Molecular biology of colorectal cancer: Review of the literature

Elrasheid A. H. Kheirelseid, Nicola Miller, Michael J. Kerin

Department of Surgery, National University of Ireland, Galway, Ireland
Email: rashmed1111@gmail.com

Received 10 February 2013; revised 26 March 2013; accepted 20 April 2013

ABSTRACT

Colorectal cancer (CRC) results from the progressive accumulation of genetic and epigenetic alterations that lead to the transformation of normal colonic epithelium to colon adenocarcinoma. From the analysis of the molecular genesis of colon cancer, four central tenets concerning the pathogenesis of cancer have been established. The first is that the genetic and epigenetic alterations that underlie colon cancer formation promote the cancer formation process because they provide a clonal growth advantage to the cells that acquire them. The second tenet is that cancer emerges via a multi-step progression at both the molecular and the morphologic level. The third is that loss of genomic stability is a key molecular step in cancer formation. The fourth is that hereditary cancer syndromes frequently correspond to germ line forms of key genetic defects whose somatic occurrences drive the emergence of sporadic colon cancers.

Keywords: Genetics of Colorectal Cancer; Molecular Biology of Colorectal Cancer; Colorectal Cancer

1. INTRODUCTION

Oncogene and tumour-suppressor gene mutations all operate similarly at the physiological level: they drive the neoplastic process by increasing tumour cell number through the stimulation of cell birth or the inhibition of cell death or cell-cycle arrest. The increase can be caused by activating genes that drive the cell cycle, by inhibiting normal apoptotic processes or by facilitating the provision of nutrients through enhanced angiogenesis. A third class of cancer genes, called stability genes, promotes tumourigenesis in a completely different way when mutated. This class includes the mismatch repair (MMR), nucleotide-excision repair (NER) and base-excision repair (BER) genes responsible for repairing subtle mistakes made during normal DNA replication or induced by exposure to mutagens. Other stability genes control processes involving large portions of chromosomes, such as those responsible for mitotic recombination and chromosomal segregation (e.g., BRCA1, BLM and ATM). Stability genes keep genetic alterations to a minimum, and thus when they are inactivated, mutations in other genes occur at a higher rate [1]. All genes are potentially affected by the resultant increased rate of mutation, but only mutations in oncogenes and tumour-suppressor genes affect net cell growth and can thereby confer a selective growth advantage to the mutant cell. As with tumour-suppressor genes, both alleles of stability genes generally must be inactivated for a physiologic effect to result.

Mutations in these three classes of genes can occur in the germline, resulting in hereditary predispositions to cancer, or in single somatic cells, resulting in sporadic tumours. It is important to point out that a mutation is defined as any change in the sequence of the genome. These changes include those affecting single base pairs as well as those creating large or small deletions or insertions, amplifications or translocations. In the germline, the most common mutations are subtle (point mutations or small deletions or insertions), whereas all types of mutation can be found in tumour cells. In fact, cancers represent one of the few disease types in which somatic mutations occurring after birth are pathogenic. The first somatic mutation in an oncogene or tumour-suppressor gene that causes a clonal expansion initiates the neoplastic process [2]. Subsequent somatic mutations result in additional rounds of clonal expansion and thus in tumour progression [3]. Germline mutations of these genes cause cancer predisposition, not cancer per se. Such individuals therefore often develop multiple tumours that occur at an earlier age than in individuals whose cancer gene mutations have all occurred somatically [4].
2. GENETICS AND EPIGENETICS ALTERATIONS IN CRC

Colorectal cancer results from the progressive accumulation of genetic and epigenetic alterations that lead to the transformation of normal colonic epithelium to colon adenocarcinoma. From the analysis of the molecular genesis of colon cancer, four central tenets concerning the pathogenesis of cancer have been established. The first is that the genetic and epigenetic alterations that underlie colon cancer formation promote the cancer formation process because they provide a clonal growth advantage to the cells that acquire them. The second tenet is that cancer emerges via a multi-step progression at both the molecular and the morphologic levels [5]. The third is that loss of genomic stability is a key molecular step in cancer formation [6]. The fourth is that hereditary cancer syndromes frequently correspond to germ line forms of key genetic defects whose somatic occurrences drive the emergence of sporadic colon cancers [7].

2.1. Genetic Alterations

Much progress has been made in understanding the molecular mechanism of CRC since 1990, when Fearon and Vogelstein proposed their genetic model for CRC tumorigenesis [5]. A progression from normal mucosa to adenoma to carcinoma was supported by the demonstration of accumulating mutations in genes of APC, K-RAS, P53 and DCC, all of which are thought to be of significance, but are not able successfully to account for all CRCs. The earliest identifiable lesion in colon-cancer formation is the aberrant crypt focus (ACF). The true neoplastic potential of this lesion is still undefined, but it does appear that some of these lesions can progress to frank adenocarcinoma and harbor mutations in K-RAS or APC. In particular, dysplastic aberrant crypt foci frequently carry mutations in APC and appear to have the highest potential for progressing to colon cancer. Thus, alterations in APC, which result in overactivation of the Wingless/Wnt signalling pathway, appear to initiate tumour formation in the colon. Subsequent alterations in other genes then play a role in tumour growth and the eventual acquisition of other malignant characteristics such as tissue invasiveness and the ability to metastasize.

2.1.1. APC

The Adenomatous polyposis coli (APC) gene encodes a protein that possesses multiple functional domains that mediate oligomerization as well as binding to a variety of intracellular proteins including β-catenin, γ-catenin, glycogen synthase kinase (GSK)-3β, axin, tubulin, EB1, and hDLG [7]. Germline mutations in APC result in FAP or one of its variants, Gardner’s syndrome, attenuated FAP, Turcotte’s syndrome, or the flat adenoma syndrome [8,9]. In addition, studies have shown that APC is mutated in up to 70% of all sporadic colon adenocarcinomas, which is a high APC mutation frequency unique to colorectal cancers [10,11]. These mutations are present beginning in the earliest stages of colon-cancer formation and precede the other alterations observed during colon-cancer formation [12,13]. One of the central tumour-promoting effects of these mutations results in over-activation of the Wingless/Wnt signalling pathway, with the subsequent expression of genes that favor cell growth. APC mutations disrupt the association of APC with β-catenin, resulting in excessive amounts of β-catenin and overactivation of the Wnt signalling pathway. Consequently, genes that promote tumour formation are transcribed. The over-activation of the Wnt signalling pathway occurs because normally GSK-3β forms a complex with APC, β-catenin, and axin and phosphorylates these proteins. The phosphorylation of β-catenin targets it for ubiquitin-mediated proteasomal degradation. Truncating APC mutations prevent this process from happening and cause an increase in the amount of cytoplasmic β-catenin, which can then translocate to the nucleus and interact with other transcription factors.

2.1.2. K-RAS

Kirsten rat sarcoma (K-RAS) is a member of the RAS family of genes and present one of the most prominent proto-oncogenes in colon carcinogenesis. The RAS family genes encode highly conserved proteins that are involved in signal transduction. One major function of the RAS protein family is to couple growth factors to the Raf-mitogen-activated protein (MAP) kinase kinase-MAP kinase signal transduction pathway, which leads to the nuclear expression of early response genes [14]. K-RAS mutations have been found in 37% - 41% of colon carcinomas and appear to occur relatively early in colon-cancer formation [15,16]. Vogelstein et al. [13] found K-RAS mutations in 13% of small tubular adenomas, 42% of large adenomas, and 57% of adenomas that contained areas of invasive carcinoma. In fact, 58% of adenomas greater than 1 cm in size had RAS mutations, compared to 9% of adenomas less than 1 cm in size [13]. These results have been supported by other investigators who have found an incidence of approximately 40% in colon adenomas [15]. The K-RAS mutations appear to follow APC mutations and are associated with advanced adenomatous lesions. Evidence for this model comes from the observation that small adenomas with APC mutations carry K-RAS mutations in approximately 20% of the tumours, whereas approximately 50% of more advanced adenomas have K-RAS mutations [12,17]. Thus, alterations of K-RAS appear to promote colon-cancer formation early in the adenoma-carcinoma sequence by mediating adenoma growth. Of interest,
however, they do not appear necessary for the malignant conversion of adenomas to adenocarcinomas.

2.1.3. P53
Tumour protein-53 (P53) was initially identified as a protein forming a stable complex with the SV40 large T antigen and was originally suspected to be an oncogene [18]. Subsequent studies demonstrated that P53 is a transcription factor with tumour suppressor activity, located at chromosome 17p13.1, and is mutated in 50% of primary human tumours, including tumours of the gastrointestinal tract [19]. P53 is currently believed to be a transcription factor that is involved in maintaining genomic stability through the control of cell cycle progression and apoptosis in response to genotoxic stress [19]. In colon cancers, P53 mutations have not been observed in colon adenomas, but rather appear to be late events in the colon adenoma-carcinoma sequence that may mediate the transition from adenoma to carcinoma [13]. Furthermore, mutation of P53 coupled with loss of heterozygosity (LOH) of the wild-type allele was found to coincide with the appearance of carcinoma in an adenoma, thus providing further evidence of its role in the transition to malignancy [20,21]. The function of P53 to recognize DNA damage and induce cell cycle arrest and DNA repair or apoptosis has led to P53 being called the “guardian of the genome” [22]. Thus, P53 normally acts as a tumour suppressor gene by inducing genes that can cause cell cycle arrest or apoptosis and also by inhibiting angiogenesis through the induction of TSP1 [23]. Mutant P53 can block these functions by forming oligomers with wild-type TP53, thereby causing diminished DNA-binding specificity [24].

2.1.4. DCC
Since it was first discovered in a colorectal cancer study in 1990 [25], DCC (Deleted in colorectal cancer) has been the focus of a significant amount of research. DCC held a controversial place as a tumour suppressor gene for many years, and is well known as an axon guidance receptor that responds to netrin-1 [26]. More recently DCC has been characterized as a dependence receptor, and theories have been put forward that have revived interest in DCC’s candidacy as a tumour suppressor gene, as it may be a ligand-dependent suppressor that is frequently epigenetically silenced. One of the most frequent genetic abnormalities that occur in advanced colorectal cancer is loss of heterozygosity (LOH) of DCC in region 18q21. DCC elimination is not believed to be a key genetic change in tumour formation, but one of many alterations that can promote existing tumour growth.

2.2. Epigenetic Alterations
The finding of aberrant hMLH1 promoter methylation in sporadic MSI colon cancers dramatically illustrated the role of epigenetic changes as potential pathogenetic alterations in cancer [27-30]. The term DNA methylation refers to the methylation of cytosine residues (5-methyl cytosine) at CpG sites found throughout the genome [31]. These epigenetic alterations are characteristically clustered in so-called CpG islands in gene promoter regions, and hypo and hypermethylation of these regions are related to activation and inhibition of transcription, respectively. This type of gene regulation is essential to cell differentiation as well as embryological development [32]. Furthermore, DNA methylation is closely related to the mechanism by which one copy of a gene is preferentially silenced according to parental origin, generally referred to as genomic imprinting [33]. aber rant methylation of the cancer genome, and associated silencing of the genes whose promoters demonstrated such methylation, has been well described at multiple genetic loci [34,35]. Reversion of the methylation using demethylating agents such as 5-deoxy-azacytidine frequently restores expression of these genes, demonstrating methylation in fact induces gene silencing. As inactivation of hMLH1 plays an initiating role in the pathogenesis of MSI colon cancers, the finding of aberrant methylation of hMLH1 in sporadic MSI colon cancers, and the restoration of hMLH1 expression by demethylating the hMLH1 promoter in cell lines derived from such cancers, strongly suggests that such aberrant methylation could be a cause rather than a consequence of colon carcinogenesis [28-30]. Moreover, Grady et al. [36] provided additional evidence for the primary role of aberrant methylation in gastrointestinal carcinogenesis. They demonstrated that, loss of expression of E-cadherin (CDH1) in association with CpG methylation of the wild-type CDH1 allele in tumours occurs in the setting of the cancer family syndrome Hereditary Diffuse Gastric Cancer. Epigenetic and genetic changes also appear to cooperate to promote cancer formation [37]. Moreover, the aberrant hypermethylation of 50 CpG dinucleotides that has been demonstrated to silence a variety of tumour suppressor genes including CDH1, CDKN2A/p16, TSP1, and GSTP1 may be similarly pathogenic in the tumours in which these changes have been identified [28,37-40].

3. GENETIC CLASSIFICATION OF CRC
Colorectal carcinoma (CRC) is generally classified into three categories, based on increasing hereditary influence and cancer risk [41]. Sporadic CRC (60%) comprises patients with no notable family history and, by definition, with no identifiable inherited gene mutation that accelerates cancer development. Familial CRC (30%) refers to patients who have at least one blood relative with CRC or an adenoma, but with no specific germline mutation or clear pattern of inheritance. Hereditary CRC
syndromes (10%) which result from germline inheritance of mutations in highly penetrant cancer susceptibility genes. Although the last group is observed with the lowest frequency, however they have been instrumental in the elucidation of molecular mechanisms of carcinogenesis applicable to sporadic CRC.

3.1. Sporadic CRC

Sporadic colorectal cancers arise at a median age of 70-75 years. Seventy percent arise in the left side of the colon and there are differences in the age, sex and regional distribution of both adenomas and carcinomas between both sides of the large bowel. Sporadic cancers caused by the development of a series of genetic abnormalities in tumour suppressor genes and oncogenes that give cells an evolutionary advantage over their neighbours.

3.2. Hereditary and Familial CRC Syndromes

3.2.1. Hereditary Non-Polyposis Colorectal Cancer (Lynch Syndrome)

Hereditary non-polyposis colorectal cancer, also referred to as the Lynch syndrome, is the most common form of hereditary colorectal cancer. It is inherited in an autosomal dominant fashion its clinical consequences develop from germline mutations in mismatch repair (MMR) genes. The lack of functional MMR proteins leads to genomic instability and development of various cancers. Multiple generations are affected with colorectal cancer at an early age (mean, approximately 45 years) with a predominance of right-sided colorectal cancer (approximately 70 percent proximal to the splenic flexure). There is an excess of synchronous colorectal and metachronous colorectal cancers, namely, carcinoma of the endometrium, colon and there are differences in the age, sex and regional distribution of both adenomas and carcinomas between both sides of the large bowel. Sporadic cancers caused by the development of a series of genetic abnormalities in tumour suppressor genes and oncogenes that give cells an evolutionary advantage over their neighbours.

3.2.2. Familial Adenomatous Polyposis

Familial adenomatous polyposis (FAP) is characterized by numerous (>100, usually several hundreds in fully developed cases) of adenomatous colorectal polyps. In general, less than 1% of all new CRC arise in FAP patients. FAP is an autosomal dominant hereditary cancer syndrome caused by a germline mutation in the APC gene (adenomatous polyposis coli). Because this syndrome may be associated fewer number of colonic polyps (“attenuated” FAP), it may first present with extra-intestinal manifestations and because as many as 50% of FAP patients result from new germline mutations in APC patients result from new germline mutations in APC patients result from new germline mutations in APC patients result from new germline mutations in APC gene, pathologists may be the first to suspect this hereditary condition. Gardner syndrome is characterized by epidermoid cysts, osteomas, dental anomalies and desmoid tumours (fibromatoses). Turcot syndrome is an association between colorectal polyposis and primary central nervous system (CNS) tumour (usually medulloblastomas) [47,48]. Extra-gastrointestinal manifestations may be of importance for practicing pathologists in diagnosis of unsuspected FAP. Desmoid tumour (fibromatosis) is rare in the general population, but it is commonly seen in FAP and it may be the first manifestation of disease. Patients with FAP typically develop in retroperitoneal tissues or in the abdominal wall following surgical trauma (abdominal desmoids), while fibromatoses unrelated to FAP are more common in extra-abdominal localizations [49]. Papillary carcinoma of the thyroid and its rare cribriform-morular variant may be associated with FAP, and this could lead to detection of unsuspected FAP [50,51]. The risk of hepatoblastoma in children of patients with FAP is highly increased and new germline mutations can be identified in 10% of cases [52].

3.2.3. Hamartomatous Polyposes Syndromes

The hamartomatous polyposis syndromes include Peutz-Jeghers syndrome, juvenile polyposis, Cronkhite-Canada, and Cowden disease/Bannayan-Riley-Ruvalcaba syndrome [53-55]. Hamartomatous polyposes syndromes are distinguished by their characteristic clinicopathologic and radiologic features.

All of these syndromes are characterized by hamartomatous polyps and most of them are associated with increased risk of development of gastrointestinal and extraintestinal carcinomas [56].

1) Peutz-Jeghers Syndrome (PJS):

Peutz-Jeghers syndrome (PJS) is characterized by mucocutaneous pigmentation and GI hamartomas, which occur anywhere from stomach to anus. It was first described by Peutz in 1921 [57] and Jeghers in 1944 [58]. It is inherited in an autosomal dominant fashion with no sex predilection [59]. A prototypic PJS polyp is a hamartoma of the muscularis mucosae. Therefore, the core of the polyp consists of smooth muscle covered by lamina propria and mature glandular epithelium [60,61] which gives rise to a characteristic arborising smooth muscle core of the polyp. Germ-line mutations in the serine/threonine kinase gene (STK11/LKB1) on chromosome 19p13.3 cause Peutz-Jeghers syndrome in about half of the affected families. Additional loss of the wild-type allele in hamartomas and adenocarcinomas suggests that STK11/LKB1 is a tumour suppressor gene [62].

2) Juvenile polyposis coli (JP):

Juvenile polyposis (JP) coli is inherited in an auto-
4. GENETIC INSTABILITY OF CRC

Colorectal cancer is a heterogeneous disease that can develop through different genetic pathways. The most common is termed the chromosomal instability pathway and accounts for 70% to 85% of colorectal cancers (13). These tumours are characterized by mutations in *APC*, *P53*, and *KRAS* and by frequent allelic loss at 18q [68]. Aneuploidy, amplifications, and translocations are also common in these tumours. Familial adenomatous polyposis is the hereditary syndrome associated with these changes [46]. The microsatellite instability (MSI) pathway, comprising the remaining 15% of colorectal cancers, is characterized by loss of proficiency of the DNA mismatch repair (MMR) system and MSI.

Microsatellite Instability (MSI)

Microsatellites are repeated DNA sequences, usually 1 to 10 nucleotides long, present throughout the genome. Instability is mostly characterized by single base-pair insertions or deletions in these repeat loci, causing widespread genomic instability due to the failure of the cell’s mismatch repair (MMR) mechanism. MSI occurs as a consequence of inactivation of the mutation mismatch repair system and is recognized by frame shift mutations in microsatellite repeats located throughout the genome. Inactivation of the MMR system due to germ line gene defects accounts for the colon cancer family syndrome, hereditary non-polyposis colon cancer syndrome (HNPCC). Somatic inactivation of the mismatch repair system additionally gives rise to approximately 15% of sporadic colon cancers. In either instance the resulting colon cancers display the phenotype of microsatellite instability. The demonstration of microsatellite unstable cancers is generally performed by assaying for alterations at microsatellite loci that are particularly frequently mutated in the setting of MMR inactivation. Since many colon cancers demonstrate frame shift mutations at a small percentage of microsatellite repeats, the designation of a colon tumour as showing microsatellite instability depends on the detection of at least two unstable loci out of five from a panel of loci that were selected by a National Cancer Institute consensus conference [69].

Genes that possess such microsatellite-like repeats in their coding regions appear to be the targets relevant to carcinogenesis. Indeed, frequently, many genes that possess microsatellite repeats are observed to be mutated in MSI colon cancers. The relationship between the microsatellite mutator pathway and other genetic alterations frequently found in colon cancer is only partially understood. Alteration of the Wnt/Wingless pathway can be observed in tumours irrespective of MSI [70]. Mutations in *APC* and *CTNNB1* can be found in 21% and 43% of MSI tumours, respectively [71,72]. In addition, the incidence of *K-RAS* mutations appears to be as high as 22% - 31%, which is similar to the incidence observed in microsatellite stable (MSS) colon cancers [73]. Mutations in *P53* do appear to be less frequent in MSI cancers than in MSS cancers. The mutation incidence in MSI colon cancers has been demonstrated to range between 0% - 40%, whereas the incidence in MSS tumours is between 31% - 67% [71,73,74].

The MSI tumour formation process has been termed the microsatellite mutator phenotype and is a pathway to tumour formation that is distinct from those in colon cancers that are microsatellite stable [75-77]. The most frequently targeted gene for mutation in this pathway is the TGF-β receptor type II tumour suppressor gene (*TGFBR2*). Other less frequently targeted genes include the *ACVR2*, *BAX*, *RIZ*, *CDX2*, *SEC63*, *AIM2*, *MSH3* and *MSH6* [77-80]. *CTNNB1* mutations are also found in 25% of MSI colon cancers but are not found in MSS cancers.

The Lynch syndrome is caused mainly by germ-line mutations in the DNA mismatch repair genes and heterozygosity for a mutation results in susceptibility to the cancer. Lynch syndrome can be identified based on age at onset, previous medical history and the characteristics of family history that fulfill the Amsterdam criteria and Bethesda guidelines for the diagnosis of hereditary non-polyposis colorectal cancer (HNPCC) [81-84]. The early recognition of Lynch syndrome is essential to identify patients at high risk who will require intensive surveillance. Nevertheless; its diagnosis can be difficult to make due to incomplete family history information and lack of characteristic clinical phenotype. Although the Amsterdam criteria and Bethesda guidelines continue to be used widely, several studies have underscored the limitations of their accuracy in predicting the presence of MMR gene mutations [85,86]. Therefore, new strategies for screening for and diagnosis of Lynch syndrome need to be investigated.

In addition to screening for Lynch syndrome, testing for MSI is important because of its possible prognostic and therapeutic implications. Cancers with high mi-
crosatellite instability (H-MSI) were reported to have a more favourable clinical outcome than non-MSI tumours and the survival advantage conferred by the MSI phenotype is independent of tumour stage and other clinicopathological variables [87-89]. Moreover, tumours with H-MSI are thought to be less responsive to 5-fluorouracil and other anticancer agents in vitro and in vivo [90-92].

5. ACKNOWLEDGEMENTS

The authors would like to acknowledge the National Breast Cancer Research Institute (NBCRI) their funding of this study.

REFERENCES

lation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. Cancer Research, 57, 808-811.


