Genetic and epigenetic biomarkers for diagnosis, prognosis and treatment of colorectal cancer

Fabio Coppedè, Angela Lopomo, Roberto Spisni, Lucia Migliore

Abstract

Colorectal cancer (CRC) is one of the most common cancer worldwide and results from the accumulation of mutations and epimutations in colonic mucosa cells ultimately leading to cell proliferation and metastasis. Unfortunately, CRC prognosis is still poor and the search for novel diagnostic and prognostic biomarkers is highly desired to prevent CRC-related deaths. The present article aims to summarize the most recent findings concerning the use of either genetic or epigenetic (mainly related to DNA methylation) biomarkers for CRC diagnosis, prognosis, and response to treatment. Recent large-scale DNA methylation studies suggest that CRC can be divided into several subtypes according to the frequency of DNA methylation and those of mutations in key CRC genes, and that this is reflected by different prognostic outcomes. Increasing evidence suggests that the analysis of DNA methylation in blood or fecal specimens could represent a valuable non-invasive diagnostic tool for CRC. Moreover, a broad spectrum of studies indicates that the inter-individual response to chemotherapeutic treatments depends on both epigenetic modifications and genetic mutations occurring in colorectal cancer cells, thereby opening the way for a personalized medicine. Overall, combining genetic and epigenetic data might represent the most promising tool for a proper diagnostic, prognostic and therapeutic approach.

INTRODUCTION

It is now clear that cancer is a multi-step process resulting
from the accumulation of both genetic and epigenetic alterations of the genome\[9\]. Gene mutations and epigenetic modifications have been initially viewed as two separate mechanisms participating in carcinogenesis. However, recent evidence points to a crosstalk between these two mechanisms in cancer formation, suggesting that gene mutations have the potential to disrupt DNA methylation patterns, and epigenetic modifications can drive genome instability and mutagenesis\[2,3\]. For example, the whole exome sequencing of thousands of human cancers revealed unexpected mutations in genes involved in epigenetic mechanisms, and those mutations have the potential to disrupt DNA methylation patterns, histone modifications, and nucleosome positioning\[8\]. Similarly, epigenetic inactivation of DNA repair genes, such as hMLH1, hMSH2, MGMT and BRCA1, is often associated with genome instability and increased frequency of point mutations of cancer-related genes\[9\].

Colorectal cancer (CRC) is one of the most frequent cancers in humans, with over one million new cases diagnosed worldwide every year\[9\]. The disease occurs sporadically in most of the cases (75%-80%) as a result of the accumulation of both mutations and epigenetic modifications of several genes\[5\], and large-scale DNA methylation studies suggest that CRC can be divided into at least three-four subtypes according to the frequency of DNA methylation and those of mutations in key CRC genes\[6,7\]. The sequential process of gene mutations and epigenetic alterations is believed to drive the progression toward malignant adenocarcinomas because those events affect signalling pathways that regulate hallmark behaviours of cancer. Gene mutations create a clonal growth advantage that leads to the outgrowth of progressively more malignant cells, which ultimately manifests itself as invasive adeno-carcinoma. The 5-year survival rates are approximately 90% for early CRC patients but decrease to less than 10% in patients with distant metastases, by this the need to identify biomarkers to improve the prediction of clinical outcomes in CRC\[8\]. Further progress is very much desirable in non-invasive diagnostic methods to enable early diagnosis, pre- and postoperative staging, and to assist in selecting the most suitable neo-adjuvant and adjuvant therapeutic methods and post-treatment. Novel biomarkers which are absent in healthy persons and present in CRC are still being investigated, especially those that can be detected at early development stage of the disease and used in screening tests. Unfortunately, no molecule that would meet all of the foregoing criteria has been identified so far. Carcinoembryonic antigen still remains the only tumour marker of recognised efficacy in monitoring patients during and after CRC therapy\[9\].

There is an increasing interest to identify mutations in key genes of tumourigenesis, such as APC, CTNNB1, BRAF and KRAS because they are involved in the Wnt and the Ras-Raf-MEK-MAPK signalling cascades (MAPK, mitogen-activated protein kinase; MEK, MAPK/ERK kinase) and therefore play a substantial role in the adenoma-carcinoma and in the serrated adenoma pathways. There are also attempts to “personalise” chemotherapy based on presence or absence of specific genetic biomarkers. For example, therapy with anti-EGFR (epidermal growth factor receptor) antibodies is desirable in patients with advanced CRC and absence of KRAS\[6\] or BRAF\[7\] mutations, and defining tumours phenotype - microsatellite instability (MSI) or microsatellite stability (MSS) and testing for the presence or absence of 18q chromosome deletion is very much desirable in standard 5-fluorouracil (5-FU)-based therapy\[9,10\].

DNA methylation represents one of the most studied epigenetic marks in CRC\[11\], since methylation of CpG islands in the promoter region of a gene might induce chromatin conformational modifications and inhibit the access of the transcriptional machinery, thus altering gene expression levels. Promoter hypermethylation is commonly associated with gene silencing as well as promoter demethylation with gene expression. The ever-growing number of genes that show epigenetic alterations in cancer emphasizes the crucial role of these epigenetic alterations, and particularly of DNA methylation, for future diagnosis, prognosis and prediction of response to therapies\[12\]. Lao et al\[11\] (2011) reviewed the genes that seem to be more commonly methylated in the multi-step process leading from normal colonic epithelium to adenocarcinoma, observing that some of them are frequently methylated in the passage from a normal colon epithelium to an abberant crypt focus, whilst others are methylated in the passage from an abberant crypt focus to polyp/adenoma, or could have a role in CRC progression and metastasis. Concerning CRC diagnosis, there is increasing interest in searching for aberrantly methylated genes in plasma DNA and in the DNA obtained from faecal material, as non-invasive diagnostic tools\[8,13,14\]. Methylation of certain genes, such as for example those involved in the extracellular matrix (ECM) remodelling pathway, were associated with worse survival in CRC, suggesting that epigenetic biomarkers could gain prognostic value\[8\]. There is also active research focusing on epigenetic signatures in CRC for their possible interaction with chemotherapeutic agents\[14\].

Given the enormous potential of both gene mutations and DNA methylation biomarkers in CRC diagnosis, staging, prognosis and response to treatment, active research is currently ongoing to develop rapid, cost-effective and reproducible tools for the detection of those markers\[12\]. Aim of this article is to review currently available genetic and DNA methylation biomarkers for CRC diagnosis, staging, prognosis and treatment.

**GENETIC BIOMARKERS IN CRC**

**Genetic and cytogenetic biomarkers**

In 1990, Fearon and Vogelstein proposed a model for colorectal cancer tumourigenesis, which defines the genetic alterations involved in transformation from normal intestinal mucosa to colorectal carcinoma. This aberrant transformation is a multi-step process that includes genet-
ic alterations such as mutation of the \textit{APC} (adenomatous polyposis coli gene), located on chromosome 5q, which is thought to occur early on during the development of adenomatous polyps, the activation of \textit{KRAS} (v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog gene), an oncogene located on chromosome 12p12, during the adenomatous stage and loss of chromosomal regions 17p and 18q that contain tumoural suppressor genes as tumour protein p53 (TP53) and \textit{DCC} (deleted in colorectal carcinoma), in the transition to carcinoma \textit{in situ}\cite{27}. A lot of studies, by different approaches, identified these common alterations described by Fearon and Vogelstein and, in addition, others changes such as gain of chromosomes 7, 8q, 13q and 20q, together with loss of the 1p, 4p and 22q chromosomal regions were also identified\cite{28,29}. Some studies have suggested that the loss of heterozygosity (LOH) of 17p and 18q could be associated with more advanced stages of the disease; the loss of 17p and 18q are believed to play an important role in the pathogenesis of CRC since these two chromosomes carry genes relevant to the malignant transformation of the gut epithelium and also probably play an important role in the metastatic process. In this regard, recent findings showed that breakpoints in the 17p11.2 chromosomal region were preferentially found in primary colonic tumours in CRC patients with liver metastases\cite{30,32}. The deletion of the long arm of chromosome 18 (loss of 18q or LOH of 18q) is the most common cytogenetic abnormality in CRC and seems to be associated with poor prognosis as 18q contains several important tumour suppressor genes, such as \textit{SMAD7}, \textit{SMAD4}, and \textit{SMAD2} that are transcriptional mediators in the TGF-\beta signalling pathway and \textit{DCC}\cite{34,35}. Mouse studies demonstrate that loss of \textit{SMAD4} expression changes the role of TGF-\beta from growth suppressor to growth promoter, thus increasing the tumorigenic and metastatic potential of colorectal cancer cells\cite{35}. Loss of SMAD activity occurs in 10\% of the colorectal cancers and is associated with advanced-stage disease, the presence of lymph node metastases and shorter overall survival and it has been shown to be a significant independent prognostic factor for worse recurrence-free and overall survival, particularly in patients with stage Ⅲ disease. Patients with stage Ⅲ disease and intact \textit{SMAD4} expression with microsatellite instability were found to have similar outcomes compared with patients with stage Ⅱ disease, whereas patients with stage Ⅱ disease and loss of \textit{SMAD4} expression without microsatellite instability status had outcomes similar to patients with stage Ⅲ disease\cite{36}. Retention of \textit{SMAD4} expression has also been found to be a predictive marker for a threefold increase in benefit from 5-FU-based chemotherapy\cite{37} while the loss of \textit{SMAD4} seems to be a predictive marker for a poorer response to 5-FU\cite{38}. So this chromosome instability (CIN) could have a prognostic value, as patients with CIN+ disease have a poorer prognosis\cite{39}.

**Microsatellite Instability**

Over the CIN, another form of genomic instability frequent in CRC is the microsatellite instability, observed at the nucleotide level, frequently resulting in deletions or insertions of a few nucleotides. Microsatellites are polymorphic tandem repeats of short nucleotide sequences distributed through the genome prone to frame shifts and base-pair substitutions during replication if DNA mismatch repair (MMR) genes are impaired. So MSI refers to a clonal change in the number of repeated DNA nucleotide units in microsatellites and appears in tumours with deficient mismatch repair due to the inactivation of the four MMR genes: \textit{MSH2}, \textit{MLH1}, \textit{MSH6} and \textit{PMS2} and while it is typically associated with hereditary non-polyposis colorectal cancer (HNPCC), most MSI-high tumours occur sporadically\cite{40}. Sporadic MSI tumours tend to be more proximal, to occur in older females, to be poorly differentiated and mucinous, and to show marked lymphocytic infiltration\cite{41,35}. Despite their resistance to alkylating agents and cisplatin, MSI-high tumours have better recurrence-free and overall survival. In patients with stage Ⅱ disease, MSI-high status was found to confer the same advantage in long-term outcomes as that conferred by stage T3 over T4\cite{42}. MSI positive tumours are associated with a better prognosis in all stages of the disease. Patients with MSI tumours have a significant survival advantage compared with patients with non-MSI tumours\cite{43} and are associated with resistance to 5-FU chemotherapy and shorter survival of patients after treatment with the drug\cite{34,35}. MSI can be thus seen as one of the most promising positive prognostic markers for CRC patients and can be detected using a panel of five markers (\textit{BAT25}, \textit{BAT26}, \textit{D2S123}, \textit{D5S346}, and \textit{D17S2720}, particularly analyzing this 5 loci, MSI-H is defined as instability at 2 loci or more, and MSI-L, as instability at 1 locus) and a recent study using this panel observed that the presence of MSI-H was significantly higher in carcinomas than in adenomas, confirming the prognostic value of MSI in CRC\cite{44}.

**APC gene**

Mutations in the \textit{APC} gene are responsible for familial adenomatous polyposis (FAP) and the majority of sporadic CRC. The \textit{APC} gene encodes a multifunctional protein with important roles in Wnt signaling pathway, intercellular adhesion, cytoskeleton stabilization, cell cycle regulation, and apoptosis. Mutations of \textit{APC} may lead to unregulated transcription of oncogenes such as c-myc and cyclin D1, thereby promoting tumourigenesis. Mutations in \textit{Wnt/APC/CTNNB1 (b-catenin)} signalling pathway members have been found in many CRC and more than 90\% of patients have alterations that affect it. In light of the critical role of the \textit{Wnt/APC/CTNNB1} signalling pathway in maintaining proper colorectal cell function, it is possible that genetic variants in this pathway might affect CRC progression. A meta-analysis provides a complete and systematic picture of the role of three \textit{APC} polymorphisms (D1822V, E1317Q, H1307K) in the risk of colorectal neoplasia, particularly the H1307K variant was associated with a significantly increased risk
for colorectal neoplasia while the E1317Q one was associated with a significantly elevated adenoma risk. This meta-analysis may provide genetic insight into possible strategies for the prevention of colorectal neoplasia\(^{[8]}\).

A very recent study, applying a comprehensive approach to systematically evaluate the tag single-nucleotide polymorphisms (tSNPs) in two key genes of the Wnt pathway, APC and CTNNB1, identified, by survival tree analysis, a higher-order genetic interaction profile consisting of the APC rs565453, CTNNB1 rs2293303 and \(APC\) rs1816769 and this was significantly associated with overall survival; these SNPs might influence APC/CTNNB1 splicing and expression by altering the consensus splicing site sequences, the transposable elements, and the transcription factor binding sites. If validated, these biomarkers might be valuable to facilitate the identification of good treatment\(^{[8]}\).

Recently, it was found that, in advanced-stage cancer, patients with \(APC\) mutation/high miR-21, an activator of the Wnt signaling pathway, had poorer overall survival so \(APC\) mutation and miR-21 expression could be used to predict the clinical outcome of CRC\(^{[8]}\).

**KRAS gene**

The status of \(KRAS\) is generally accepted as a predictive marker for response to established EGFR inhibitors used for CRC because mutant \(KRAS\) is associated with resistance to anti-EGFR monoclonal antibody (mAb) immunotherapy with agents such as cetuximab or panitumumab (Figure 1). It is the only established biomarker in clinical practice for CRC. The \(KRAS\) encodes a 21-kD protein (p21ras) involved in the G-protein signal transduction pathway, modulating cellular proliferation and differentiation. \(KRAS\) abnormalities are one of the earliest events in the stepwise progression of colorectal neoplasms, being detectable even in histologically unremarkable epithelium and aberrant crypt foci adjacent to cancers. Mutations of the \(KRAS\) oncogene result in constitutive activation of this signal transduction pathway and, consequently, unregulated proliferation and impaired differentiation\(^{[9]}\). The \(KRAS\) wild-type (WT) protein is transiently activated during tightly regulated signal transduction events. The binding of mAbs to EGFR normally induces receptor internalization, causing a direct inhibition of tyrosine kinase activity and the blockage of downstream RAS/RAF/MAPK signalling (Figure 1). However, activating \(KRAS\) mutations result in constitutively active GTP-bound protein which consequently renders the downstream pathway permanently “switched on” irrespective of the activation status of upstream receptors including EGFR. In such an instance, the binding of an anti-EGFR mAb to EGFR and the inhibition of ligand-mediated receptor activation will fail to elicit any pathway suppressive effects. This constitutive pathway activation leads to unregulated proliferation, impaired differentiation, and resistance to anti-EGFR therapies\(^{[9]}\) (Figure 1). Up to 90% of activating mutations of \(KRAS\) are detected in codons 12 (82%-87%) and 13 (13%-18%), but less frequently in codons 61, 63 and 146 and they are generally observed as somatic mutations. The most common types of \(KRAS\) mutations in CRC are point mutations, particularly G > A transitions and G > T trans-
versions. The codons 12 and 13 code for two adjacent glycine residues located in the proximity of the catalytic site. Specific KRAS mutations may be heterogeneous in their phenotype. For example, codon 12 mutations were associated with a mucinous phenotype of CRC. By contrast, CRCs associated with codon 13 mutations were rather non-mucinous, but were characterized as more aggressive tumours with a greater metastatic potential. However, a recent study observed that patients with an isolated p.G12A mutation (no other KRAS mutations) had aggressive disease (stage III or IV and extensive metastatic or recurrent disease). KRAS mutations have also emerged as a major predictor or resistance to anti-EGFR mAbs, as confirmed by small data sets, and both retrospective and prospective trials, and both retrospective and prospective trials, and both retrospective and prospective trials, and both retrospective and prospective trials, and both retrospective and prospective trials.

In these studies, patients with metastatic CRC harbouring KRAS mutations had no benefits from treatment with cetuximab or panitumumab either alone or in combination with standard chemotherapy. This discovery led to the first practical implementation of personalized medicine in metastatic CRC, and KRAS mutations can be considered a highly specific negative biomarker for benefit of anti-EGFR mAbs. However it is intriguingly now coming to light that not all KRAS mutations are equal in their biological characteristics and their impact on mediating EGFR resistance, and that not all KRAS mutations will confer resistance to EGFR inhibitor therapy, probably due to heterogeneity of tumours.

Other clinicopathological and prognostic biomarkers of CRC

BRAF (V-raf murine sarcoma viral oncogenes homolog B1) is a member of the R-AF viral oncogenes family, it encodes a serine-threonine protein kinase, a downstream effector of activated RAS. In the past decade, many studies have shown that BRAF somatic mutation presents in approximately 10% of CRC. A hotspot for BRAF mutation is the conversion of valine 600 to glutamic acid (V600E) within the kinase activation domain of the BRAF protein and this account for 80% of the BRAF mutations in CRC. This hot spot is suggested to be biologically distinct from other infrequent BRAF mutations, because the cancer cells having the V600E mutation can grow without functional RAS, and thus the BRAF-V600E mutation has not been found in CRCs with KRAS mutations.

BRAF mutations have been linked with high grade, right side tumours, female gender, older age and MSI-H tumours. A distinct pattern of metastatic spread has also been observed in BRAF mutant tumours, namely higher rates of peritoneal metastases, distant lymph node metastases and lower rates of lung metastases. A very recent study demonstrated sex-related differences in the prognostic value of BRAF mutations in CRC, being particularly evident in men, in fact, BRAF mutation was associated with a significantly reduced cancer-specific survival in overall adjusted analysis.

Phosphatidylinositol-3-kinases (PI3K) are lipid kinases that promote various biological processes including cellular proliferation and survival. Mutations in the PIK3CA gene, which encodes the p110α catalytic subunit of PI3K, have been identified in many human solid tumours. In colorectal cancers, PIK3CA mutations, which are found in 10%-20% of the cases, have been reported to be associated with specific clinicopathological features and molecular events, tumour proximal colonic location, mucinous differentiation, KRAS mutation, high levels CIMP and loss of MGMT expression. It is unclear whether PIK3CA mutation defines a clinically and/or biologically relevant subset of tumours as there is significant overlap with KRAS and BRAF V600E mutation; a recent study observed that the adverse prognostic effect of PIK3CA mutation on survival was restricted to patients with a BRAF wild type tumour. The majority of activating PIK3CA mutations map to three sites: exon 9, codons 542 and 545 in the helical domain, and exon 20, codon 1047 in the kinase domain. Mutation at any one of these sites has been shown to result in a gain of enzymatic function and to promote oncogenic transformation in vitro and in vivo (Figure 1). Co-existence of PIK3CA exon 9 and 20 mutations is associated with poor prognosis of CRC patients. More recently, PIK3CA mutation was associated with longer survival in patients who use aspirin regularly after diagnosis.

Loss of PTEN expression (PTEN) is a key tumour suppressor gene involved in the homeostatic maintenance of PI3K/AKT signalling was associated with a higher rate of distant metastasis. However, patients with PTEN expression had significantly longer overall survival than patients with PTEN loss tumour.

TP53 is a tumour suppressor gene encoding a protein involved in the regulation of cell division, growth arrest and apoptosis. A recent study demonstrated that p53 expression was a significant prognostic factor for disease-free survival for the patients with stage III tumour and also found stage III tumours with wt-p53 high expression were associated with a significantly better prognosis after chemotherapy, according to previous findings.

A very recent study, exploring candidate tumour suppressor genes at chromosome 4q25-q28.2, found a novel candidate tumour suppressor gene, namely NDST4, identified at 4q26. This gene was markedly downregulated in CRC tumours and this genetic aberration was increased considerably in tumours with higher pathological stages (T3 and T4). NDST4 is one member of the N-deacylase/N-sulfotransferase (heparan glucosaminyl) (NDST) family, which is responsible for heparan sulfate (HS) biosynthesis on a core protein to form heparan sulphate proteoglycans (HSPGs) that contribute to the tissue structure and function during development and adult homeostasis. The loss of function of NDST4 might impair the modification of HS chains of specific HSPGs, leading to more invasive tumour cells through remodelling of the interaction of cell adhesion receptors and ligands. The genetic loss of NDST4 might serve as a biomarker of adverse prognosis for patients with CRC.
Genetic biomarkers of response to treatment

During the last few years, chemotherapeutic agents, such as oxaliplatin, irinotecan, cetuximab, panitumumab, bevacizumab, afibercept and regorafenib, have been approved as an addition to the traditional fluorouracil (5-FU) treatment, increasing the median overall survival. The survival of patients with metastatic CRC (mCRC) progressively improved over the past decades was due primarily to new chemotherapeutic combinations (5-FU, irinotecan, oxaliplatin) and the introduction of new therapies, among which there are two monoclonal antibodies against the receptor of epidermal growth factor receptor, cetuximab and panitumumab, effective in the treatment of mCRC. However, these treatments are toxic and expensive, by this, the necessity to select patients most likely to have benefit with the treatment. Several analyses have revealed that patients with KRAS mutations receiving first and subsequent lines of treatment do not respond to cetuximab or panitumumab, and that they show no survival benefit from such treatments (Figure 1). Therefore, patients with mCRC with KRAS codon 12 or KRAS codon 13 mutated tumours are presently excluded from treatment with anti-EGFR mAb. Recent study has demonstrated, for the first time, that KRAS wt status is associated with better response to bevacizumab based chemotherapy and represents a positive prognostic factor for patients with advanced CRC treated in the first-line setting[23]. Codon 12 KRAS and B-RAF mutations predict for adverse outcome of CRC patients receiving cetuximab[23]. KRAS mutations are, also, predictive of resistance to anti-EGFR antibodies when combined with irinotecan[17], response negatively affected also by NRAS, B-RAF and PI3KCA mutations[49,57,79], as well as by a wild-type TP53[50]. So KRAS status has emerged as a major predictor of resistance to anti-EGFR mAb in the clinical setting but probably it is not the only to determine this type of resistance, in fact, a recent meta-analysis showed that BRAF mutation is associated with poor response to anti-EGFR mAbs and it is an adverse prognostic biomarker of the survival of patients with mCRC[37]. Regular aspirin use was associated with lower risk of BRAF wt colorectal cancer but not with BRAF mutated cancer risk. These findings suggest that BRAF mutant colon tumour cells may be less sensitive to the effect of aspirin[39]. In addition to KRAS and BRAF mutations, loss of PTEN expression and PIK3CA mutation is likely to be predictive of a lack of benefit to anti-EGFR therapy in metastatic colorectal cancer[39] (Figure 1); PTEN expressing tumours had statistically higher response rate for cetuximab based treatment than tumours with PTEN loss[40].

CRCs with MSI are reported to have a significantly better prognosis compared with CRCs without MSI (non-MSI), while MSI CRCs show resistance to 5-FU based chemotherapies. Although high frequency MSI tumours have better stage independent prognosis compared to those with CIN, MMR deficient CRC appears to be resistant to fluorouracil based treatment, but sensitive to other therapeutic regimens[40]. A summary of genetic biomarkers for CRC is shown in Table 1.

DNA METHYLATION BIOMARKERS IN CRC

Global DNA hypomethylation and depletion of overall 5-methylcytosine content in CRC tissues was observed for the first time in 1983, by Feinberg and Vogelstein. It was observed predominantly at CpG dinucleotides in repetitive sequences, occurring gradually, age-dependently, and early in carcinogenesis[81]. It was also clear from the beginning that global DNA hypomethylation in CRC tissues was accompanied by hypermethylation and transcriptional silencing of tumours suppressor genes or genes coding for DNA repair proteins[82]. Subsequent studies revealed that hundreds of genes are aberrantly methylated in the average CRC genome, and their number is ever-growing, including genes of the Wnt signalling pathway such as APC, AXIN2, DKK1, SFRP1, SFRP2, WNT5A, the DNA repair genes MGMT, hMLH1, and hMLH2, cell cycle-related genes such as CDKN2A/p16[14], CDKN2A/p14[15,16], the RAS signalling genes RASF1A and RASF1B, and many more[80,83,84]. Although all CRCs are characterized by the presence of hypermethylation, a specific subgroup of them, denoted as the CpG island methylator phenotype (CIMP+), displays extensive levels of methylated genes[85]. By an epigenetic point of view, CRCs can be broadly divided into CIMP+ and non-CIMP tumours, but taking into account also genetic alterations, several subgroups have been proposed. For example, Hinoue and coworkers recently proposed the following four subtypes: (1) CIMP-high tumours exhibiting a very high frequency of cancerspecific DNA hypermethylation associated with MLH1 methylation, microsatellite instability, and the BRAF V600E mutation; (2) CIMP-low tumours associated with KRAS mutations and characterized by methylation of a subset of CIMP-high associated genes; (3) Non-CIMP tumours characterized by TP53 mutations and frequent occurrence in the distal colon; and (4) Non-CIMP tumours showing a low frequency of cancer-specific gene mutation and hypermethylation, and enriched of rectal tumours[84]. Increasing evidence suggests that several of those epigenetic modifications can be valuable biomarkers for CRC diagnosis, progression, prognosis, tendency to metastasis, and response to treatment (Table 2).

Methylation biomarkers of CRC diagnosis

Aberrant patterns of DNA methylation from CRC cells can be detected in tumours-derived cell-free DNA found in blood or feces of cancer patients, and there is also evidence that often DNA methylation profiles in blood reflect those in CRC tissues[80]. This led researchers to search for DNA methylation biomarkers in those specimens and to develop blood-based and stool-based non-invasive and cost-effective epigenetic CRC diagnostic tools[87]. The presence of aberrantly methylated septin 9
(SEPT9) in plasma is a valuable and minimally invasive blood-based PCR test (Figure 2), showing a sensitivity and a specificity of almost 90% in the detection of CRC\cite{13,14,88}, and represents a currently commercialized test as it is able to detect CRC at all stages and locations\cite{94}. Researchers have evaluated the possibility to include the methylation analysis of additional genes, such as for example ALX4 and HLT\cite{95}, to increase the sensitivity of this blood-based test\cite{94,95}. Others are searching for different blood-based biomarkers than SEPT9. For example, the methylation status of secreted frizzled-related protein 2 gene (SFRP2) in CRC tissues, serum and fecal DNA was able to detect almost 67% CRCs\cite{93}, and recent genome-scale search of DNA-methylation biomarkers for blood-based detection of CRC revealed that methylated thrombomodulin (THBD) gene detects 74% of stage I/II CRCs at a specificity of 80%\cite{93}, and that methylation of the syndecan 2 (SDC2) gene has a sensitivity of 92% for stage I CRC\cite{94}.

A stool-based test for the methylation analysis of the vimentin (VIM) gene is available in the United States (Figure 2) and has a specificity and sensitivity of almost 80%\cite{14}. Several hypermethylated genes isolated from stool samples have been utilised as biomarkers for the detection of CRC or colorectal adenomas, including APC, p16, MLH1, MGMT, SFRP1, SFRP2 and VIM\cite{93}. Two meta-analyses of those studies revealed that the sensitivity for the detection of CRC or adenomas ranged from 62% to 75%\cite{93,94}. Recently, hypermethylation of fibrillin-1 (FBN1) was detected in CRC stool samples, and showed 72% sensitivity and 93% specificity for detecting CRC\cite{95}.

### DNA methylation biomarkers of CRC staging and prognosis

A few years ago, Lao and Grady reviewed the genes that seem to be more commonly methylated in the multistep process leading from normal colonic epithelium to adenocarcinoma. At least six genes (SLC5A8, SFRP1, SFRP2, CDH13, CRBP1, and RUNX3) and two loci (MINT1 and MINT31) have been consistently found to be methylated in the passage from a normal colon epithelium to an aberrant crypt focus. Other genes (p14, HLT, ITGA4, p16, CDH1 and ESR1) resulted frequently methylated in the passage from an aberrant crypt focus to polyp/adenoma, and four additional genes (TIMP3, CXCL12, ID4, and IRF8) could have a role in CRC progression and metastasis\cite{11}. More recent studies revealed additional epigenetic biomarkers linked to CRC staging and progression. A high degree of LINE-1 hypomethylation was found in early-onset CRC, a clinically distinct form of CRC that is often associated with a poor prognosis\cite{99}. LINE-1 hypomethylation leads to the activation of proto-oncogenes in CRC metastasis\cite{99}. There is also indication that the clinical outcome of MSI CRCs depends on LINE-1 methylation, suggesting that lower LINE-1 methylation status serves as a significant prognostic parameter of adverse prognosis\cite{93}. Serrated adenomas form a distinct subtype of colorectal pre-malignant lesions that may progress to malignancy along a different molecular pathway than the conventional adenoma-carcinoma pathway, and loss of expression of the slit homolog 2 (SLIT2) gene by promoter hypermethylation and loss of heterozygosity events are significantly associated with serrated adenoma development\cite{101}. Methylation of the WNT5A gene, a member of the WNT gene family, has been frequently detected in early gastric carcinomas\cite{102}. Also somatic mutations, allele loss, and DNA methylation of the cub and sushi multiple domains 1 (CSMD1) gene, whose function is still unclear, correlate with earlier clinical presentation in CRC\cite{103}.

Concerning CRC prognostic biomarkers, it was shown that DNA methylation of p14, RASSF1A and APC genes, defines a poor prognosis subset of CRC patients independently of both tumour stage and differentiation\cite{104}, whilst MGMT methylation seemed to play a protective role\cite{104}, and MLH1 inactivation through hypermethylation was found to be related to improved survival\cite{105}. A meta-analysis of 11 studies indicated that p16

---

**Table 1 Examples of genetic biomarkers for colorectal cancer**

<table>
<thead>
<tr>
<th>Biomarkers</th>
<th>Prognosis</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic biomarkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breakpoints of 17p11.2</td>
<td>Found in primary colonic tumours in CRC patients with liver metastasis</td>
<td>[22]</td>
</tr>
<tr>
<td>Loss of 18q</td>
<td>Poor prognosis</td>
<td>[23,24]</td>
</tr>
<tr>
<td>Loss of SMAD</td>
<td>Advanced stage disease (II), lymph node metastases, shorter overall survival</td>
<td>[26]</td>
</tr>
<tr>
<td>APC mutations</td>
<td>Poorer overall survival</td>
<td>[39]</td>
</tr>
<tr>
<td>KRAS mutations</td>
<td>Heterogeneous phenotype of CRC</td>
<td>[43,44,45]</td>
</tr>
<tr>
<td>BRAF mutations</td>
<td>Specific phenotype and metastasis</td>
<td>[61,62]</td>
</tr>
<tr>
<td>PIK3CA mutations</td>
<td>Poor prognosis and specific clinicopathological features</td>
<td>[64]</td>
</tr>
<tr>
<td>Loss of PTEN</td>
<td>High rate of distant metastasis</td>
<td>[68]</td>
</tr>
<tr>
<td>TP53 expression</td>
<td>Worse prognosis</td>
<td>[70,71]</td>
</tr>
<tr>
<td>Loss of NDRG4</td>
<td>Adverse prognosis</td>
<td>[72]</td>
</tr>
<tr>
<td>Candidate biomarkers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss of SMAD4</td>
<td>Poorer response to 5-FU</td>
<td>[28]</td>
</tr>
<tr>
<td>MSI</td>
<td>Resistance to 5-FU</td>
<td>[83]</td>
</tr>
<tr>
<td>KRAS, BRAF, PIK3CA, PTEN mutations</td>
<td>Resistance to anti-EGFR mAb</td>
<td>[46-52,74,75,79-81]</td>
</tr>
</tbody>
</table>

CRC: Colorectal cancer; mAb: Monoclonal antibodies; 5-FU: 5-fluorouracil.
Table 2 Examples of DNA methylation biomarkers for colorectal cancer diagnosis, progression, prognosis and treatment

<table>
<thead>
<tr>
<th>DNA methylation biomarkers</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methylated genes/loci</td>
<td></td>
</tr>
<tr>
<td>SLCSA8, SFRP3, SFRP2, CDH13, CRABP1, RUNX3, MINT1, MINT31, WNT5A</td>
<td>[11,105]</td>
</tr>
<tr>
<td>p14, HLTF, ITGA4, CDKN2A/p16, CDH1, ESR1</td>
<td>[11]</td>
</tr>
<tr>
<td>TIMP3, CXCL12, ID4, IRF8, MGMT, hMLH1</td>
<td>[11]</td>
</tr>
<tr>
<td>SPARC, miR-34b/c, miR-126, miR-128</td>
<td>[114-117]</td>
</tr>
<tr>
<td>Methylation biomarkers</td>
<td></td>
</tr>
<tr>
<td>SEPT9, SFRP2, THD8, SBC2</td>
<td>[91,97]</td>
</tr>
<tr>
<td>VIM, FN1</td>
<td>[14,100]</td>
</tr>
<tr>
<td>Methylation biomarkers</td>
<td></td>
</tr>
<tr>
<td>p14, RASSF1A, and APC</td>
<td>[107]</td>
</tr>
<tr>
<td>MGMT, hMLH1</td>
<td>[107,108]</td>
</tr>
<tr>
<td>p16</td>
<td>[109]</td>
</tr>
<tr>
<td>HOPX-β</td>
<td>[110]</td>
</tr>
<tr>
<td>Extracellular matrix genes (IGFBP3, EVL, CD109 and FLNC)</td>
<td>[111]</td>
</tr>
<tr>
<td>IGF2 hypomethylation</td>
<td>[112]</td>
</tr>
<tr>
<td>Polycrom genes (SFRP1, MYOD1, HIC1, and SLIT2)</td>
<td>[113]</td>
</tr>
<tr>
<td>miR-34b/c, miR-126, miR-128</td>
<td>[115-117]</td>
</tr>
<tr>
<td>Candidate biomarkers</td>
<td></td>
</tr>
<tr>
<td>TFAP2E</td>
<td>Chemoresistance/Chemosensitivity</td>
</tr>
<tr>
<td>DYPD, TYMP, UMPK, SPARC</td>
<td>[118]</td>
</tr>
<tr>
<td>LCTTA1</td>
<td>Their methylation might affect 5-FU treatment</td>
</tr>
<tr>
<td>MGMT</td>
<td>Clinical response to dacarbazine is restricted to those with MGMT</td>
</tr>
</tbody>
</table>

Suggested biomarkers from cell culture studies, with limited or no evidence in humans. CRC: Colorectal cancer; miR: micro-RNA; PCR: Polymerase chain reaction; 5-FU: 5-fluorouracil.

DNA methylation biomarkers and CRC chemotherapy

Epigenetic signatures in CRC are also of interest for their possible interactions with chemotherapeutic agents. Indeed, the epigenetic silencing of a particular gene might result in chemosensitivity or chemoresistance toward a particular therapeutic agent. Crea et al. proposed a panel of genes whose aberrant methylation could contribute to chemosensitivity or chemoresistance to 5-FU, irinotecan, and oxaliplatin, three of the most frequently used drugs in CRC treatment. 5-FU antitumor activity is mainly exerted by inhibiting thymidylate synthase, in the de novo synthesis of pyrimidines. Increased TMY5 expression is one of the major mechanisms of 5-FU chemoresistance, and there is indication that histone acetylation/deacetylation processes, rather than DNA methylation of the promoter, might be of relevance in epigenetically regulating TMY5 expression in CRC. Several other genes that participate in pyrimidine metabolism might represent potential molecular determinants of 5-FU chemoresistance, including dihydropyrimidine dehydrogenase (DYPD), thymidine phosphorylase (TYMP), and uridine monophosphate/cytidine monophosphate kinase (UMPK) genes. Their potential epigenetic contribution to 5-FU resistance in CRC patients is under investigation.

Hypermethylation of the gene encoding the transcription factor AP-2 epsilon (TF-AP2E) was found in 51% of CRC patients and resulted in clinical nonresponsiveness to chemotherapy (5-FU, irinotecan or oxaliplatin).

Hypermethylation might be a predictive factor for unfavourable prognosis of CRC patients. Homeodomain-only protein X-β gene (HOPX-β) promoter methylation was recently shown to be frequent in human cancers and was suggested to act as a tumours suppressor gene. Particularly, HOPX-β promoter methylation was associated with worse prognosis of stage III CRC patients and also with poor differentiation. Methylation of genes in the extracellular matrix (ECM) remodelling pathway, such as IGFBP3, EVL, CD109 and FLNC, was associated with worse survival, suggesting that methylation of this pathway might represent a prognostic signature in CRC. Similarly, hypomethylation of the insulin growth factor II (IGF2) differentially methylated region in colorectal tumours was associated with poor prognosis. Conversely, methylation of the polycomb group target genes, including SFRP1, MYOD1, HIC1, and SLIT2, resulted in favourable prognosis in non-CIMP male patients.

Lymphovascular invasion of CRC was related to methylation of the gene encoding the secreted protein acidic and rich in cysteine (SPARC) in stromal cells. Others analysed DNA methylation in mucosal wash fluid from patients undergoing colonoscopy, observing that methylation of the micro-RNA (miR-34b/c) had the greatest correlation with invasive tumours. Methylation of miR-128 in CRC samples led to an upregulation of its target gene NFK2 that resulted in lymphatic invasion and peritoneal dissemination. It was also shown that epigenetic silencing of miR-126 contributes to tumour invasion and angiogenesis in CRC.

Epigenetic signatures in CRC are also of interest for their possible interactions with chemotherapeutic agents. Indeed, the epigenetic silencing of a particular gene might result in chemosensitivity or chemoresistance toward a particular therapeutic agent. Crea et al. proposed a panel of genes whose aberrant methylation could contribute to chemosensitivity or chemoresistance to 5-FU, irinotecan, and oxaliplatin, three of the most frequently used drugs in CRC treatment. 5-FU antitumor activity is mainly exerted by inhibiting thymidylate synthase, in the de novo synthesis of pyrimidines. Increased TMY5 expression is one of the major mechanisms of 5-FU chemoresistance, and there is indication that histone acetylation/deacetylation processes, rather than DNA methylation of the promoter, might be of relevance in epigenetically regulating TMY5 expression in CRC. Several other genes that participate in pyrimidine metabolism might represent potential molecular determinants of 5-FU chemoresistance, including dihydropyrimidine dehydrogenase (DYPD), thymidine phosphorylase (TYMP), and uridine monophosphate/cytidine monophosphate kinase (UMPK) genes. Their potential epigenetic contribution to 5-FU resistance in CRC patients is under investigation.

Hypermethylation of the gene encoding the transcription factor AP-2 epsilon (TF-AP2E) was found in 51% of CRC patients and resulted in clinical nonresponsiveness to chemotherapy (5-FU, irinotecan or oxaliplatin). Functional assays showed that TF-AP2E chemoresistance is mediated through its downstream target gene DKK4.
encoding for dickkopf homolog 4 protein[115]. Other genes whose methylation might be associated with decreased sensitivity to 5-FU or irinotecan chemotherapy are SPARC coding for the matricellular protein osteonectin[116], and UGT1A1 coding for the UDP glucuronosyltransferase-1A1 enzyme, the major enzyme involved in irinotecan detoxification.[117]

CRC patients who had failed standard therapies (oxaliplatin, irinotecan, 5-FU, or cetuximab or panitumumab if KRAS wild type) were treated with dacarbazine, an alkylating agent that exerts its antitumor activity inducing base pair mismatches[117]. Hypermethylation of the MGMT gene occurs in almost 40% of CRC patients, and clinical responses to dacarbazine are confined to those tumours harbouring epigenetic silencing of the MGMT gene.[118]

FUTURE PERSPECTIVES

In addition to genetic aberrations, DNA methylation also plays important roles in the development of CRC. Recent genome-scale approaches revealed that CRCs exhibit multiple genetic alterations, including allelic imbalances (copy number alterations) at various chromosomal loci. For example, alongside with mutations of TP53, KRAS, BRAF, and PIK3CA, genomic losses commonly occurred at 3q26.1, 4q13.2, 6q21.32, 7q34, 8p12-23.3, 15qcen and 18, while gains were commonly found at 1q21.3-23.1, 7p22.3-q34, 13q12.11-14.11, and 20. Moreover, the total number of copy number alterations were significantly associated with the aberrant DNA methylation of six marker genes[118]. Similarly, transcriptome analyses are revealing thousands of genes whose expression is altered, likely through promoter methylation, in CRC tissues[119].

Goal of present and future research is to identify those biomarkers that could allow a feasible, cost-effective and non-invasive diagnosis of CRC, as well as to understand which panel of biomarkers can be used to better define patient’s prognosis and the best choice of available treatments (Figure 2). Several examples are provided within this review suggesting the need to combine genetic and epigenetic data for a better diagnostic, prognostic and therapeutic approach. Integration of those data with transcriptome and proteome profiles could represent a valuable strategy to further understand the molecular pathways involved in CRC, as well as to improve life expectancies and quality of life of the patients.

REFERENCES


28 **Losso** GM, Moraes Rda S, Gentilici AL, Messias-Reason IT. Microsatellite instability--MSI markers (BAT26, BAT25, D2S125, D5S346, D17S250) in rectal cancer. *Anq Bras Cir Dig* 2012; 25: 240-244 [PMID: 23141922]


Bokemeyer C, Bondarenko I, Hartmann JT, de Braud F, Schuch G, Zuelb A, Celik I, Schlichting M, Koralewski P. Efficacy according to biomarker status of cetuximab plus
Coppedé F et al. Biomarkers of colorectal cancer.


Neumann J, Weheveck I, Maatz S, Engel J, Kirchner T, Jung A. Alterations in the EGFR pathway coincide in colorectal cancer patients refractory to chemotherapy with wild-type KRAS. Tumour Biol 2013 Sep 1; Epub ahead of print [PMID: 23996432]


Di Nicolantonio F, Martini M, Molinari F, Sartore-Bianchi A, Arena S, Saletti P, De Dosso S, Mazzucchelli L, Frattini M, Siena S, Bardelli A. Wild-type BRAF is required for response to panitumumab or cetuximab in metastatic colorectal can-
[PMID: 23459561 DOI: 10.7150/jca.5839]


Kostin PA, Zakharzhkevskia NB, Gnerovoz EV, Govorun VM, Chernyshov SV, Shchelynig IuA. [Hypermethylation of the CDH1, SEPT9, HLF and ALX4 genes and their diagnostic significance in colorectal cancer]. Vopr Onkol 2010; 56: 162-168 [PMID: 20552891]


Coppedè F et al. Biomarkers of colorectal cancer


