Invited Review

Recent advances in the biology of colorectal cancer

L. Grogan, K.A. Behan and P.G. Johnston
NCI-Navy Medical Oncology Branch, National Cancer Institute, Bethesda, USA

Summary. The identification of the precise molecular defects responsible for the common forms of inherited colorectal cancer has significantly advanced our understanding of both inherited and sporadic disease. These advances coupled with a rapid accumulation of information on the molecular genotype and biological phenotype of colorectal cancer have identified potential markers that may be not only of prognostic value but also important as screening tools and therapeutic targets. These molecular and biological features include replication errors, mutations of oncogenes and tumor suppressor genes and expression of tumor specific antigens and cytokeratins. This review highlights important recent advances that further our understanding of the biology and genetics of colorectal cancer.

Key words: Replication errors, Adenomatous polyposis coli, Tumor suppressor genes, Oncogenes, Cytokeratins

Introduction

Colorectal cancer is a major public health problem in the United States and other Western countries. In 1996 it is estimated that there will be 159,000 new cases and approximately 58,000 deaths from colorectal cancer in the United States. At the time of initial diagnosis seventy-five percent of patients have disease resectable with potential curative intent. Nevertheless, forty to fifty percent of these patients will die from their disease. Colorectal cancer is the third leading cause of death from cancer in males and females and accounts for approximately 18% of all cancer deaths (Boring et al., 1994). At present the major determinant of prognosis and treatment benefit relies on pathological and clinical staging criteria. Patients with Dukes' A disease have an estimated 5 year survival of 90%, whereas, patients with Dukes' D metastatic disease have a very poor survival of approximately 8-12 months. In patients with Dukes' B and C disease 5 year survival ranges from 70%-75% and 45-55% respectively (Moertel, 1994). Over the last 20 years, many treatment strategies to improve disease free survival and survival have been unsuccessful; however, recent clinical trials have demonstrated that adjuvant 5-fluorouracil/levamisole or 5-fluorouracil/leucovorin chemotherapy may improve survival with certain patient subgroups (Wolmark et al., 1993; Moertel, 1994).

The ability to study colonic adenocarcinomas at various stages of development has led to a better understanding of the biology of colorectal cancer, as has the clinical and molecular characterization of familial syndromes in which the inheritance of specific genetic defects leads to the development of colorectal carcinoma. These important biological properties have potential as diagnostic and prognostic tools and may impact adjuvant therapy for this disease. This review highlights important advances in the understanding of the biology and genetics of colorectal cancer that are likely to result in further improvements in the treatment of patients with this disease.

Hereditary factors in colorectal cancers

The role of inheritance in the development of colorectal cancer is most evident in two syndromes, familial adenomatous polyposis (FAP) and hereditary non polyposis colorectal cancer (HNPCC). An individual with a positive family history has a 2-3 fold increased risk of developing cancer of the colon and over 20% of all colorectal cancers are thought to have an inherited component (Cannon-Albright et al., 1988; St. John et al., 1993; Cunningham and Dunlop, 1994; Fuchs et al., 1994; Marra and Boland, 1995).

FAP is an autosomal dominant trait with almost complete penetrance and a varied expression of a characteristic clinical phenotype (Cunningham and Dunlop, 1994). FAP typically presents with hundreds to thousands of adenomas, which appear early during the second decade of life and in extreme cases carpet the entire colon and rectum. The progression to colorectal cancer is inevitable and occurs in a small number of these colonic adenomas usually by the fourth decade of life. It is associated with a variety of extracolonic
features both benign and malignant (Cunningham and Dunlop, 1994). Benign features include a pigmented fundal lesion called congenital hypertrophy of the retinal pigment, epidermoid cysts, dental abnormalities, jaw cysts, sebaceous cysts, fibromas, lipomas and osteomas (Järvinen and Sipponen, 1986; Chapman et al., 1989). More ominous extracolonic features include gastric, duodenal polyps and endocrine adenomas (Järvinen and Sipponen, 1986). Periampullary duodenal carcinoma is the second most common cancer observed in patients with this disease and occurs in FAP patients treated successfully with colectomy (Herrera-Ornelas et al., 1987; Lynch et al., 1983; Watson and Lynch, 1993; Bertario et al., 1994). Ileal carcinoma is also common (Herrera-Ornelas et al., 1987; Lynch et al., 1993; Watson and Lynch, 1993; Bertario et al., 1994). The syndrome is associated with an increased incidence of a variety of other cancers which arise outside of the gastrointestinal tract including hepatoblastoma in children, papillary thyroid cancer, adrenal cortical tumors, medullablastoma and mesenteric fibromatosis which are slow growing desmoid tumors that usually develop after abdominal surgery in adults (Järvinen, 1987; Klemmer et al., 1987; Hughes and Michels, 1992). The presence of specific extracolonic features was used to describe distinct clinical syndromes such as Gardner’s syndrome (osteomas, epidermal inclusion cysts, fibromas and colorectal cancer) and Turcot’s syndrome (brain tumors especially medullablastoma and occasionally glioblastoma and colorectal polyposis). However the majority of patients with FAP have one or more extracolonic features suggesting that the features identifying a distinctive phenotype in Gardner’s and Turcot’s syndromes are really part of the spectrum of disease abnormalities that are associated with FAP (Bodmer et al., 1987; Nakamura et al., 1988; Nishisho et al., 1991). Confirmation that these clinical entities are not associated with a specific genotype has come from the identification of the molecular basis of FAP (Bodmer et al., 1987; Nakamura et al., 1993; Nishisho et al., 1991).

Linkage studies in FAP families have facilitated the mapping of the adenomatous polyposis gene (APC) to the long arm of chromosome 5 (5q21-22) (Herrera et al., 1986; Bodmer et al., 1987; Groden et al., 1991; Joslyn et al., 1991; Kinzler et al., 1991; Nishisho et al., 1991; Cross et al., 1992). The characterization of germine mutations in the genomic DNA of FAP kindreds established this gene as being responsible for the FAP syndrome (Cotrell et al., 1992; Fodde et al., 1992; Nagase et al., 1992b; Groden et al., 1993; Olschwang et al., 1993a; Powell et al., 1993; Varesco et al., 1993). It is a classical tumor suppressor with inactivation of both alleles required for the development of colorectal cancer in the majority of cases studied (Levy et al., 1994; Luongo et al., 1994). Over 80% of FAP families have a germine mutation of the APC gene and importantly over 25% of new cases represent new germine mutation (Mahe et al., 1993; Bisgaard et al., 1994). The APC gene is very large, encoding a protein of 2843 amino acids in 15 exons and although mutations can occur throughout the gene the majority occur in a 600 bp region of exon 15 (Groden et al., 1991; Kinzler et al., 1991; Nagase et al., 1992a; Groden et al., 1993; Nagase and Nakamura, 1993; Nakamura, 1993). The mutations in general are single base pair changes resulting in termination codons, small deletions, insertions, or splicing mutations which result in a truncated abnormal protein. An attenuated form of FAP is characterized by a lower number of polyps, which although variable, range from 2 to 100 (Leppert et al., 1990; Sprio et al., 1992, 1993a). Despite a lower absolute number of polyps the lifetime penetrance of colon cancer remains very high in these families, however the average age of onset colorectal cancer is usually 10-15 years later than the average for FAP patients. In this variant of FAP the sites of germine mutations are very close to each other at the 5' end of the APC gene (Leppert et al., 1990; Sprio et al., 1992, 1993a). The function of the APC protein is unknown and it shares very little homology with any known proteins or with recognized protein functional regions (Marra and Boland, 1995). It does form heterodimers with itself and binds to E-cadherin associated proteins β- and γ-catenin suggesting it may play a role in cell adhesion and possibly intercellular communication (Rubinfeld et al., 1993; Su et al., 1993a,b, 1995). Mutant APC protein can form heterodimers with wild type protein interfering with normal function and behaving in a dominant negative fashion (Su et al., 1993a). Not all mutant APC proteins however, affect wild type APC function and classical FAP occurs in patients who have constitutional chromosomal deletions resulting in the loss of a single copy of the APC gene. There are some aspects of the phenotypic variation associated with the location of mutations including the attenuated form of FAP linked to a small area in the 5' region of the gene (Sprio et al., 1993b). Moreover congenital retinal pigmentation almost never occurs if a mutation is localized before exon 9 and is almost always present if a mutation occurs after exon 9 (Nagase et al., 1992a,b; Olschwang et al., 1993b). However, identical APC mutations are associated with a wide variety of FAP phenotypes and family members inheriting the same APC mutation may differ markedly in tumor burden suggesting that other factors, may affect the phenotypical expression of a defect in the APC gene (Leppert et al., 1990; Paul et al., 1993; Giardiello et al., 1994).

An ideal model system for studying this syndrome has been developed. MIN mice (Multiple Intestinal Neoplasia) have a germine mutation in the mouse homologue of APC and develop an identical clinical phenotype to FAP with multiple intestinal polyps and subsequent neoplasia (Su et al., 1992; Fodde et al., 1994; Luongo et al., 1994). Like FAP, phenotypic variation exists in the MIN mouse and a gene responsible for at least 50% of this modifying effect has been identified (Dietrich et al., 1993; MacPhie et al., 1995). Increased expression of the secretory type II phospholipase A2
(PLA2s) gene decreases the number of induced intestinal tumors and is associated with a reduction in the number of polyps (MacPhee et al., 1995). PLA2s is an enzyme involved in arachidonic acid synthesis, a rate limiting substrate for the generation of prostaglandins. There is evidence suggesting a role for prostaglandins in adenoma development and studies both in humans and mice have demonstrated that prostaglandin inhibitors such as Sulindac, a non steroidal anti-inflammatory drug (NSAID), are effective in reducing adenoma formation (MacPhee et al., 1995). It was presumed that Sulindac and other NSAIDs worked by inhibiting prostaglandin synthesis; however, the effect of PLA2s in the MIN mouse suggests that expression of prostaglandins may have a protective effect and the mechanism by which NSAIDS reduce polyp numbers may be more complicated than previously thought. High levels of PLA2s may prevent adenoma formation by other mechanisms as it plays a role in controlling normal intestinal bacterial flora, inactivating harmful dietary fats, and maintaining cellular membrane asymmetry (MacPhee et al., 1995). The human homologue of PLA2s is located on chromosome 1p35 and the identification of this modifier gene in the MIN mouse may have important implications for human cancer. The intraluminal expression of PLA2s may protect against the development of adenomas and may account for the wide variation in adenoma formation among family members inheriting an identical APC mutation (MacPhee et al., 1995). It also suggests a potential direct molecular link between high dietary fat intake and increased risk of colon cancer (MacPhee et al., 1995).

**Hereditary nonpolyposis colorectal cancer syndrome (HNPCC)**

Hereditary nonpolyposis colorectal cancer syndrome (HNPCC) is an autosomal dominant inherited predisposition to colon cancer in the absence of colorectal polyposis (Marra and Boland, 1995). Family members have an excess incidence of a variety of other cancers, particularly endometrial adenocarcinoma (Marra and Boland, 1995). They do have adenomatous polyps and like sporadic colorectal cancer the tumor is thought to arise in these polyps (Marra and Boland, 1995). Until recently clinical and epidemiological criteria were used to define two broad categories of HNPCC referred to as Lynch I and Lynch II syndromes (Lynch et al., 1993). Lynch I refers to families in whom there was an excess of early onset, predominantly proximally located colorectal cancer (Lynch et al., 1993). Lynch II cancer family syndrome referred to families with an excess of extracolonic cancers in addition to colorectal cancer (Lynch et al., 1993). An international collaborative group on HNPCC defined families as carrying the HNPCC trait using what is referred to as the Amsterdam criteria (Vasen et al., 1994). These criteria require that in a HNPCC family, at least three relatives must have colorectal cancer, one relative must be a first degree relative of the other two and one affected relative must be diagnosed prior to the age of 50 (Vasen et al., 1994). The exclusion of extracolonic cancers from the definition has been a major source of criticism and may have resulted in an underestimation of the prevalence of this syndrome (Marra and Boland, 1995).

Epidemiological studies using these criteria suggest that HNPCC accounts for between 1% and 5% of all colorectal cancers (Mecklin, 1987; Ponz de Leon et al., 1989, 1993; Kee and Collins, 1991; Ponz de Leon, 1994). Like sporadic colorectal cancer, HNPCC tumors are thought to arise in preexisting adenomas. Polyps are found more frequently in HNPCC patients than in the general population (Lanspa et al., 1990; Jass et al., 1994; Marra and Boland, 1995). In HNPCC sixty percent of colorectal cancers are found proximal to the splenic flexure and the average age is 44 years which is about 20 years younger than sporadic cases. HNPCC are frequently poorly differentiated, associated with excess mucin production and frequently are surrounded by Crohn's like lymphoid aggregates (Lynch et al., 1993). Over 90% of HNPCC tumors are diploid or near diploid (Lynch et al., 1993). A large study of 1300 hundred high risk members of HNPCC families found a significantly increased incidence of cancers of the stomach, small intestine, upper urological tract (Renal pelvis and ureter) and ovary (Watson and Lynch, 1993). No excess incidence of breast, pancreatic or bladder cancer was found despite previous reports linking these tumors to HNPCC (Watson and Lynch, 1993). There was considerable heterogeneity among the families with respect to the types and frequencies of specific extracolonic tumors suggesting that Lynch I and Lynch II are not distinct clinical syndromes but part of a spectrum of a single disease trait (Watson and Lynch, 1993).

Endometrial cancer was accepted as a tumor involved in HNPCC for the purposes of the study with one case of endometrial cancer found for every 5.4 cases of colorectal cancer among family members. Surprisingly these family members had a significantly lower incidence of lung cancer (Watson and Lynch, 1993). Muri-Torre particularly and some cases of Turcots syndromes are both part of the disease spectrum seen in HNPCC (Lynch et al., 1993; Honchel et al., 1994; Kolodner et al., 1994; Hamilton et al., 1995; Marra and Boland, 1995).

Characterization of the molecular defect responsible for HNPCC occurred when almost simultaneously a number of groups discovered that 10-15% of sporadic colorectal cancers and almost all cancers from HNPCC kindreds had an unusual form of genomic instability identified initially in microsatellite sequences which they variously referred to as replication error (RER), microsatellite instability or ubiquitous somatic mutations (Aaltonen et al., 1993; Ionov et al., 1993; Thibodeau et al., 1993). Microsatellites are repeat sequences of mono, di, tri or tetra or more nucleotides which are abundant with over 100000 in the genome and although highly polymorphic throughout the population they are uniform
in the DNA of all cells from an individual (Marra and Boland, 1995). They noticed that microsatellites, which are scattered frequently throughout the genome were mutated in these tumors and these mutations consisted of changes in the lengths of these microsatellites (Aaltonen et al., 1993; Ionov et al., 1993; Thibodeau et al., 1993). This was correctly identified as evidence of errors of insertion or deletion occurring during the replication of DNA in these tumors. This mutational pattern is characteristic of an identical pattern identified in bacteria resulting from defective mismatch repair (Modrich, 1991). The mismatch repair "proofreading system" was already identified and functionally characterized in bacteria. Germline mutations in at least four DNA mismatch repair genes which are homologues of bacterial mismatch repair genes Muts (human MSH2) and MutL (human MLH1, PMS1 and PMS2) have recently been identified as the genetic abnormalities responsible for HNPCC syndromes (Fishel et al., 1993; Leach et al., 1993; Bronner et al., 1994; Nicolaides et al., 1994; Papadopoulos et al., 1994). Mutations of these genes are associated with insertions and deletions in microsatellite sequences and are responsible for defective DNA mismatch recognition and repair (Fishel et al., 1993; Leach et al., 1993; Bronner et al., 1994; Nicolaides et al., 1994; Papadopoulos et al., 1994). Approximately 40%-50% of HNPCC families have germline mutations of hMSH2 and most mutations result in substantial change in the predicted protein (Liu et al., 1994). Approximately 30-40% of families have germline mutations of hMLH1 and less than 10% have mutations of PMS1 and PMS2 (Limdblo et al., 1993; Nystrom-Laiti et al., 1994). The human hMSH-2 is located on chromosome 2p and the human MLH1 gene is located on chromosome 3p23-p21 with two additional genes designated PMS1 and PMS2 located on chromosomes 2p and 7p respectively (Fishel et al., 1993; Leach et al., 1993; Bronner et al., 1994; Nicolaides et al., 1994; Papadopoulos et al., 1994). The identification of defective mismatch repair as the mechanism responsible for HNPCC and a subgroup of sporadic colorectal cancer suggested a new mechanism for cancer development. The inactivation of the mismatch repair system may promote a hypermutable phenotype which is manifested as replication errors in microsatellites and potentially increases the frequency of random mutations in critical oncogenes and tumor suppressor genes (Loeb, 1994). The mutator phenotype has been found in 57% of adenomas and over 85% of colorectal tumors from HNPCC patients and in 100% of patients known to harbor a germline mutation in hMSH2 (Aaltonen et al., 1994). A study by Parsons and colleagues demonstrated that the mutation rate of dinucleotide repeats in RER positive tumor cells is at least 100-fold that in RER negative tumor cells and RER positive cells have a profound defect in strand-specific mismatch repair indicating that a true mutator phenotype exists in a subset of human tumors (Parsons et al., 1993). In vitro studies indicate that heterozygosity of mutations in DNA mismatch repair genes, unlike homozygosity, does not affect mismatch repair suggesting that DNA mismatch repair genes resemble tumor suppressor genes in that two hits are required to cause a phenotypic effect (Parsons et al., 1993). The generation of MSH2-deficient knockout mice confirms that MSH2 is involved in safeguarding the genome from DNA errors and errors in homologous recombination. MSH2-deficient mice displayed no major abnormalities, however, lymphomas developed at an early age and at a high frequency in homozygous MSH2 deficient mice suggesting that a deficiency of a mismatch repair gene is directly involved in the pathogenesis of cancer (de Wind et al., 1995; Reitmair et al., 1995). PMS2 knockout mice are prone to develop lymphomas however the males are infertile (Baker et al., 1995). MSH2 mutations are associated with defects in repair of all types of mismatches but a newly discovered protein that binds to hMSH2, GT binding protein (GTBP) confers a specific repair function namely correcting G-T mismatches or single base loops (Drummond et al., 1995; Palombo et al., 1995; Papadopoulos et al., 1995). The inactivation of this gene may result in less profound microsatellite instability and this suggests that defects in mismatch repair are not associated with a homogeneous replication error phenotype. To date germline mutations in this gene have not been found in any HNPCC families (Drummond et al., 1995; Palombo et al., 1995; Papadopoulos et al., 1995).

A link between the mutator phenotype and tumor progression was established when investigators demonstrated that RER positive cell lines and tumors had mutations in the TGF beta receptor II gene (Markowitz et al., 1995). This gene when inactivated allows cells to escape from transforming growth factor (TGF) beta mediated growth control and the majority of mutations occur in a mononucleotide A repeat in the 5' region of the gene in RER positive tumors (Markowitz et al., 1995). This is the first clear example of a gene targeted for mutation in mismatch repair deficient cells. Moreover recent studies have suggested that cells that have a defect in mismatch recognition and repair are more resistant to alkylators (Kat et al., 1993; Branch et al., 1994; Koi et al., 1994). Whether this association is true for other chemotherapeutic agents and its implications for cancer therapy remains to be determined.

Sporadic non-familial colorectal cancer

Studies of colorectal cancer specimens using probes that detect restriction fragment length polymorphism have demonstrated that widespread loss of chromosomal material is common in the sporadic form of colorectal cancer (Vogelstein et al., 1988). Loss of alleles from chromosomes 5, 17 and 18 are most frequently found in colorectal cancer (Vogelstein et al., 1988). A molecular model based on the accumulation of multiple genetic abnormalities in the transformation from adenoma to
The molecular biology of colorectal cancer

carcinoma was first described by Vogelstein to explain the molecular events in the genesis of colorectal cancer (Fearon and Vogelstein, 1990). In this model it was the accumulation of genetic abnormalities involving both oncogenes and tumor suppressor genes over time rather than the exact order in which these genetic changes occurred that was important (Fearon and Vogelstein, 1990). Single abnormalities such as K-ras or APC mutations occur in early adenomas and are followed by loss of alleles on 5q (APC gene), 17p (p53 gene) and 18q (DCC gene). The APC gene is mutated or inactivated by allelic loss in over 60% of all colorectal cancers and adenomas (Powell et al., 1992). It is inactivated in the smallest adenomas and the frequency of these mutations remains constant as tumors progress from the smallest adenomas to frank malignancy (Powell et al., 1992). In an analysis of small benign colorectal lesions APC mutations were found in 82% of dysplastic polyps, 0% of non dysplastic polyps, 1/1 dysplastic aberrant crypt foci (ACF) and 0/19 non dysplastic ACF. In comparison, K-ras mutations were identified in 25% of dysplastic polyps, 22% of non dysplastic polyps, 0/1 dysplastic ACF and 19/19 of non dysplastic ACF (Jen et al., 1994b). In these early lesions APC mutations were closely associated with dysplasia one of the earliest hallmarks of malignant potential while K-RAS mutations were found to be common in non dysplastic lesions thought to have a low potential for malignant transformation (Jen et al., 1994b). These results suggested that the nature and order of mutations are important in colorectal cancer (Jen et al., 1994b). In the progression from normal colonic mucosa to cancer APC gene may function as a «gatekeeper» whereby inactivation of the gene may be required to initiate malignant transformation.

Microsatellite instability in sporadic colorectal cancer

The frequency of RER positivity in sporadic colorectal cancers has ranged from 13-28% in several studies (Aaltonen et al., 1993; Ionov et al., 1993; Thibodeau et al., 1993; Kim et al., 1994). These studies have also noted that RER-positive tumors are distinct from RER-negative tumors and are associated with certain clinicopathological variables (Aaltonen et al., 1993; Ionov et al., 1993; Thibodeau et al., 1993; Kim et al., 1994). Sporadic RER positive tumors are frequently more HNPCC like than HNPCC itself, almost always occurring on the right side of the colon (80% proximal versus 56% proximal for HNPCC tumors), are diploid, associated with poorly differentiated or mucin producing adenocarcinomas and a marked host lymphocyte response (Kim et al., 1994). There are differences, sporadic RER positive patients develop colorectal cancer over 20 years later on average than HNPCC kindreds and they don’t usually have extracolonic cancers (Kim et al., 1994). An analysis of microsatellite instability in sporadic colorectal cancers suggested that only one in ten RER positive tumors has a germline mutation of the known mismatch repair genes and while investigators have demonstrated that somatic mutations can generate this instability in cell lines the majority of cell lines from RER positive sporadic colorectal cancer do not have any mutation of the genes now known to cause HNPPC (Liu et al., 1995b). Patients under 35 years of age who develop sporadic colorectal cancer more frequently exhibited microsatellite instability, 58% of 31 patients compared to 12% of 158 patients over 35 years of age in a study from Johns Hopkins (Liu et al., 1995a). Forty two percent of patients under 35 and none of the 8 patients over 35 years with RER positive tumors had germline mutations in the mismatch repair genes (Liu et al., 1995a). The majority of patients with colorectal cancer under 35 years of age exhibit defects in mismatch repair and many of these patients had germline mutations in contrast to older patients who rarely had microsatellite instability and none of whom had a germline mutation (Liu et al., 1995a). Mutations of K-ras (61%) p53 (64%) and APC (57%) were similar in patients with familial and sporadic colorectal cancer in one study although an earlier study found that the mutation rate p53 and K-ras was lower in RER positive tumors (Aaltonen et al., 1993; Ionov et al., 1993).

Several of these studies have also attempted to correlate RER status with patient outcome (Thibodeau et al., 1993; Jen et al., 1994a). Patients with RER positive tumors appear to have a better survival than patients with RER negative tumors (Thibodeau et al., 1993; Jen et al., 1994a). However, this association with improved survival may not be an independent prognostic factor. Jen et al. noted that the survival of 18 patients with RER positive tumors was indistinguishable from that of 28 patients who were RER negative and had retained chromosome 18q (Jen et al., 1949a). They also demonstrated that RER positive tumors do not tend to lose chromosome 18q alleles. Thus, the better outcome in patients with RER positive tumors may be due to a retention of both alleles of chromosome 18q or may relate to the marked peritumoral inflammatory cell infiltrate. More studies will be needed to determine the clinical importance and potential therapeutic application of RER status in both HNPPC and sporadic colorectal cancers.

Tumor suppressor genes

Loss of chromosomal alleles usually involves only one of the two parental chromosomes present in normal cells and suggests that these regions contain tumor suppressor genes whose protein products regulate cell division and growth. A candidate tumor suppressor gene located on chromosome 5q has recently been identified called the mutated in colorectal cancer gene (mcc gene). This encodes a 829 amino acid protein with sequence homology to the G-protein coupled to the muscarinic receptor (Lechleiter et al., 1990; Powell et al., 1993). The region of sequence homology coincides with that
region of the receptor necessary for G protein activation. Members of the G protein family are thought to play an important role in cellular signal transduction; however, the exact role of ncc in colorectal cancer remains to be determined.

Over 75% of colorectal cancers have lost a portion of chromosome 17p or 18q or both, due to chromosomal loss or mitotic recombination. The 17p segment that is lost contains the tumor suppressor gene p53. When 17p allelic segments are lost on one chromosome, the remaining p53 allele is generally inactivated by a point mutation.

Mutations of the p53 gene occur at many codons; however, the majority are clustered to 4 regions (codons 5-8) (Baker et al., 1989; Shaw et al., 1991). Most of these mutations are missense mutations that result in amino acid substitution. This gene is strongly associated with the adenoma-carcinoma sequence and encodes a 53 kD nuclear protein which is thought to control the activation of transcription by binding to specific DNA sequences that result in inhibition of cell cycle progression (Yin et al., 1992). A mutation in p53 results in loss of cell cycle control and eventual malignant transformation (Levine et al., 1991). Mutant p53 is an oncogene because of its ability to immortalize normal cells and to transform primary rat embryo cells in conjunction with an activated ras gene. Subsequent transfection studies have demonstrated that while mutated p53 results in cellular transformation wild type p53 inhibits this process (Lane and Benchimol, 1990). Furthermore, in colorectal adenoma cell lines transfection of the wild type p53 gene has been shown to inhibit adenomatous growth (Baker et al., 1990). Missense mutations in p53 gene produce an abnormal protein with a longer half-life which results in accumulation of p53 protein in the nucleus and cytoplasm. While several studies have shown a strong concordance (85-90%) between molecular and immunohistochemical detection techniques; increased levels of p53 protein detected immunohistochemically do not always equate with genomic mutation (Hall and Lane, 1994). Approximately 5-10% of colorectal tumors with elevated levels of p53 expression do not contain mutated p53 protein but reflect increased levels of wild type p53 protein that has been stabilized by binding to other intracellular proteins. Approximately 45-60% of colorectal cancers demonstrate increased expression of p53 by immunohistochemical analysis. Several studies have correlated p53 expression in colorectal cancer with established clinicopathological criteria. These studies suggest that increased p53 expression may be more common in left sided colon tumors with a more advanced disease stage, but no correlation with patients age, sex, and degree of tumor differentiation have been noted. The currently available data on the clinical prognostic relevance of p53 overexpression in colorectal cancer is controversial (Lazaris et al., 1995); the second most common site of allelic loss in colorectal tumors is chromosome 18q. This allele is lost in more than 70% of colorectal carcinomas and in 50% of late adenomas (Fearon et al., 1990; Jen et al., 1994b). The common site of allelic loss contains a gene called the deleted in colon cancer gene (DCC gene). The DCC gene is a putative tumor suppressor gene whose protein product is thought to be a cell-cell, cell-matrix, adhesion molecule (Fearon et al., 1990). The DCC gene product is homologous to neural cell adhesion protein (N-CAM), a cell surface glycoprotein. Deletion of the DCC gene is frequent in colorectal adenocarcinomas but infrequent in colorectal adenomas. Because the DCC gene may represent a cell adhesion molecule, decreased levels of DCC expression may lead to altered adhesion and contribute to enhanced tumor growth and metastatic spread of colorectal tumors. Studies by Zetter et al. have demonstrated normal DCC expression in the majority of non-metastatic colorectal tumors; however, in metastatic colorectal cancers DCC was undetectable (Zetter, 1993). Allelic loss of many other chromosomes involving 1q, 4p, 6q, 8p, 9q, and 22q are also seen in colorectal cancer (Goelz et al., 1985).

The frequency of these chromosomal losses (25-50%) is markedly less than with chromosomes 5q (APC), 17q (p53) and 18q (MCC, DCC) (Jen et al., 1994b). These chromosomes may contain important as yet unidentified suppressor genes that may account for some of the variation in the biological patterns of colorectal tumors in patients. In addition to these specific genetic alterations, abnormal patterns of DNA methylation also appear to be important in colorectal cancer (Silverman et al., 1989; Vogelstein et al., 1989). Methylation is particularly important in the metabolism of the cytosine nucleoside and ultimately in the regulation of genetic expression (Feinberg et al., 1988). In adenomas there is evidence of both hypomethylation and hypermethylation of genes which suggests that DNA methylation abnormalities may play an important role in colorectal tumorigenesis. In one study at least one third of DNA regions studied from small adenomas (<1.0 cm) had lost methyl groups present in the DNA of normal colonic mucosal cells (Vogelstein et al., 1989). Loss of DNA methylation inhibits chromosomal condensation, which may lead to mitotic non-disjunction resulting in loss or gain of chromosomal fragments. This may contribute to cellular genetic instability and alter the rate at which genetic changes such as allelic loss occur.

**Cellular oncogenes**

Molecular analysis of many adenoma and colorectal carcinomas has identified a number of alterations in oncogenes involved in cellular transformation and carcinogenesis. Among the best studied of cellular oncogenes are the ras family of genes, which were originally identified as the cellular homologues of the transforming genes of the Harvey and Kirsten murine sarcoma viruses. Ras gene mutations are commonly found in colorectal carcinomas and were thought to be one of the earliest events in the adenoma-carcinoma sequence (Bos et al., 1987; Bos, 1989). The ras gene
family encodes a 21 kD membrane bound protein involved in signal transduction. Ras proteins associate with the inner surface of the cell membrane and bind guanine nucleotides resulting in hydrolysis of GTP to GDP (Bos, 1989). Mutations localized to ras codons 12, 13 and 61 produce amino acid substitutions that result in inappropriate activation of transmembrane signal transduction (Forrester et al., 1987). Activating mutations of ras oncogenes involving the K-ras gene on the short arm of chromosome 12 and the N-ras gene on chromosome 1p have been identified in approximately 45-50% of colorectal cancers cell lines and tumor specimens (Spandidos and Kerr, 1984; Gallick et al., 1985; Fearon et al., 1987; Finkelstein et al., 1993). Approximately 50% of adenomas greater than 1 cm in size will contain ras mutations compared to only 10-15% of adenomas less than 1 cm (Fearon et al., 1987). The significance of ras gene mutations in colorectal cancer remains unclear. As previously described in this review, k-ras mutations were found to be very common in small non-dysplastic lesions, which have limited potential to progress to larger tumors, whereas APC mutations were correlated with dysplasia, an indicator of potential malignancy (Jen et al., 1994b). Studies of K-ras mutations in dysplastic polyps versus smaller adenomas indicate that K-ras mutations often occur after tumor initiation (Fearon and Vogelstein, 1990). Together, this data would suggest the nature and order of genetic changes may effect tumor morphology and progression. Several studies have correlated the presence or absence of ras mutations with clinicopathological features. These studies have been unable to demonstrate any association between the presence of ras gene mutations and disease free survival or overall survival (Sidransky et al., 1992; Bell et al., 1993; Finkelstein et al., 1993). However, in a study of Dukes B and C colorectal tumors, G-T and G-C transversions on codon 12 of the K-ras gene were associated with metastatic behavior of colorectal carcinomas, while G-A transitions were not. Moreover, the presence of multiple point mutations in codon 12, and an increased number of mutations was associated with the stage of the disease (Moerkerk et al., 1994). These results indicate that colon tumor progression in Duke's B to Duke's C is accompanied by genetic instability, suggesting that the type and number of K-ras point mutations may affect the biological behavior of the tumor. In a study by Breivik et al. DNA from primary tumors from 251 male and female Norwegian patients with colorectal cancer were analyzed for K-ras mutations at codons 12 and 13 (Breivik et al., 1994). The frequency of K-ras mutations was found to be significantly related to age and sex of the patients, and to the location of the tumors. K-ras mutations were much less frequent in colonic tumors in male patients at younger ages (<40). Since the incidence of cancer in the proximal colon is reported to be higher in women than in men at all ages, these results suggest that this sex difference may be related to a low frequency of K-ras mutations. The higher incidence in women of ras mutations in proximal colonic tumors may possibly be explained by intrinsic or environmental factors related to females.

Recently, attempts have been made to develop practical tests which enable the detection of potential oncogenes in patients. Two recent studies were able to identify ras oncogene mutations in cells isolated in the stool from patients with non-metastatic colon cancers and adenomas (Sidransky et al., 1992; Hasegawa et al., 1995). While ras mutations are only detectable in 50% of colorectal tumors it is possible that this approach may be able to detect other mutant genes that may be present in colorectal cancers and would facilitate the development of a sensitive and specific test for detecting colorectal tumors at an early stage of disease.

Another oncogene that may play a role in colorectal carcinogenesis is the c-myc oncogene. The c-myc oncogene was first identified as the homologue of the transforming sequence of the avian myelocytomatosis virus MC29 (Stewart et al., 1986; Sikora et al., 1987). The c-myc gene is located on chromosome 8q24 and encodes a 64 kD phosphoprotein found in the nucleus which is thought to be involved in DNA replication possibly by binding to regions on the DNA that serve as initiation sites for DNA synthesis. It may also affect the processing or half life of other proteins that are essential for cellular proliferation. Although c-myc expression is correlated with entry and exit from the cell cycle its exact role in cell proliferation remains unclear. The majority of colorectal cancers have elevated expression of c-myc mRNA and increased c-myc oncoprotein expression without evidence of gene amplification or rearrangement (Stewart et al., 1986; Sikora et al., 1987).

Constitutively elevated expression of c-myc has been implicated as an initiating event in colorectal carcinogenesis in several studies. These studies suggest that c-myc deregulation is due to a defect in c-myc gene transcription resulting in synthesis of increased levels of c-myc mRNAs. The APC gene on chromosome 5q has been implicated as a possible gene that may deregulate c-myc transcriptional activity (Fearon and Vogelstein, 1990). There have been several studies correlating the level of c-myc expression in primary tumors with various clinicopathological features. These studies suggest that c-myc expression is more common in left sided than right sided colonic lesions; however, no association has been found between tumor stage or degree of differentiation. Similarly, no association has been detected between c-myc expression and disease free or overall survival (Watson et al., 1985; Stewart et al., 1986; Sikora et al., 1987). Collectively these studies suggest that c-myc expression may be independent of clinical prognostic parameters and the use of this gene as a prognostic molecular marker would not appear to be useful.

Other oncogenes which may play a role in colorectal tumorigenesis include the L-myc oncogene, located on chromosome 1p32, the bcl-2 oncogene, first discovered due to its involvement in the 14;18 translocation
common in human follicular lymphomas and the cyclin D1 oncogene, first cloned because of its clonal rearrangements in parathyroid adenomas (Tsujimoto et al., 1985; Motokura et al., 1991). The L-myc gene, sometimes activated late in tumorigenesis, is expressed by three genotypes in the population, LL, LS and SS, and previous studies have suggested that the presence of the S-allele of the L-myc gene in the DNA of patients with cancer is associated with a higher risk of metastasis in lung, breast, and renal cell carcinomas (Nau et al., 1985). Young et al. examined the relation of the S-allele to metastasis and the suppression of tumor suppressor genes in colorectal cancer (Young et al., 1994). Of 124 colorectal cancer patients studied for the L-myc genotype, 19% of Dukes A; 19% of Duke’s B; 25% of Dukes C, and 40% of Dukes D had the SS genotype. In addition, the L-myc genotype was associated with a significantly higher level of loss of sequences from 18q, in that all SS patients had 18q deletions in their tumors, possibly as a result of the inactivation of the tumor suppressor gene DCC. This study suggests that the S-allele of the L-myc gene in colorectal cancer, possibly plays a role in modifying the development of colorectal cancer. Another potential oncogene, the bcl-2 gene, is a known inhibitor of apoptosis and is differentially expressed in the colonic crypt, suggesting a role for bcl-2 in the regulation of apoptosis in colonic epithelium. It is thought that the loss of bcl-2 expression may be a normal event during colonic epithelial cell differentiation, and that deregulation of this gene may be an early event in colorectal carcinogenesis through the promotion of benign and malignant colorectal tumors (Hague et al., 1994; Sinicrope et al., 1995). In a study investigating the relationship of the bcl-2 gene and p53, Sinicrope surmised that there is a potential down-regulation of bcl-2 by mutant p53 in premalignant polyps, suggesting that they may regulate a common cell pathway. In addition, cells expressing bcl-2 or mutant p53 are resistant to inducers of apoptosis, such as radiation and some chemotherapeutic agents, including 5-FU, an agent widely used in the treatment of colorectal cancers. The overexpression of the bcl-2 protein in colorectal cancers suggests a genetic basis for clinical drug resistance. Cyclin D1, localized on chromosome 11q13, positively regulates transition through the G1 phase of the cell cycle and its function is critical for both normal diploid cells. Cyclin D1’s oncogenic properties have been demonstrated through its deregulation, which disrupts the cell cycle by shortening the length of the G1 phase (Jiang et al., 1993; Quelle et al., 1993). A relationship between D1 and H-ras has also been recently reported with the demonstration that overexpressed cyclin D1 cooperates with an H-ras oncogene and transforms primary cells into tumorigenic cells (Lovec et al., 1994). In a study by Bartkova, analyzing cyclin D1 expression in normal/non-involved, primary and metastatic colorectal carcinomas by immunohistochemical staining, over 20% of human colorectal carcinomas accumulated high levels of the cyclin D1 oncoprotein, indicating that the cyclin D1 protooncogene may be deregulated in a significant subset of colorectal tumors (Bartkova et al., 1994).

**Tumor associated antigens**

Carcinoembryonic antigen (CEA) is a tumor associated antigen which can be measured in patients serum and is used in the clinical diagnosis and prognosis of colorectal carcinoma (Wanebo et al., 1978, 1989; Minton et al., 1985; Northover, 1986). It is produced by colon carcinoma cells and plays a role in cellular recognition and cell attachment. Transfection studies have demonstrated that CEA expressing cells aggregate with each other and promote metastatic spread by binding to CEA receptors which are heavily expressed in the liver and lung. When athymic nude mice bearing human colon xenografts are pretreated with CEA they rapidly develop metastatic lesions in the liver and lung. Preoperative CEA levels are elevated in 40-60% of patients with colorectal cancer. Increased CEA levels correlate with the diagnosis of well differentiated tumors; however, the antigen is also expressed in 30% of patients with poorly differentiated tumors. Elevated serum CEA levels are also found in smokers and in patients with benign liver, breast and lung diseases (Clarke et al., 1980; Tomoda and Furusawa, 1981). While it was originally hoped that it might be useful as a screening test CEA has proved to be too insensitive and non-specific for detection of the vast majority of early stage colorectal cancers (Tomoda and Furusawa, 1981). Resection of the primary tumor usually results in the CEA returning to normal levels. The reappearance of increasing CEA levels after surgery usually heralds the onset of tumor recurrence (Mach et al., 1974; Wood et al., 1980; Lunde and Havig, 1982; Denstman et al., 1986). Several investigators have suggested that postoperative CEA monitoring may be used as a guide for second look surgery which they suggest may improve the cure rates for as many as 20% of patients and decrease mortality by 5% (Denstman et al., 1986). However, a recent study by Moertel and colleagues have shown this approach to be of no value except for patients with potentially resectable hepatic metastasis (Moertel et al., 1993). Collectively these studies suggest that patients with higher levels of CEA tend to have a worse disease free survival and overall survival; however, when compared to other markers such as disease stage, CEA loses its significance. Thus, the use of CEA as an important adjunct to colorectal cancer is confined to monitoring of patients with known colorectal carcinoma.

Cytokeratin antigens are an integral part of the cytoskeleton that are widely expressed on human colon carcinoma cells. Using an antibody to the cytokeratin antigen (CK18) Lindemann and coworkers looked for disseminated colon cancer cells in the bone marrow of 88 patients with Dukes C colorectal cancer post surgical resection. Twenty eight percent of patients had positive CK18 cells in the bone marrow (Lindemann et al.,
The molecular biology of colorectal cancer

1992). These patients had a significantly poorer disease free and overall survival both by univariate and multivariate analysis. While the bone marrow is not a usual site for metastatic disease of the colon this study points to the presence of disseminated disease at the time of surgery beyond the confines of the abdominal lymph nodes and suggests that this may be of prognostic significance. Another novel and interesting randomized study by Reithmuller and colleagues (1994) examined the effect of a mouse monoclonal antibody directed against the cytokeratinic antigen 17-1A as adjuvant immunotherapy in 166 patients with resected Dukes C colorectal carcinoma. This monoclonal antibody was used to target minimal residual disease in these patients. A previous phase I study had shown that this antibody was well tolerated and induced an antibody dependent cytotoxic response in patients. Patients were randomly assigned to surgery alone or surgery followed by post operative treatment with 17-1A once a month. After a follow up of 5 years the group that received adjuvant antibody therapy had a significant improvement in disease free and overall survival compared to those randomized to surgery alone. Antibody therapy led to a decrease in the overall death rate by 30% and recurrence rate by 27%. Eight percent of patients who received the antibody developed a human antibody response against the murine antibody but this did not differ between responders and non-responders. This study is the first to suggest that antibody therapy directed at cytoskeletal antigens expressed by colorectal cancer cells may have a therapeutic role in patients with this disease. Other groups have begun to investigate the efficacy of immunotherapy using the anti-idiotyp antibody 3H1, which mimics an epitope on CEA, as well as a monoclonal anti-idiotyp vaccine against colon cancer cells using CEA expressing recombinant vaccinia virus (Foon et al., 1994; Hamilton et al., 1994). These immunotherapeutic approaches are at present being investigated in patients with advanced metastatic disease and further study will be needed to determine their role.

Mucins related antigens

Colonic epithelium is characterized by the production of mucus the main component of which is mucin. Mucins are a family of glycoproteins which are thought to lubricate the epithelial lining of the gastrointestinal tract. They also serve as a protective barrier and promote intercellular adhesion and attachment. They are composed of oligosaccharide side chains connected by O-glycosidic linkages to serine and threonine amino acid residues. The carbohydrate portion of the glycoprotein express a variety of specific antigenic determinants including the Tn, sialosyl-Tn and T antigens which can be identified and distinguished using specific monoclonal antibodies. Expression of the Tn, and sialosyl-Tn mucin glycoprotein has been identified as one of the earliest steps in colorectal carcinogenesis. The sialosyl-Tn antigen is a simple disaccharide structure that is not expressed in normal colonic mucosa and only rarely in hyperplastic polyps. However, at least 50% of adenomatous polyps (>1 cm) express this antigen, the level of which increases in dysplastic polyps (Itzkowitz et al., 1989). Approximately 90% of colonic cancers express sialosyl-Tn antigen. In a retrospective study Itzkowitz and co-workers examined the expression of sialosyl-Tn antigen in primary colorectal cancers to determine its prognostic significance (Itzkowitz et al., 1990). Of the 128 samples analyzed 86% were positive for Sialosyl-Tn antigen expression. This expression was found to be independent of tumor location, Dukes stages, histological type, depth of invasion, and DNA content. By multivariate analysis sialosyl-Tn antigen expression was an independent prognostic factor for disease free survival and survival. Those patients with tumors that were sialosyl-Tn antigen negative had a 100% survival and 94% disease free survival after 5 years; whereas, in patients whose tumors were sialosyl-Tn antigen positive, the 5 year survival was 73%. The prognostic effect of sialosyl-Tn antigen appears specific since expression of two closely related mucins (T and Tn) did not correlate with prognosis. Sialosyl-Tn is the first mucin associated antigen whose expression by colon cancer cells is independently associated with prognosis. The reason for the correlation of Sialosyl-Tn antigen expression and poor prognosis is not clear and will need further investigation. If these observations are confirmed in subsequent prospective studies Sialosyl-Tn antigen detection may help define a subgroup of colorectal cancer patients with a more aggressive clinical course. Furthermore, the recent cloning of the mucin gene may help provide further insights into the biological role of mucin glycoproteins in colorectal cancer.

Conclusion

Advances in understanding the biology of colon cancer have progressed rapidly over the last several years. The improvements in our understanding of colorectal cancer biology have identified potential markers that may prove to be not only of prognostic value but also important as screening tools and therapeutic targets. Biological properties such as the RER phenotype, tumor suppressor genes, allelic loss of 18q, 17p and other chromosomal alleles may eventually be of use as screening tools or markers of prognosis and outcome to chemotherapy. The understanding of the biology of colorectal cancers will now need to be applied in the clinical setting.

References

The molecular biology of colorectal cancer


The molecular biology of colorectal cancer
N-methyl-N-nitro-N-nitrosoguanidine tolerance in colon tumor cells with homozygous hMLH1 mutation. Cancer Res. 54, 4308-4312.


Nagase H., Miyoshi Y., Horii A., Aoki T., Ogawa M., Utsunomiya J., Baba S., Sasazuki T. and Nakamura Y. (1992a). Correlation between the location of germ-line mutations in the APC gene and the number of colorectal polyps in familial adenomatous polyposis
The molecular biology of colorectal cancer

patients. Cancer Res. 52, 4055-4057.


Spirio L., Otterud B., Stauffer D., Lynch H., Lynch P., Watson P.,
The molecular biology of colorectal cancer


