. PIP3 activates **AKT** kinases, which in turn phosphorylate tuberous sclerosis protein 1 (**TSC1**) and 2 (**TSC2**). Phosphorylation of TSC results in dissociation of the TSC1–TSC2 complex. The intact TSC complex inhibits the kinase mTOR. Thus, activation of AKT results in activation of the **mTOR** complex **1** (**mTORC1**). Ras homolog enriched in brain (**Rheb**) in its GTP-bound form activates mTORC1. mTORC1 is involved in negative feedback to prevent overactivation of AKT. Depicted is the interaction and activation of the TSC1–TSC2 complex by serine/threonine kinase **11** (***STK11***)/liver kinase **B1** (***LKB1***) and adenosine monophosphate-activated protein kinase (**AMPK**). Not depicted here are interactions with the MAPK pathway, which on different levels can result in activation of the PI-3K pathway. The PI3K-AKT-mTOR signaling is involved in multiple cellular processes including metabolism, motility, proliferation, growth, and cell survival.

of the protein and opening of the SH2, SH3, and catalytic domains, which permits access of src substrates to the catalytic domain. Autophosphorylation of the tyrosine Y416 and phosphorylation of src substrates is enhanced by src dimerization. Src can attach to cellular membranes via myristoylation and can be activated by various transmembrane receptors; most studied examples are the activation by PDGFR and EGFR.

Nuclear Oncogenes

Myc (Myelocytomatosis)

The myelocytomatosis (Myc) proteins, c-myc, l-myc, and n-myc, constitute a family of transcription factors. ***C-myc* is located on chromosome 8 and reported to regulate the expression of 15% of all genes.** In mouse models, tissue-specific overexpression of myc proteins causes cancer development in those tissues, demonstrating the potency of the myc proteins. **Myc proteins are activated by various mitogenic stimuli including the Wnt** (derived from wingless and int-1) **and MAPK** (Fig. 1.1) **signaling pathway.** Myc proteins bind to DNA via a basic helix-loop-helix motif (**bHLH**); with the leucine zipper motif, they heterodimerize with another bHLH transcription factor. It is thought that myc proteins activate gene expression by binding to enhancer box sequences (E-boxes), that is, directly through activation of transcriptions, and recruiting histone acetyltransferases, that is, indirectly through the spatial reorganization of nucleosomes that make DNA more accessible for the transcription machinery.