PART I

Biology of the Gastrointestinal Tract

1 Cellular Growth and Neoplasia

Ezra Burstein

CHAPTER OUTLINE

MECHANISMS OF NORMAL TISSUE HOMEOSTASIS 1
Cellular Proliferation 1
Apoptosis 2
Senescence 3
Signaling Pathways That Regulate Cellular Growth 3
INTESTINAL TUMOR DEVELOPMENT 5
Multistep Formation 5
Clonal Expansion 5
Cancer Stem Cells 5
Epithelial-Mesenchymal Transition 5
NEOPLASIA-ASSOCIATED GENES 6
Oncogenes 6
Oncogenic Growth Factors and Growth Factor Receptors . . 7
Nuclear Oncogenes 7
Tumor Suppressor Genes 8
DNA Repair Genes 9
Noncoding RNAs 9
Oncogenic Signaling Pathways 10
TUMOR MICROENVIRONMENT 10
TUMOR METABOLISM 10
Inflammation and Cancer 10
Microbiome 10
BIOLOGICAL FEATURES OF TUMOR METASTASIS 11
Angiogenesis and Lymphangiogenesis 11
ENVIRONMENTAL INFLUENCES 11
Chemical Carcinogenesis 11
Dietary Factors 11
MOLECULAR MEDICINE: CURRENT AND FUTURE APPROACHES IN GASTROINTESTINAL ONCOLOGY 12
Next Generation Sequencing 12
Molecular Diagnostics 12

Normal cellular proliferation and differentiation are essential to tissue homeostasis in all organs, including the digestive tract. The neoplastic process involves a fundamental disruption of these mechanisms, which can give rise to cancer development and metastasis with the additional acquisition of other hallmarks of cancer. As a group, malignancies of the GI tract are the leading cause of cancer-associated mortality, and it is therefore essential to understand the underlying biology that gives rise to tumor formation. This chapter reviews mechanisms of normal cell growth and the fundamental cellular and molecular alterations that facilitate malignant transformation. The basic concepts discussed in this chapter provide the framework for discussion of specific GI neoplasms in later chapters.

MECHANISMS OF NORMAL TISSUE HOMEOSTASIS

Cellular Proliferation

Tissue homeostasis is maintained by the delicate balance of cellular proliferation and differentiation, which provide new cellular elements to replace dying cells as part of normal tissue function or during tissue repair. At a fundamental level, neoplasia arises when cell proliferation escapes the homeostatic mechanisms that maintain this process in balance with senescence and programmed cell death. Cell proliferation occurs as cells divide, a process that occurs through an orderly set of steps referred to as the cell cycle (Fig. 1.1). In preparation for cell division, there is a period of biosynthetic activity called the *G_1 phase* that is typically associated with an increase of cell size. This phase is followed by precise duplication of the genome, designated the *S phase*. After an intervening gap period designated as the *G_2 phase*, mitosis occurs during the *M phase*.

The commitment to proceed to DNA replication occurs at the G_1/S checkpoint or restriction (R) point. Cells may exit this cycle of active proliferation before reaching the R point and enter a quiescent phase known as G_0. Cells can subsequently reenter the cell cycle from the G_0 state (see Fig. 1.1). Another checkpoint exists at the boundary between the G_2 and M phases. The G_2/M checkpoint ensures that mitosis does not proceed prior to the repair of any damaged DNA after genome replication. Impaired function of these checkpoints is frequently observed in cancers.

Regulation of cell cycle progression is achieved principally by a set of proteins known as cyclins and cyclin-dependent kinases (CDKs). These proteins are expressed in specific parts of the cell cycle and regulate the G_1/S and G_2/M checkpoints. During the G_1 phase, cyclins D and E are most active.[1] Overexpression of cyclin D1 in fibroblasts results in more rapid entry of cells into the S phase, and, consistent with a role in cancer, cyclin D1 is frequently overexpressed in a number of GI and non-GI malignancies.[2] During the S phase, cyclin A is predominantly expressed, and by the G_2 phase cyclin B is the main regulator (see Fig. 1.1).

Each cyclin forms a complex with a CDK and function as catalysts for CDK activity in a cell cycle–dependent fashion (see Fig. 1.1). The cyclin-CDK complexes regulate cell cycle progression through phosphorylation of key target proteins. For example, cyclin D1–dependent progression from G_1 to S phase is the result of cyclin D1/CDK4 phosphorylation of the tumor suppression pRb, the product of the retinoblastoma gene, as well as the Rb family members p130 and p107.[3] These proteins sequester E2F transcription factors that promote expression of factors required for S phase, and their phosphorylation by CDK4 leads to their functional inhibition. Thus, loss of Rb expression also accomplishes more rapid progression to S phase and is another genetic lesion seen in many tumors. An analogous circuit is found in the G_2/M transition, where cyclin A/CDK2 mediates the activation of another transcriptional regulator, FoxM1, required for the expression of factors involved in mitosis.[4]

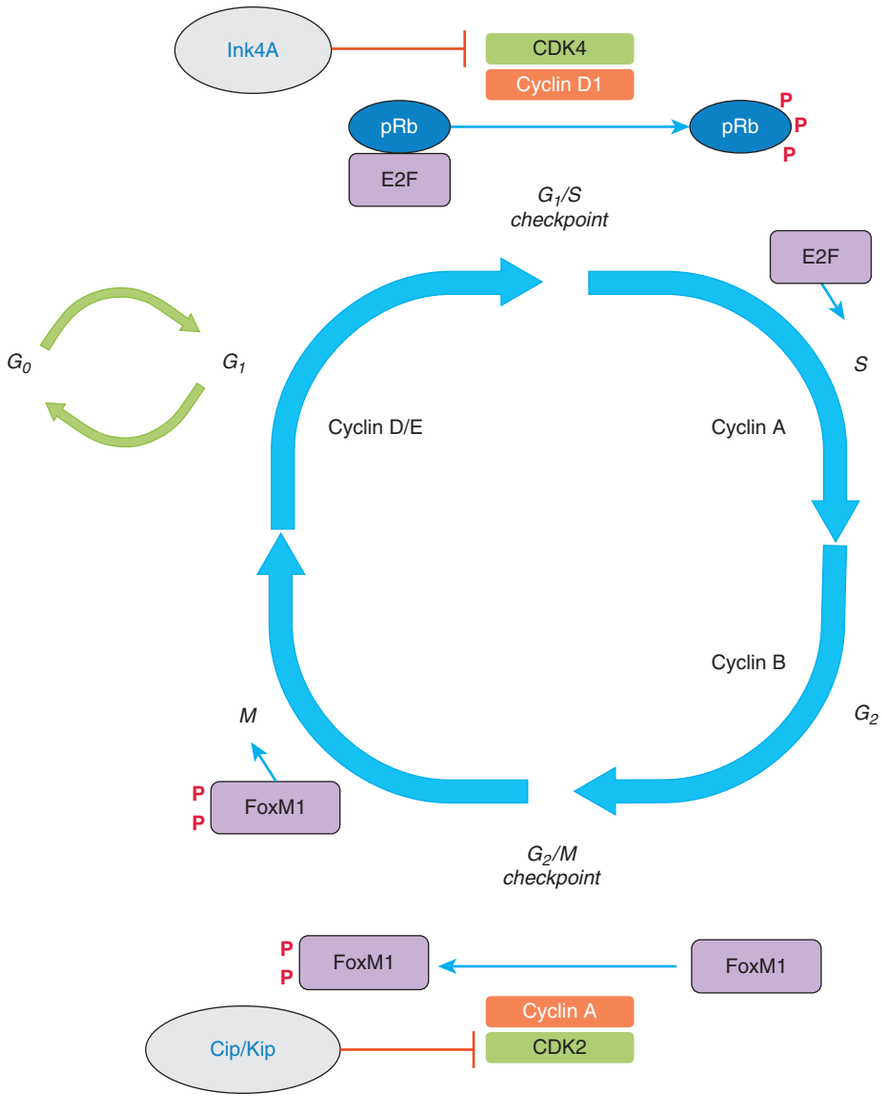

Fig. 1.1 Regulation of the cell cycle by (cycs), cyclin-dependent kinases (cdks), and cdk inhibitors. In the normal cell cycle, DNA synthesis (in which chromosomal DNA is duplicated) occurs in the S phase, whereas mitosis (in which nuclei first divide to form a pair of new nuclei, followed by actual cellular division to form a pair of daughter cells) takes place in the M phase. The S and M phases are separated by two gap phases: the G_1 phase after mitosis and before DNA synthesis, and the G_2 phase following the S phase. During these gap phases, the cell is synthesizing proteins and metabolites, increasing its mass, and preparing for the S phase and M phase. Cell cycle progression is regulated primarily at two points, the *G2/M* and *G1/S* checkpoints, through the coordinated activities of cyclins and CDKs, which in turn are negatively regulated by CDK inhibitors (*Ink4* and *Cip/Kip* families).

The cell cycle is also regulated by multiple CDK inhibitors, which are classified into various classes and are referred by multiple names.[5] CDK4 and CDK6 are inhibited by members of the Ink4 family of inhibitors known as $p16^{INK4a}$ (encoded by the *Cdkn2a* gene), $p15^{INK4b}$ (*Cdkn2b*), $p18^{INK4c}$ (*Cdkn2c*), and p19INK4d (*Cdkn2d*)].[6] Thus these factors also impinge on Cyclin D1/CDK4 regulation of pRb, and consequent E2F activity and S phase entry. $p16^{INK4A}$ loss in cancer results in greater activation of CDK4 and is frequently inactivated in GI cancers, a finding consistent with its function as a tumor suppressor gene.[7,8] Members of the Cip/Kip family of CDK inhibitors are known as $p21^{Cip1}$ (*Cdkn1a*), $p27^{Kip1}$ (*Cdkn1b*), and $p57^{Kip2}$ (*Cdkn1c*)] and are more promiscuous and interfere with multiple cyclin/CDK complexes, including CDK2.

Apoptosis

Apoptosis is a form of programmed cell death that is genetically programmed and executed by specific proteases known as caspases.[9] Similar to other protease cascades, such as the coagulation system, caspases become active upon cleavage of an inactive pro-form, typically through the action of another caspase or as a result of focal accumulation of inactive caspases. Apoptosis is an important mechanism that counterbalances cell proliferation; thus, escape from normal apoptotic mechanisms plays a critical role in oncogenesis. Morphologically, apoptosis is characterized by distinctive features that include chromatin compaction, condensation of the cytoplasm, nuclear fragmentation, and marked alterations at the plasma membrane, resulting

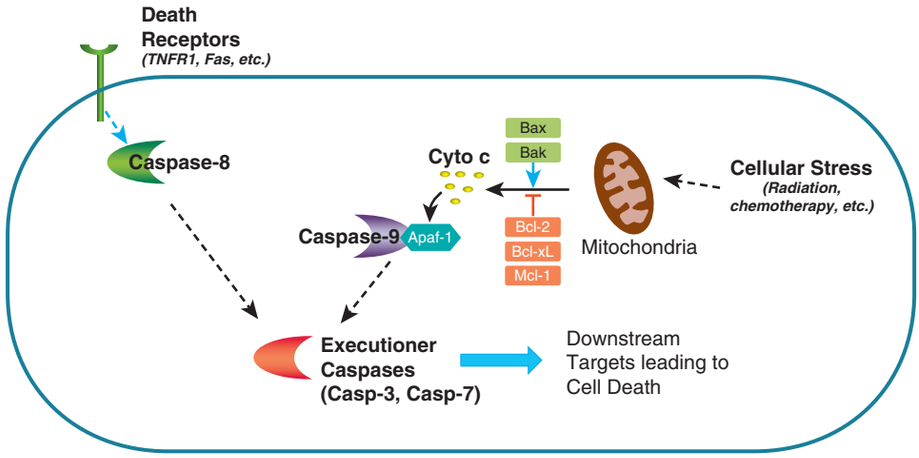

Fig. 1.2 Apoptosis (programmed cell death) counterbalances cellular proliferation to regulate overall tissue growth. A complex interplay of proapoptotic and antiapoptotic molecules results in downstream activation of caspases that mediate cell death. Some of these signals are initiated through cellular stress that can destabilize mitochondrial membranes, and some are initiated through death receptors, including *TNFR1* and *Fas*. The mitochondrial step is regulated by the interplay between proapoptotic *(Bax, Bak)* and antiapoptotic *(Bcl-2, Bcl-xL)* molecules. Upon mitochondrial permeabilization, cytochrome c release promotes the formation of the apoptosome complex (*APAF1*, *caspase 9*, and *cytochrome c*). Activation of caspase-8 (downstream of death receptor) or of caspase-9 (as a result of apoptosome formation), leads to activation of executioner caspases (3 and 7) which are responsible for targeting downstream targets that are responsible for cell death.

in compacted apoptotic bodies that are eventually phagocytosed and eliminated.

Apoptosis may be triggered by internal or external stimuli. Internal stimuli of apoptosis may include nutrient deprivation, hypoxia, DNA damage, or other stressors, including specific toxins, chemical signals, and pathogens. Apoptosis routinely occurs during normal development to facilitate tissue patterning. Similarly, a number of stress situations, including tissue inflammation, can trigger apoptosis. Apoptosis may also be stimulated by specific cell surface receptors belonging to the tumor necrosis factor receptor superfamily, including tumor necrosis factor R1 and Fas, which are referred to as death receptors (Fig. 1.2).

At the intracellular level, the last common event in all forms of apoptosis is the activation of so-called executioner caspases, caspase 3 and 7, which mediate the cleavage of a large number of downstream targets that eventually precipitate cell death. Proapoptotic signals frequently converge at the level of the mitochondria, where they destabilize the mitochondrial membrane and collapse the electrical gradient required for aerobic respiration (see Fig. 1.2). Besides the effects that result in cellular energetics, this process leads to the release into the cytosol of proteins normally present in the intermembrane space of the mitochondria, including cytochrome c, a component of the respiratory chain. In the cytosol, cytochrome c helps in the assembly of a multiprotein complex known as the apoptosome, which contains Apaf1 and facilitates the activation of caspase 9, which can directly activate caspases 3 and 7. On the other hand, death receptors activate executioner caspases through receptor initiated intracellular signaling events that result in the upstream activation of caspase 8.

The mitochondrial membrane permeabilization events that lead to apoptosome formation are controlled by proteins of the Bcl-2 family. On the one hand, Bax and Bak help form the pore, whereas Bcl-2, Bcl-xL, and Mcl-1 inhibit pore formation. The stoichiometric ratio between proapoptotic and antiapoptotic members of the Bcl-2 family can determine the balance between cell survival and cell death.[10] In cancer, alterations in the balance of proapoptotic and antiapoptotic factors, including member of the Bcl-2 family, are common events.

Senescence

Senescence is the process by which cells permanently lose their ability to divide. Senescence may occur in response to the stress induced by activation of oncogenes or DNA damage or after a fixed number of cellular divisions (replicative senescence). Associated with the exit from the cell cycle, senescence is associated with a secretory phenotype that includes a variety of proinflammatory factors. As a physiologic event, senescence limits dysregulated or excessive proliferation. However, when dysregulated, senescence can also contribute to aging and depletion of stem cells.[11] During carcinogenesis, senescence is frequently bypassed or lost.

Replicative senescence is triggered shortening of telomeres, repetitive sequences at the end of chromosomes that protect genomic integrity. Telomeres shorten with each cell division, and when they reach a critically short length, they initiate DNA damage signaling and cellular senescence. This phenomenon can be routinely seen in vitro when primary cells undergo repeated rounds of replication, eventually acquiring critically short telomeres.[12] To prevent senescence from being triggered by sustained replication, cancer cells activate the telomerase enzyme, which adds additional telomeres to the end of chromosomes.[13]

Signaling Pathways That Regulate Cellular Growth

Cellular proliferation is achieved through transition of cells from G_0 arrest into the active cell cycle (see Fig. 1.1). Although progression through the cell cycle is controlled by the regulatory mechanisms just described, overall proliferation is also modulated by external stimuli. Growth factors that bind to specific transmembrane receptors on the cell surface are especially important. Also acting through transmembrane cell surface receptors, extracellular matrix and cell-cell adhesion molecules (i.e., integrins, cadherins, selectins, proteoglycans) can also have a significant impact on cell proliferation. Alterations in cell-matrix or cell-cell interactions are particularly important in contributing to the invasive phenotype of malignant cells.

After ligand binding, the cytoplasmic tails of these transmembrane receptor proteins activate intracellular signaling cascades

that alter gene transcription and protein expression. Based on the nature of the intracellular signaling cascades that these receptors initiate, they can be classified into three major categories: (1) tyrosine kinases, (2) serine and threonine kinases, and (3) G protein–coupled receptors (GPCRs).

The receptors for many peptide growth factors contain intrinsic tyrosine kinase activity within their intracellular tail. After ligand binding, tyrosine kinase activity is stimulated, leading to phosphorylation of tyrosine residues in target proteins within the cell. Most receptors also autophosphorylate tyrosine residues present in the receptors themselves to magnify signaling, and, in some cases, this also causes attenuation of their own activity to effect an intramolecular feedback regulatory mechanism. The receptors for many peptide growth factors, including the receptor for EGF and related growth factors, belong to this receptor class.

Other receptors on the cell surface possess kinase activity directed toward serine or threonine residues rather than tyrosine. These receptors also phosphorylate a variety of cellular proteins, leading to a cascade of biological responses. Multiple sites of serine and threonine phosphorylation are present on many growth factor receptors, including the tyrosine kinase receptors, suggesting the existence of significant interactions among various receptors present on a single cell.[14] The transforming growth factor (TGF)-α receptor complex is one important example of a serine-threonine kinase–containing transmembrane receptor.

Many receptors are members of the so-called 7-membrane–spanning receptor family. These receptors are coupled to guanine nucleotide binding proteins, also known as G proteins, and thus, the receptors are referred to as G protein–coupled receptors. G proteins undergo a conformational change that is dependent on the presence of guanosine phosphates.[15] Activation of G proteins can trigger a variety of intracellular signals, including stimulation of phospholipase C and the generation of phosphoinositides (most importantly, inositol 1,4,5-triphosphate) and diacylglycerol through hydrolysis of membrane phospholipids, as well as modulation of the second messengers cyclic adenosine monophosphate and guanosine monophosphate.[16] Somatostatin receptors exemplify a GPCR prevalent in the GI tract.

Binding of growth factors and cytokines to cell surface receptors typically produces alterations in a variety of cellular functions that influence growth. These functions include ion transport, nutrient uptake, and protein synthesis. However, the ligand-receptor interaction must ultimately modify one or more of the homeostatic mechanisms discussed to affect cellular proliferation.

The Wnt pathway is one important example of a signaling pathway that regulates a diverse number of homeostatic mechanisms to control proliferation of intestinal epithelial cells (Fig. 1.3). Evolutionarily conserved among several species, Wnt signaling, as a rule, regulates proliferation in the stem cell niche and is essential for epithelial homeostasis in the GI tract. From a

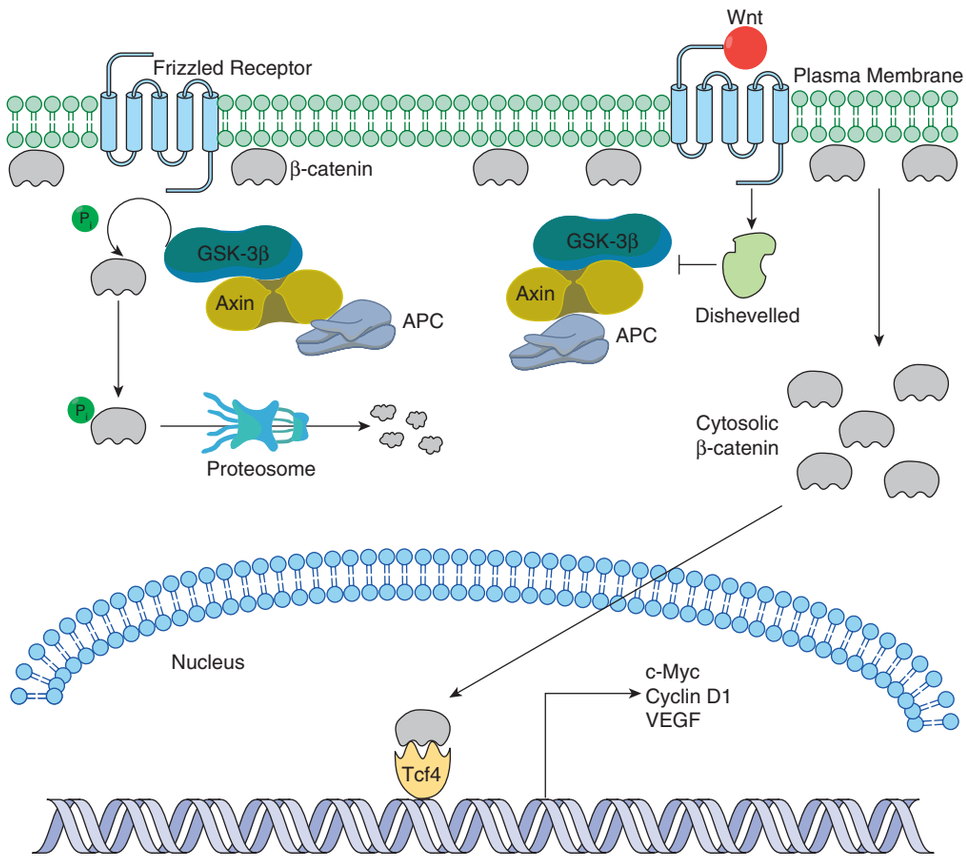

Fig. 1.3 The Wnt signaling pathway is an important regulator of intestinal epithelial cell proliferation and tumorigenesis. In the absence of a Wnt signal (*left top*), cytosolic β-catenin is regulated by the destruction complex, consisting of *APC*, *Axin*, and glycogen synthase kinase-3β (*GSK-3β*). The destruction complex phosphorylates α-catenin and targets it for degradation via the ubiquitin-proteosome pathway. In the presence of an active Wnt signal (*right top*), α-catenin degradation is prevented and the protein is stabilized, leading to excess cytoplasmic α-catenin which is translocated to the nucleus. Nuclear α-catenin interacts with the Tcf-4 transcription factor to regulate the expression of many key target genes. *APC*, Adenomatous polyposis coli; *P*, phosphate group; *Ub*, ubiquitin; *VEGF*, vascular endothelial growth factor.

signaling perspective, its actions are largely the result of the accumulation of α-catenin in the nucleus, where it binds with the transcription factor Tcf-4 to activate a set of target genes.[17] In normal cells, α-catenin is largely associated with adherens junctions, and the cytoplasmic pool of this protein is rapidly degraded through a phosphorylation and ubiquitination pathway. This is mediated by the so-called destruction complex, which includes the tumor suppressor APC. When secreted Wnt ligands bind to cell surface receptors of the Frizzled family, the constitutive degradation of α-catenin is inhibited (disheveled) which results in the nuclear accumulation of this factor, and the subsequent transcriptional activation of genes that promote cell proliferation. Inhibition of the Wnt signal in mice can be achieved by deletion of Tcf-4 or overexpression of the Wnt inhibitor Dickkopf1, which results in dramatic hypoproliferation of the intestinal epithelium.[18,19] Wnt signaling is most active in the base of the crypt, and as differentiation ensues, tissue homeostasis is maintained by growth-inhibiting signals that counterbalance proliferative signals and promote differentiation, including members of the TGF-α family such as BMP4.[20] Specific members of this family have unique functions is tissue homeostasis, including promoting a differentiated and fibrogenic phenotype of mesenchymal cells, induction of specific T cell subtypes, and myriad other activities. In broad terms, the effects of TGF-α family members are mediated intracellularly through the Smad family of proteins, which are transcription factors that are activated in response to ligand-receptor binding.[21] TGF-α induces transcription of the cell cycle inhibitors $p15^{INK4B}$ and $p21^{CIP1/WAF1}$ and is a potent growth-inhibiting factor that mediates arrest of the cell cycle at the G_1 phase. Furthermore, it also enhances the inhibitory activity of $p27^{KIP1}$ on the cyclin E/CDK2 complex.[22]

INTESTINAL TUMOR DEVELOPMENT

Multistep Formation

Multiple sequential genetic alterations are required for the transformation of normal intestinal epithelium to neoplasia. This multistep nature of tumorigenesis is most directly illustrated by the changes that accrue in the development of colonic neoplasia (see Chapter 127). The progression from normal epithelium through adenomatous polyps to malignant neoplasia is paralleled by the accumulation of genetic alterations that change key pathways that control proliferation and tissue homeostasis. Studies on the molecular pathogenesis of colon cancer have served as a paradigm for the elucidation of genetic alterations in other GI cancers, including gastric and pancreatic cancer.

Genomic instability is observed in almost all cancers in the GI tract. This genetically unstable environment promotes the accumulation of the multiple alterations that characterize GI cancers. Instability of the genome may result from several mechanisms, including changes in the genome DNA sequence or through modifications of the nucleotides to alter their functionality, a process called epigenetic change. In colon cancer, there are now 3 well-recognized forms of genetic/epigenetic instability that promote carcinogenesis (Fig. 1.4), and they have been termed *chromosomal instability*, *microsatellite instability* (MSI), and *CpG island methylator phenotype* (CIMP).[23,24] Chromosomal instability refers to alterations in chromosomal structure resulting in large chromosomal deletions, duplications, and translocations, which in aggregate result in a state of aneuploidy. In contrast, MSI refers to frequent alterations in tracts of repetitive DNA sequences (referred to microsatellite DNA) and are often diploid or near-diploid on a chromosomal level (see later discussion on DNA repair). CIMP refers to the accumulation of an epigenetic modification, methylation of guanine residues in so-called CpG-islands, areas rich in cytidine and guanine in gene promoter sites. This modification has a potent effect on gene transcription and results in gene silencing. Other forms of epigenetic change involve the chemical modification of the histone proteins that are required for the assembly of the nucleosome and that control chromatin compaction and DNA access. Although mutations in histones themselves are rare in cancer, mutations in the enzymes that modify histones are emerging as an important group of tumor-associated mutations. It is important to note that involvement by these pathways is not mutually exclusive.

Clonal Expansion

Clonal expansion is essential to tumor development.[25] The acquisition of a mutation that may provide a growth or survival advantage to a cell is followed by clonal expansion of these mutated cells. As this population grows, and particularly with the acquisition of genetic/epigenetic instability, a second round of clonal expansion occurs as a cell within this population sustains still another genetic alteration that further enhances its growth properties. This iterative process of selection, with accumulating genetic alterations, results in malignancy. Because of the nature of the clonal expansion process, once frank malignancy has developed, it is often the case that multiple clones are present in the same tumor, with a different catalog of mutations harbored among various cancer cells. Referred to as *tumor heterogeneity*, this ongoing process may give certain cells selection advantages.[26] Metastasis may be facilitated by the evolution of a subset of tumor cells that acquire the capability of traversing the circulatory system and thriving in a new environment.

Cancer Stem Cells

Recognition of tumor heterogeneity has led to the *cancer stem cell (CSC) hypothesis*, which asserts that there exists a subset of tumor cells that have stem cell–like properties. CSCs are believed to be the tumor-initiating cells from which clonal expansion occurs. Moreover, it is hypothesized that eradication of these cells is a key therapeutic goal because failure to do so may result in relapse of disease. Within this CSC hypothesis, there are 2 models.[27] The first is a hierarchical model in which CSCs serve as progenitors for all cells in in a given tumor, whereas other cells have limited long-term reproductive potential. The basic evidence for this model is the finding that only cells with specific surface markers can repopulate the tumor in xenotransplantation experiments. In the GI tract, analysis of putative CSCs demonstrate transcriptional programs and markers shared with normal intestinal stem cells, such as Lgr5 and EphB2, which identify and purify colon CSCs.[28] The second stochastic model posits that each cancer cell has the same potential to be a CSC, but this determination is stochastically based on internal factors in addition to external environmental cues.

Epithelial-Mesenchymal Transition

It has been noted that within tumors of epithelial origin, some cells acquire features of mesenchymal cells. A similar process occurs during normal embryogenesis, when polarized epithelial cells no longer recognize the boundaries imposed by adjacent epithelial cells or their basement membrane and adopt features of migratory mesenchymal cells. This phenomenon, designated *epithelial-mesenchymal transition* (EMT), endows cells with the ability to move through tissue planes that normally serve as boundaries for epithelial cells, such as the basement membrane, a dense matrix of collagen, glycoproteins, and proteoglycans. The transmigration of tumor cells through the basement membrane likely involves production of key proteolytic activities. Alternatively, the tumor cell may produce factors capable of activating proenzymes present in the extracellular matrix. For example, the tumor may produce urokinase, itself a protease, or plasminogen

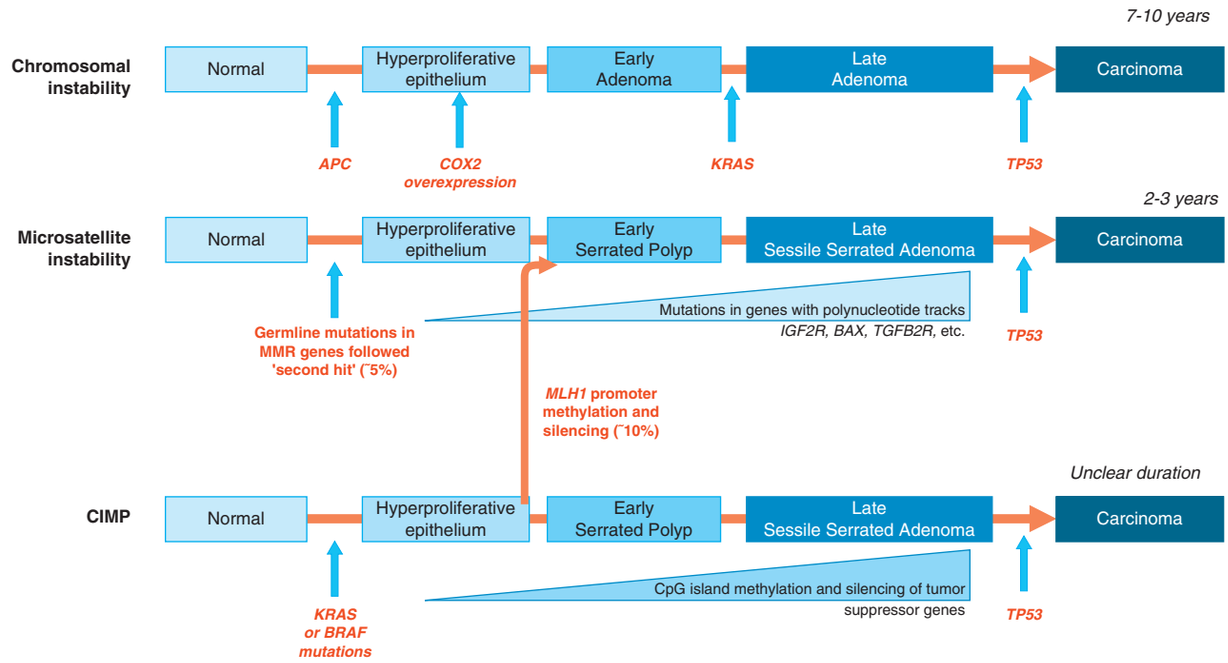

Fig. 1.4 Multistep models of colorectal cancer based on underlying genetic instability. As shown on the left, there are 3 major pathways: chromosomal instability (*top pathway*), microsatellite instability (*middle pathway*), and the CpG island methylation, or *CIMP* (*lower pathway*). The progression from normal colonic epithelium to carcinoma is associated with the acquisition of several genetic and epigenetic alterations. In the chromosomal instability pathway (*top pathway*), these alterations include the early loss of *APC*, followed by activation of oncogenes (e.g., *KRAS*) through a point mutation and inactivation of tumor suppressor genes (e.g., *APC, TP53*) through a point mutation or deletion. An increasing aggregate number of mutations can be correlated with progression from early benign adenoma to cancer, as reflected by analysis of polyps by size. In the microsatellite instability model (*middle pathway*), mutations in DNA mismatch repair (*MMR*) genes create a mutator phenotype in which mutations accumulate in specific target genes (see section on DNA mismatch repair). Tumors develop much more rapidly through this pathway than through the chromosomal instability pathway (2-3 years compared to 7-10 years). Germline mutations in *MMR* genes account for 5% of all colorectal tumors. In the *CIMP* pathway (*lower pathway*), the initiating event is hypothesized to be a *BRAF* or *KRAS* activating mutation that somehow triggers extensive CpG island methylation, particularly of gene promoters, resulting in gene silencing. Among the potential gene targets is *MLH1*, a component of the MMR pathway, and when silenced as part of the *CIMP* pathway, the tumor evolves along a similar molecular as microsatellite unstable cancers (MSI-H). Sporadic *MLH1* methylation and silencing accounts for nearly 10% of sporadic colorectal cancers. Alternatively, serrated adenomas arising in the *CIMP* pathway can undergo a pathway similar to that of chromosomal instability to become microsatellite stable tumors.

activator. Having gained access to the interstitial stromal compartment, tumor cells can then enter lymphatic and blood vessels and metastasize.

In addition to these properties, it has been recognized that cells that undergo EMT acquire not only invasive features but also CSC-like features.[29]

One key feature of EMT is the loss of adherens junctions that normally maintain epithelial cell–cell interactions. The molecular correlate of this phenomenon is the loss of expression of E-cadherin, a critical component of the adherens junction.[30] Mutations in E-cadherin are common in many GI cancers, particularly gastric cancer, where germline mutations in E-cadherin are also linked to hereditary diffuse gastric cancer.

NEOPLASIA-ASSOCIATED GENES

Genes that become altered during the neoplastic process belong to two distinct groups: (1) oncogenes, which actively confer a growth-promoting property, or (2) tumor suppressor genes, the products of which normally restrain growth or proliferation. An important category within tumor suppressor genes includes DNA repair genes, which prevent accumulation of new mutations. Activation of oncogenes or inactivation of tumor suppressor genes contributes to malignant transformation. Although most of these genes encode for proteins, many cancer-promoting genes that harbor oncogenic and tumor suppressive functions do not encode for proteins but rather for RNAs that modulate genomic function, so-called noncoding RNAs.

Oncogenes

According to the Catalog of Somatic Mutations in Cancer (COSMIC),[31] there are close to 80 oncogenes with strong evidence of involvement in cancer. Genes that encode a normal cellular protein, whose function may promote the neoplastic process (e.g., antiapoptotic function, cell proliferation stimulation, etc.), may function as oncogenes when they are expressed at inappropriately high levels. A typical mechanism for this phenomenon is gene amplification, when tumors acquire multiple copies of a normal gene resulting in a dosage effect that leads to increased gene expression.

In other cases, a variety of mutations may lead to inappropriately high activity of a normal gene, leading to cancer-promoting

activities. Point mutations or large gene rearrangements resulting in fusion proteins are examples of mutations that can lead to oncogene activation. For example, several genes that encode tyrosine kinase–containing growth factor receptors become oncogenes after a mutation results in unregulated tyrosine kinase activity that is no longer dependent on the presence of the appropriate ligand (e.g., EGF). Because of their tumor-promoting activity, these mutations tend to be recurrent among specific cancer classes. The normal cellular genes from which the oncogenes derive are designated *proto-oncogenes*. Most of these genes are widely expressed in many different types of tumor cells.

Finally, another source of oncogenes are virally encoded proteins that may affect cellular growth or survival.[32] These factors, while evolved to favor the viral cycle, may in some instances favor neoplastic development and this is the reason why specific viruses are associated with increased cancer risk. In addition, in the case of retroviruses, the ability of the viral genome to insert itself in the genome of the host can lead to disruptions in the expression of genes in the vicinity of insertion sites, which at times, may have oncogenic activities.

The proteins encoded by oncogenes may affect any of the hallmarks of cancer, such as stimulate growth factor pathways, promote tumor invasion, prevent cell death, or have other tumor-promoting actions. With regards to promoting growth factor pathways, oncogenes may encode for (1) growth factors or their receptors, or for (2) intracellular signal transduction molecules downstream of the receptor itself, including transcription factors that mediate the actions of the growth factor at the level of the nucleus.

Oncogenic Growth Factors and Growth Factor Receptors

The transforming effects of enhanced expression of a variety of growth factors have been demonstrated both in vitro and in vivo. Several growth factor–related proteins encoded by oncogenes have now been recognized, including the family of Wnt and Sis proteins, which encodes the α chain of platelet-derived growth factor. Cancer cells may engage in autocrine signaling to promote their growth, or coax the adjacent stroma to hypersecrete such growth-stimulating factors. More frequently, a variety of receptors are upregulated in expression or dysregulated leading to constitutive action. Among them, are receptor tyrosine kinases of the EGF receptor family (ERBB1-4), which are frequently upregulated in a variety of GI cancers.

Signal Transduction–Related Oncogenes

Intermediate steps that effectively translate ligand-receptor binding to an intracellular signal are essential in mediating functional responses of the cell. Mutations in genes that encode key proteins that participate in signal transduction can also lead to cellular transformation (Fig. 1.5). In this regard, the largest family of oncogenes encodes proteins with protein kinase activity. Many members of this large oncogene group are expressed by neoplasms of the GI tract, and these include the Src nonreceptor tyrosine kinase that associates with the inner surface of the plasma membrane.

G proteins regulate signaling of the large family of GPCRs through the exchange of guanosine triphosphate with guanosine diphosphate. In this regard, the *ras* family of genes, which encodes a family of proteins related to the G proteins, are among the most commonly detected oncogenes in GI tract cancers. The *ras* family contains 3 genes: *H-ras*, *K-ras*, and *N-ras*. These factors are essential to transduce signals from various growth receptor signaling cascades and point mutations that result in activating amino acid substitutions at critical hot spot positions convert the normal gene into an oncogene.

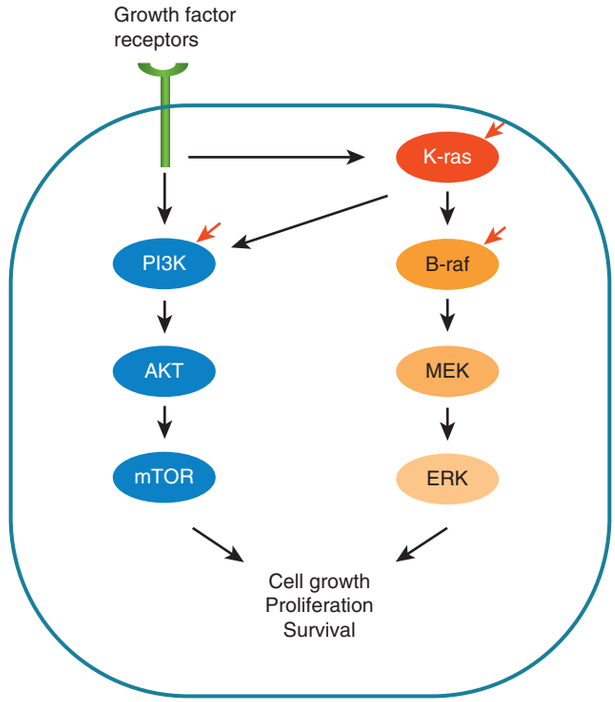

Fig. 1.5 Signal transduction downstream of growth factor receptors, where *K-ras* plays a major role. Oncogenic *K-ras* can activate multiple signaling pathways. Molecules that are frequently mutated in colorectal cancer are noted by a red arrow and include *K-ras* (40%), *B-raf* (10%), and PI3K (15%). *AKT*, Cellular homolog of v-Akt oncogene; *ERK*, extracellular signal regulated kinase; *MEK*, MAPK/ERK kinase; *mTOR*, mammalian target of rapamycin; *PI3K*; phosphoinositide-3 kinase.

To date, almost all *ras* mutations in GI malignancies occur in the *K-ras* oncogene. The highest mutation frequency is found in tumors of the exocrine pancreas (>90%).[33] *Ras* genes activated through point mutation have been identified in approximately 50% of colonic cancers as well as a subset of serrated tumors (see Fig. 1.4).[34]

Most oncogenic mutations in *ras* cause biochemical changes that maintain it in the active, guanosine triphosphate–bound state by reducing guanosine triphosphatase activity or by destabilizing the inactive guanosine diphosphate–bound form. However, several *ras* mutants retain significant guanosine triphosphatase activity; therefore, other mechanisms that convert *ras* to a transforming protein may be involved.[35]

A functional consequence of ras activation is the phosphorylation and activation of key downstream serine/threonine kinases. One important target of ras is B-raf. In colon cancers without an identifiable *K-ras* mutation, 20% possess an activating *B-raf* mutation,[36] consistent with the concept that activation of an oncogenic pathway can be achieved through an alteration in any of several sequential components of a particular pathway (see Fig. 1.5).

Nuclear Oncogenes

Many cellular oncogenes encode proteins that localize to the nucleus. In essence, these nuclear oncogene products are the final mediators of signal transduction pathways that are also affected by cytoplasmic and plasma membrane–bound oncoproteins, because they act as transcription factors that regulate expression of certain genes that enhance cellular proliferation and suppress normal differentiation.

The role of nuclear oncogenes is illustrated by the myc family. The c-Myc protein product is involved in critical cellular functions like proliferation, differentiation, apoptosis, transformation, and transcriptional activation of key genes.[37] Frequently, c-Myc is overexpressed or amplified in many GI cancers. c-Myc has been found to be a transcriptional target of the α-catenin/TCF-4 complex in colorectal cancers (see Fig. 1.3), which may explain the overexpression of c-Myc observed in this cancer type.[38]

Tumor Suppressor Genes

Mutations of tumor suppressor genes are associated with all GI cancers, and a number of these genes and their products have been identified and characterized (Table 1.1). Unlike gain-of-function mutations, which are characteristic of oncogenes, mutations in tumor suppressor genes are loss-of-function mutations and are therefore biallelic.

Initial recognition of the existence of tumor suppressor genes was derived from genetic analyses of cancer-prone families. In the GI tract, hereditary colon cancer, gastric cancer, and pancreatic cancer syndromes are the best described and are discussed elsewhere in this text (see Chapters 54, 60, and 127). In these syndromes, there is a marked increase in risk for a particular tumor in the absence of other predisposing environmental factors. Tumors arise typically at a younger age than they do in the general population, and multiple primary tumors may develop within the target tissue.

From a genetic standpoint, cancer genetic syndromes most often have an autosomal dominant mode of mendelian inheritance. Based on observations in hereditary retinoblastoma, Knudson proposed the "2-hit" hypothesis,[39] which explains the relationship between sporadic and familial forms of cancer. Whereas sporadic tumors are initiated by somatic biallelic inactivating mutations of a tumor suppressor gene, tumors in familial cancer syndromes are accelerated by the inheritance of a monoallelic mutation of a tumor suppressor gene present in all cells in affected family members. When this germline mutation is followed by a somatic mutation in the remaining normal allele of the tumor suppressor gene, this gives rise to the development of a neoplastic clone that eventually gives rise to a tumor (Fig. 1.6). Because of the germline mutation, the likelihood of full inactivation of the tumor suppressor is diminished substantially because only one additional hit is required, leading to the younger age of onset and the potential for tumor multiplicity that accompanies these syndromes.

Although this 2-hit model has been generally observed, there are exceptions. Some tumor suppressors may function to increase cancer risk when only one allele is mutated. Moreover, some cancer genetic syndromes display somatic recessive mode of inheritance because genetic risk is conferred only when biallelic inactivating mutations are present. Another important feature of tumor suppressor genes is that they do not function identically in every tissue type. Consequently, inactivation of a particular tumor suppressor gene is tumorigenic only in certain tissues. For example, the tumor suppressor genes *RB1* and *VHL* play crucial roles in retinoblastomas and renal cell cancer, respectively, but are rarely mutated in GI malignancies. Tumor suppressor genes shown to have a critical role in the pathogenesis of GI malignancies, *APC*, *TP53*, and *SMAD4*, are described later. Furthermore, we will discuss DNA repair pathways that, when lost, can give rise to neoplasia and therefore function as tumor suppressor factors.

Adenomatous Polyposis Coli Gene

Genetic linkage analysis revealed markers on chromosome 5q21 that were tightly linked to polyp development in affected members of kindreds with familial adenomatous polyposis (FAP) and Gardner's syndrome.[40] Further work led to identification of the gene responsible for FAP, the *APC* gene.[41-43] The full spectrum of adenomatous polyposis syndromes attributable to *APC* is discussed in detail in Chapter 126. Somatic mutations in *APC* have also been found in most sporadic colon polyps and cancers.[44,45] Mutations in *APC* are characteristically identified in the earliest adenomas, indicating that *APC* plays a critical role as the gatekeeper in the multistep progression from normal epithelial cell to colon cancer (see Fig. 1.4).

The *APC* gene comprises 15 exons and encodes a predicted protein of 2843 amino acids, or approximately 310 kDa. Most germline and somatic *APC* gene mutations result in a premature stop codon and therefore a truncated *APC* protein product and loss of function. As discussed earlier, *APC* is a negative regulator of the Wnt signaling pathway and its inactivation results in a state that resembles constitutive activation of Wnt. Intracellularly,

TABLE 1.1 Mutations Associated with Hereditary Gastrointestinal Cancer Syndromes

Disorder	Gene(s) Mutated
FAP, AFAP, Gardner syndrome	*APC*
Lynch syndrome (HNPCC)	*MLH1, MSH2, PMS2, MSH6, EPCAM* (through disruption of the neighboring MSH2 gene)
MAP	*MUTYH*
Peutz-Jeghers syndrome	*STK11*
Cowden's disease	*PTEN*
Juvenile polyposis	*SMAD4, BMPR1A*
Hereditary diffuse gastric cancer	*CDH1*
Hereditary pancreatic cancer	*ATM, BRCA1, BRCA2, PALB2, PALLD, CDKN2A, PRSS1, SPINK1, PRSS2, CTRC, CFTR*
MEN1	*Menin*

AFAP, Attenuated FAP; *APC*, adenomatous polyposis coli; *FAP*, familial adenomatous polyposis; *HNPCC*, hereditary nonpolyposis colorectal cancer; *MAP*, MUTYH-associated polyposis; *MEN1*, multiple endocrine neoplasia, type 1; *MUTYH*, mutY homolog.

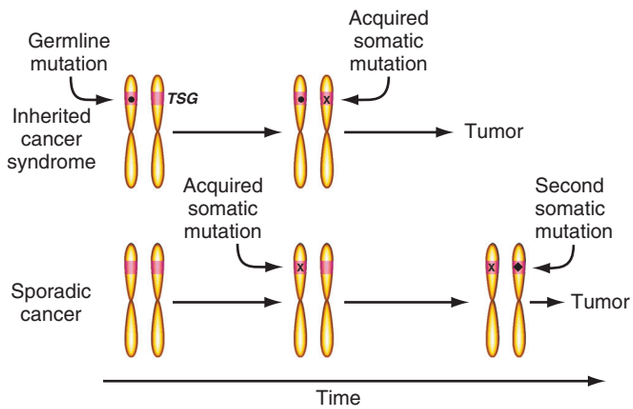

Fig. 1.6 Knudson's 2-hit hypothesis. In an inherited cancer syndrome, one chromosome has an inactive tumor suppressor gene (TSG) locus because of a germline mutation. The counterpart TSG on the remaining paired chromosome is subsequently inactivated by a somatic mutation, leading to tumor formation. In contrast, in a sporadic cancer, the two alleles of the TSG need to become inactivated through two independent somatic mutations, an event that is less likely to occur within a single cell.

this is manifested by stabilization of α-catenin, which mediates the transcriptional effects of Wnt activation and the subsequent oncogenic phenotype (see Fig. 1.3). Interestingly, another mechanism to achieve this signaling outcome are mutations in α-catenin itself that render the protein impervious to *APC*-dependent degradation.

TP53 Gene

This is the most commonly mutated gene in human cancer,[46] and point mutations in *TP53* are found with high frequency in all cancers of the GI tract.[47] In fact, point mutations in *TP53* have been identified in as many as 50% to 70% of sporadic colon cancers (see Fig. 1.4). Interestingly, these mutations arise relatively late in the oncogenic process as the gene is mutated in only a small subset of colonic adenomas.[48]

Named for a 53-kDa-sized gene product, p53 is a nuclear phosphoprotein that plays a key role in cell cycle regulation and apoptosis.[47] In the nucleus, p53 functions as a transcription factor which can be induced by conditions of cellular stress, such as ionizing radiation, growth factor withdrawal, or cytotoxic therapy. Induction of p53 arrests cells at the G_1 phase to facilitate DNA repair, senescence, or trigger apoptosis. These responses are mediated in part by its transcriptional targets such as the $p21^{CIP1/WAF1}$ inhibitor of the cell cycle or the proapoptotic gene, *PUMA*.[49] Interestingly, it is often the case that TP53 mutations occur as the combination of a genomic deletion encompassing one allele, together with a missense mutation in the second allele that targets specific hotspots within the protein. Recent evidence indicates that the genomic deletions function not only by removing *TP53* but through the loss of adjacent genes with tumor suppressive activities.[50] Furthermore, the second type of mutations, resulting in specific missense mutations are thought to contribute gain-of-function tumorigenic activities.[51] In addition to the *TP53* point mutations in sporadic cancers, germline *TP53* mutations have been observed in the Li-Fraumeni syndrome, an autosomal dominant familial disorder in which breast carcinoma, soft tissue sarcoma, osteosarcoma, leukemia, brain tumor, colon cancer, and adrenocortical carcinoma can develop in affected persons.[52]

SMAD4 Gene

SMAD4 is a tumor suppressor gene located on chromosome 18q and is deleted or mutated in most pancreatic adenocarcinomas and a subset of colon cancers. Smad4, the protein encoded by this gene is an essential intracellular mediator of factors belonging to the TGF-α superfamily. Smad4 functions as a transcription factor and is an obligate partner of other members of the Smad protein family.[53] Mutant Smad4 lacks these properties and among other effects, leads to loss of TGF-α inhibition of proliferation. Germline mutations in *SMAD4* result in the juvenile polyposis syndrome (see Chapter 126).

DNA Repair Genes

DNA replication itself and various types of DNA damaging agents can introduce errors into the genome. These errors include spontaneous mismatching of nucleotides during normal DNA replication, oxidative damage of nucleotides, and complete double-strand breaks. Therefore, a variety of cellular mechanisms have evolved to prevent or correct DNA errors. One type of error that develops during replication may occur in repetitive mononucleotide or dinucleotide stretches of DNA, so-called microsatellite regions.[54] These repetitive regions are prone to DNA mismatches, which if not resolved, can result in short insertions or deletions. The cellular machinery devoted to correct these errors is referred to the mismatch repair system. The enzymes bind mismatched DNA, cut the DNA strand with the mismatched nucleotide, unwind the DNA fragment, fill in the gap with the correct nucleotide, and finally reseal the remaining nick. The family of DNA mismatch repair genes includes two basic molecular components, a mismatch recognition complex composed of MSH2 and MSH6, and an excision inducing complex composed of MLH1 and PMS2. Mutations in any of these genes result in defective mismatch repair, and when inherited due to a germline mutation, they give rise to Lynch syndrome, also known as *hereditary nonpolyposis colorectal cancer*.[55,56] Complete loss of a mismatch repair factor leads to very high rates of DNA mutations, and mismatch repair defective tumors accumulate a high burden of cancer somatic mutations, typically over 2000 somatic mutations, resulting in a large number of tumor-specific neoantigens.[57] Affected cells are called *replication error positive*, in contrast to the replication error–negative phenotype.[58,59] Because microsatellite DNA sequences are primarily affected by this type of genetic instability, the tumor cells display insertions or deletions in these stretches of DNA when compared to nontumor tissue, a phenomenon referred to as microsatellite instability. Mechanistically, the absence of DNA repair does not directly cause cancer but creates a milieu that permits accumulation of mutations in a variety of genes that contain repetitive DNA sequences, such as the TGF-α type II receptor, IGF type II receptor, BAX, and E2F-4, among others.

Loss of mismatch repair genes represents an important mechanism for the accumulation of mutations within a tumor (see Fig. 1.4). While 5% of colon cancer are due to Lynch syndrome, i.e., germline mutations in the mismatch repair system, twice as many tumors (10%) display similar molecular characteristics without a germline mutation in any of the mismatch repair genes. These tumors are most often driven by somatic loss of function in this system, most often as a result of silencing of *MLH1* gene expression as a result of an epigenetic change in the promoter region of this gene called DNA methylation. *MLH1* promoter hypermethylation is most often observed in lesions that are serrated adenomas by histology and that also carry B-Raf mutations (see Fig. 1.4). Finally, it has been recognized that another mechanism that can lead to a state of high mutation burden is the loss of exonuclease proofreading activity of the replicative DNA polymerase Pol-ε or Πολ–δ, through a variety of missense mutations.[60]

Another important DNA repair pathway involved in carcinogenesis is mediated by the MUTYH gene. It encodes a DNA glycosylase that participates in the repair of oxidized guanine nucleotides, such as 8-oxoguanine residues, that may inappropriately pair with adenines, ultimately leading to somatic G:C→T:A mutations if uncorrected. Biallelic mutations in MUTYH results in an adenomatous polyposis syndrome that resembles FAP, except that its mode of inheritance is autosomal recessive (see Chapter 126).[61,62] Interestingly, G:C→T:A mutations in the APC gene were almost universally found in the polyps of patients with germline MUTYH mutations, indicating that there are important similarities in the molecular pathogenesis of polyps in the MUTYH and FAP syndromes.

Noncoding RNAs

Our genomes harbor a variety of genes whose products are RNAs that do not encode for a protein. The RNA products, termed *noncoding RNAs*, consist of a broad category of active RNA molecules that can mediate a variety of effects. The categories of noncoding RNAs are rapidly expanding and include so-called microRNAs and long noncoding RNAs, which are frequently dysregulated in cancers. The microRNAs play a critical role in silencing of other RNA transcripts via RNA degradation or translational inhibition and typically regulate dozens of target RNAs at a time. Their biogenesis involves conventional gene transcription, followed by processing of the resulting RNA by a variety of nuclease cleavage events, resulting ultimately in the generation of small interfering

RNAs (siRNAs) by the protein Dicer. These siRNAs bind to complementary mRNA sequences, and this binding determines the specificity for RNA targets. Long noncoding RNAs may perform diverse functions like gene silencing, splicing, and extension of telomeres.

Oncogenic Signaling Pathways

Individual oncogenes or tumor suppressor genes do not necessarily induce cellular transformation directly but typically function in concert with one another as components of larger oncogenic signaling pathways already discussed. Some of the pathways that are particularly relevant for GI tumorigenesis include the Wnt and Ras signaling pathways. These are pathways that regulate normal tissue homeostasis but become oncogenic when the signals are transduced in an aberrant or amplified manner. The key features of Wnt signaling are illustrated in Fig. 1.3. α-Catenin is translocated from the inner plasma membrane to the cytoplasm. There, it forms a macromolecular complex with the APC protein Axin and glycogen synthase kinase-3α. Phosphorylation of α-catenin by glycogen synthase kinase-3α triggers its degradation. In the presence of an active Wnt signal, α-catenin is stabilized and enters the nucleus, where it interacts with the transcription factor Tcf-4 to up-regulate a number of key target genes, including *c-Myc*, *cyclin D1*, and *vascular endothelial growth factor (VEGF)*. As discussed earlier, Wnt signaling is essential for regulating proliferation of normal intestinal epithelium, and dysregulated Wnt signaling is an almost universal feature of all colorectal cancers. The latter can result from a mutation in the *APC*, *Axin*, or *α-catenin* genes, although alterations in the *APC* tumor suppressor gene are the most common. An alteration in just one of these components is sufficient to activate the entire pathway. Thus, it is essential to consider individual genetic alterations in the context of the overall signaling pathway in which they function.

Because pathways are typically not linear, additional levels of complexity arise. There is frequent overlap among pathways, and the distinction between pathways can be somewhat arbitrary. For example, mutations in the *K-ras* oncogene result in activation of multiple distinct signaling pathways, including Raf/ERK/MAPK, PI3K/Akt, and nuclear factor-κB, all of which play an important role in tumorigenesis (see Fig. 1.5). Crosstalk between these effector pathways serves to modulate the cellular responses further. For example, Akt, a target of PI3K, can phosphorylate Raf and thereby regulate signaling through the MAPK pathway.[63] Finally, each of these signaling pathways regulates multiple biological processes related to tumorigenesis,[64] including cell cycle progression, apoptosis, senescence, angiogenesis, and invasion.

Another pathway that plays a particularly important role in GI tumors is the cyclooxygenase-2 (COX-2) pathway. The enzyme COX-2 is a key regulator of prostaglandin synthesis that is induced in inflammation and neoplasia. Although no mutations of COX-2 have been described, overexpression of COX-2 in colonic adenomas and cancers is associated with tumor progression and angiogenesis (see Fig. 1.4), primarily through induction of prostaglandin E_2 synthesis. Inhibition of COX-2 with a variety of agents (aspirin, nonsteroidal anti-inflammatory drugs, or COX-2 selective inhibitors such as celecoxib) is associated with a reduced risk of colorectal adenomas and cancer.[65]

TUMOR MICROENVIRONMENT

Cancer is ultimately a complex tissue consisting not only of neoplastic cells harboring a number of genetic lesions, as outlined previously, but the composite of a number of cellular components that endow the tumor with all of its properties. Indeed, the contribution of non-neoplastic cells to the behavior and evolution of a tumor is increasingly recognized. Cellular elements with recognized contributions to the behavior of the tumor include its mesenchymal cells, its vasculature, a variety of immune cells recruited to the tumor and particularly in tumors of the intestinal tract, and tumor-associated microbiota which contribute significantly to the tumor microenvironment. In addition, these elements acting in concert lead to a metabolic environment, such as the oxygen and nutrient supply of the tumor, that often plays a significant role in the evolution of the tumor at the primary site and its potential for distant metastasis.

TUMOR METABOLISM

Tumor cells exhibit abnormal metabolic profiles to facilitate their growth and anabolic needs. Observations in 1924 from Nobel Laureate Otto Heinrich Warburg revealed that tumor cells displayed dramatic increases in aerobic glycolysis and diminished mitochondrial respiration. This metabolic state, known as the *Warburg effect*, has been validated and is a hallmark feature of most malignancies.[66] It is becoming increasingly clear that integration of the genetic lesions that characterize cancer formation is responsible for the changes in cellular metabolism that accompany cellular transformation. Many of the genes implicated in GI cancers (*p53*, *K-Ras*, *PI3K*, *mTOR*, *HIF*, *Myc*) can in fact regulate metabolic pathways. Moreover, germline mutations in metabolic regulators (e.g., subunits of succinate dehydrogenase) that are not classical oncogenes or tumor suppressor genes have been associated with a high risk of tumorigenesis (pheochromocytoma and paraganglioma).[67,68] The selection advantage of increased glycolysis in cancer cells may include greater tolerance to hypoxic environments and shunting of metabolic byproducts (e.g., lactate) to other biosynthetic pathways. These altered metabolic pathways are promising new targets for therapy.

Inflammation and Cancer

Immune cells recruited to the tumor microenvironment can result in a variety of effects. On the one hand, tumor immune surveillance is well recognized and immunosuppressed states increase the risk of cancer development. On the other hand, a number of cellular elements of hematopoietic origin can promote primary tumor growth, prevent effective immune surveillance, or promote the acquisition of features of neoplastic cells that facilitate metastasis. Myeloid cells with immature characteristics, so-called myeloid-derived suppressor cells, are an important example of this phenomenon.[69]

In addition, a number of chronic inflammatory conditions increase the site-specific risk of cancer; examples of this include ulcerative colitis (see Chapter 115), chronic gastritis (see Chapter 52), chronic pancreatitis (see Chapter 59), Barrett's esophagus (see Chapter 47), and chronic viral hepatitis (see Chapters 79 and 80). The influences of inflammation on the development of neoplasia are multifaceted and complex. Cytokines produced by inflammatory cells can lead to activation of antiapoptotic and pro-proliferative signals in tumor cells mediated by transcription factors such as nuclear factor-κB and STAT3.[70,71] Immune cells may also promote remodeling of the vascular network and promote angiogenesis (discussed later). Inflammation may also induce DNA damage from cytokine-stimulated production of reactive oxygen species.

Microbiome

The human body possesses over 100 trillion microbes and the largest concentration of these organisms are present in the GI tract. The interaction between these organisms and the host is an area of great interest, particularly for a broad range of autoimmune, metabolic, and neoplastic disorders.[72] Interestingly, colonic tumors are associated with specific subsets of bacteria, and the tumor associated microbial species have the capacity of inducing

colonic tumors in specified animal models.[73] *Fusobacterium nucleatum*, an organism typically found in the oral cavity, is an example of this behavior as it can be found in association with colon tumors and, when introduced into colon cancer models driven by germline *APC* mutations, can drive colon tumorigenesis.[74]

BIOLOGICAL FEATURES OF TUMOR METASTASIS

The establishment of distant metastases requires multiple processes, many of which involve alterations in interactions between tumor cells and normal host cells. To metastasize, a cell or group of cells must detach from the primary tumor, gain access to the lymphatic or vascular space, adhere to the endothelial surface at a distant site, penetrate the vessel wall to invade the second tissue site, and finally proliferate as a second tumor focus. Angiogenesis is necessary for proliferation of the primary tumor and tumor metastases. Tumor cells must also overcome host immune cell killing. As a result, few circulating tumor cells (<0.01%) successfully initiate metastatic foci. A "survival of the fittest" view of metastasis has been proposed, in which selective competition favors metastasis of a subpopulation of cells from the primary site.[75] In favor of this view is the fact that the mutational landscape of the primary and distant tumor sites are often distinct, indicating that only specific tumor clones acquire the ability to metastasize.

Angiogenesis and Lymphangiogenesis

Angiogenesis is essential to sustain continued growth of the primary tumor. If new vessels are not developed as the primary tumor expands, cells most distant from available vessels are deprived of an adequate source of nutrition and oxygen, and central necrosis occurs. Neovascularization is also an important permissive factor in facilitating metastatic dissemination of tumors.[76] A number of protein growth factors produced by malignant tumor cells and stromal cells have been found to be potent stimuli of angiogenesis, including VEGF-A, basic fibroblast growth factor, and TGF-α. VEGF-A is perhaps the most critical factor that is up-regulated in most tumor types, including colorectal cancer. Multiple genetic pathways implicated in GI carcinogenesis modulate *VEGF-A* expression, including Wnt and mutant *ras*.[77]

Angiogenesis occurs in an ordered series of events. Endothelial cells in the parent vessel are stimulated to degrade the endothelial basement membrane, migrate into the perivascular stroma, and initiate a capillary sprout. The sprout develops into a tubular structure that in turn develops into a capillary network. In vitro models that recapitulate the early events of angiogenesis indicate that this process involves a balance between proteases and protease inhibitors in a manner similar to that during tumor invasion. Indeed, functional parallels between tumor invasion and angiogenesis are evident in their mutual requirement for cellular motility, basement membrane proteolysis, and cell growth.

In addition to angiogenesis, lymphangiogenesis plays an important role in tumor metastasis. Some important clues into the molecular basis of tumor lymphangiogenesis have been obtained. VEGF-C or VEGF-D bind to the VEGF receptor-3 on lymphatic endothelial cells to stimulate formation of new lymphatic vessels.[78] This results in the development of new lymphatic channels within the tumor mass and, consequently, enhanced dissemination of tumor cells to regional lymph nodes.[79]

ENVIRONMENTAL INFLUENCES

Fundamentally, cancer is a genetic disorder and genetic mutation is the common denominator of agents or mechanisms that contribute to the development of neoplasia. Environmental factors play an important role in tumorigenesis insofar as they affect the progression of the underlying genetic lesions.

Chemical Carcinogenesis

Many compounds that have carcinogenic potential often require metabolic modification by host enzymes, a process called metabolic activation. The initial compound, the procarcinogen, is converted by host enzymes to an electrophilic derivative, which then chemically modifies DNA. Mutations result from errors that occur during DNA replication as a result of distorted base pairs. Factors that influence the potency of any chemical carcinogen include the equilibrium between activation of the procarcinogen and deactivation or degradation of the carcinogen.[80] Deactivation typically occurs through a conjugation reaction, usually in the liver.

These principles are exemplified by experimental colonic carcinomas that arise in rodents fed cycasin, a glucosylated compound present in the cycad nut. The glucose residue of cycasin is cleaved in the rat liver by α-glucosidase to form methylazoxymethanol, which is subsequently deformylated by enzymes in the liver and colon to give rise to methyldiazonium, a carcinogen. These same metabolites are formed through hepatic enzymatic modification of the compound dimethylhydrazine and result in colon cancer in the rat.

In humans, regular tobacco use is strongly associated with a higher risk of multiple GI cancers, including pancreatic and colon cancer. Among active smokers with long-term tobacco use, the risk for pancreatic cancer can be elevated twofold. Multiple carcinogenic agents including arsenic, benzene, and ethylene oxide have been identified in cigarettes, but the chemicals linked specifically to the development of pancreatic or colon cancer have not yet been defined.

Dietary Factors

Chemical mutagenesis may be especially important in the development of cancers within the GI tract and related organs. The mucosal surfaces from which most primary cancers in the GI tract develop are exposed to a complex mixture of dietary constituents that are potential carcinogens or procarcinogens. The ability of dietary factors to act as mutagens in humans was demonstrated directly in 1995. The frequency of contamination of food with aflatoxins, a fungal metabolite, parallels the incidence of hepatocellular carcinoma in various areas of the world.[81] Studies demonstrating that aflatoxins cause mutations in the *TP53* gene in hepatocellular carcinoma have provided a compelling link between genes and the environment.[81]

Nitrates present in many foods appear to be additional dietary constituents that may act as procarcinogens in the GI tract. Diet-derived nitrates can be converted by bacterial action in a hypochlorhydric stomach to nitrites and subsequently to mutagenic nitrosamines.[82] These events may underlie the documented correlation between dietary intake of foods high in nitrates and the incidence of gastric cancer in different populations.

Other dietary factors may modulate the biological potency of dietary procarcinogens. Variations in the relative and absolute amounts of dietary fats may lead to alterations in the composition of the colonic microflora and their metabolic characteristics, resulting in modulation of the production of enzymes that convert dietary constituents into potentially mutagenic compounds. Changes in dietary fiber content can alter the transit time of luminal contents in the bowel, thereby changing the duration of exposure of the mucosa to potential mutagens. Bile salt content may be an additional luminal factor that can modulate the biological effect of procarcinogens. Deconjugated bile salts may promote carcinogenesis through mucosal injury and enhanced epithelial proliferation.

These mechanisms could explain well-documented correlations between the intake of various dietary constituents and the incidence of colorectal cancer in certain populations (see Chapter 127). Populations that have a high fiber intake and resulting

fast colonic transit times generally exhibit a lower incidence of colorectal cancer than populations with low fiber intake and delayed transit. The incidence of colorectal cancer in Japanese immigrants to the United States who consume a Western diet is much higher than that of native Japanese who consume a traditional Japanese diet.[83]

MOLECULAR MEDICINE: CURRENT AND FUTURE APPROACHES IN GASTROINTESTINAL ONCOLOGY

Next Generation Sequencing

DNA sequencing relies on polymerase-mediated strand synthesis and the detection of the incorporated nucleotides throughout the successive steps of the chemical reaction, by a variety of physicochemical methods. The ability to monitor billions of reactions simultaneously, so-called massively parallel sequencing, along with the ability to computationally assemble short sequence reads into a continuous long read, have revolutionized sequencing technologies. These new approaches, often referred to as next generation sequencing (NGS), are finding their way into the clinical care of patients with cancer in a variety of settings.[84] First, sequencing of germline DNA is increasingly used to define if a patient may have a cancer genetic syndrome. Secondly, these technologies can be applied to determine the mutational landscape of a tumor to guide treatment decisions.

The extent of DNA sequencing may involve the entire genome. Whole genome sequencing uses DNA from a defined source without any step of enrichment or selection. Another method is to subject the sample to preliminary step of enrichment, where areas of interest are extracted from the sample, using hybridization methods and primer libraries, with the goal of decreasing the complexity of the sample and increasing the number of reads possible during the sequencing reaction. With greater number of reads available, so-called reading depth, the accuracy of sequencing increases and the cost is also reduced. The most common enrichment method is to focus on the areas of the genome known to harbor genes, collectively referred to as the exome, which corresponds to about 1% of the entire genome. For certain applications, subsets of genes from the entire exome may be the only ones enriched for sequencing, and this is the basis for NGS-based diagnostic tests that focus on gene panels relevant to cancer. Because NGS involves short reads that are computationally assembled into predicted long reads, this technology is insensitive to gene inversions, large insertions, or generally copy number variants that affect one allele.

Cancer and Tumor Genomics

As genetic information is obtained from sequencing analysis, understanding the potential impact of the genetic changes observed becomes an important challenge. *Single nucleotide variants* refer to changes in a single base pair of the genetic code compared to a reference sequence. *Nonsense mutations* refer to the introduction of a premature stop codon. Single nucleotide variants at splice-acceptor or donor sites may result in exon loss or misexpression of intronic sequences. These types of changes are relatively easy to interpret and adjudicate. Missense changes are those that result in a change in the amino acid encoded by the codon. Given the normal genetic variation present in the human population, understanding whether these changes are deleterious can be quite difficult to accomplish. When the effect of such variants is not known, these are referred as "variants of unknown significance." Important limitations of exome sequencing at the present time include variants of unknown significance adjudication, detection of copy number variants and large rearrangements, and the potential for intronic or promoter mutations not detectable by exome capture strategies.

Molecular Diagnostics

Genetic testing is a powerful tool to identify high-risk families and define the cancer risk for individual family members. Today, sequencing panels that assess most of the genes associated with familial cancer syndromes are commercially available. Application of genetic testing must take into consideration the sensitivity and specificity of the assay as well as issues of patient confidentiality and potential impact on medical insurability. Because these tests rely on target enrichment, it is important to be aware of their potential limitations. For these reasons, genetic counseling is an essential component of the genetic testing process.

In addition to genetic germline testing, molecular phenotyping of tumors for the purpose of guiding therapeutic decisions is important. To detect tumors due to defects in mismatch repair, testing for MSI can be performed on archived colon tumor samples.[85] In addition, loss of immunohistochemical staining for any of the 4 proteins required for mismatch repair (MLH1, PMS2, MSH2, MSH6) may provide similar information. Studies have demonstrated that the MSI status of a colon tumor is predictive of the response to 5-fluorouracil–based chemotherapy.[86,87] More recently, it has been shown that mismatch repair–deficient tumors, due to their high burden of somatic mutations and tumor neoantigens, are highly responsive to immune checkpoint inhibition therapy.[88]

Therapies that target specific signaling pathways are likely to increase as our molecular understanding of GI cancers increases. Antibodies that target EGF receptors and block the EGF receptor signaling pathway have proved therapeutic benefit in colorectal cancer. However, their benefit has been shown only in cancers lacking activating mutations in *K-ras*. Testing for *K-ras* mutations in colorectal cancers is now standard of care before administration of such targeted therapy. In addition, small molecule tyrosine kinase inhibitors of the *c-KIT* oncogene now constitute routine treatment of GI stromal tumors (see Chapter 33).[89] Molecular techniques may also find a role in the staging of disease. For example, capture of small numbers of circulating tumor cells prior to the discovery of metastasis may yield prognostic and therapeutic benefits.[90] Finally, as more tests for genetic markers become available, monitoring for disease recurrence after surgery may become another important application.

Full references for this chapter can be found on www.expertconsult.com.