Microbial markers in colorectal cancer detection and/or prognosis

Romain Villeger, Amélie Lopès, Julie Veziant, Johan Gagniere, Nicolas Barnich, Elisabeth Billard, Delphine Boucher, Mathilde Bonnet

To cite this version:

HAL Id: hal-01897664
https://hal.archives-ouvertes.fr/hal-01897664
Submitted on 15 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License
Microbial markers in colorectal cancer detection and/or prognosis

Romain Villéger, Amélie Lopès, Julie Veziant, Johan Gagnière, Nicolas Barnich, Elisabeth Billard, Delphine Boucher, Mathilde Bonnet

Abstract
Colorectal cancer (CRC) is the second leading cause of cancer worldwide. CRC is still associated with a poor prognosis among patients with advanced disease. On the contrary, due to its slow progression from detectable precancerous lesions, the prognosis for patients with early stages of CRC is encouraging. While most robust methods are invasive and costly, actual patient-friendly screening methods for CRC suffer of lack of sensitivity and specificity. Therefore, the development of sensitive, non-invasive and cost-effective methods for CRC detection and prognosis are necessary for increasing the chances of a cure. Beyond its beneficial functions for the host, increasing evidence suggests that the intestinal microbiota is a key factor associated with carcinogenesis. Many clinical studies have reported a disruption in the gut microbiota balance and an alteration in the faecal metabolome of CRC patients, suggesting the potential use of a microbial-based test as a non-invasive diagnostic and/or prognostic tool for CRC screening. This review aims to discuss the microbial signatures associated with CRC known to date, including dysbiosis and faecal metabolome alterations, and...
Villéger R et al. Microbial markers for colon cancer screening

the potential use of microbial variation markers for non-invasive early diagnosis and/or prognostic assessment of CRC and advanced adenomas. We will finally discuss the possible use of these markers as predictors for treatment response and their limitations.

Key words: Colorectal cancer; Microbiota; F. nucleatum; Colibactin-producing E. coli; Prognostic markers; Diagnostic markers; Dysbiosis

© The Author(s) 2018. Published by Baishideng Publishing Group Inc. All rights reserved.

Core tip: Many clinical studies have reported a disruption in the gut microbiota balance and in the faecal metabolome in colorectal cancer. In this review, we describe the modifications in the microbiota composition and metabolome observed in colorectal cancer (CRC) tissue and stool samples. Then, we detail how these microbiota modifications may represent novel and promising non-invasive diagnostic and/or prognostic markers for CRC and advanced adenoma.


INTRODUCTION

Approximately 5% of the population in Western countries will develop colorectal cancer (CRC) during their lifetime[1]. CRC is a heterogeneous disease with a wide range of long-term outcomes and responses to treatment. Despite recent advances in the genetic and molecular characterization of tumours, CRC remains associated with a poor prognosis and very low rates of long-term survival among patients with advanced disease[2,3]. On the contrary, the 5-year survival of patients treated during the early stages of CRC, which represent only one-third of all CRC patients, is encouraging, ranging from 72% to 100%[4]. Therefore, early detection of both premalignant lesions and CRC is crucial for increasing the chances of a cure[5]. Moreover, accurate determination of the prognosis is crucial for the practitioners, in order to optimize and personalize treatment strategies. Given the poor prognosis of this cancer, it is essential to validate new diagnostic and prognostic markers. Studies on the implications of the gut microbiota in CRC are now increasing[6-11]. Identification of a CRC-associated dysbiosis and/or the targeted detection of procarcinogenic bacterial species and/or their effectors in stool or tissue samples may represent promising tools to overcome the limitations and the poor performance of current CRC screening tools and prognostic/predictive factors. This review will focus on gut microbiota-related markers, which may be new biological markers for CRC screening or prognosis determination.

CURRENT STRATEGIES FOR CRC SCREENING AND PROGNOSTIC FACTORS

Screening strategies

Individual (opportunistic) and population-based (organized) screening strategies have been established permitting the reduction of CRC mortality either by detecting CRC or removing premalignant lesions[12]. Several CRC screening tests have been used and have different degrees of performance that vary worldwide and are primarily driven by costs, lifestyle and endoscopic resource constraints. In addition to the classic invasive endoscopic approaches, there has been great interest in developing and evaluating non-invasive methods, including radiographic examinations and stool- and blood-based tests, with the goal of creating efficient, safe, convenient and cost-effective screening tools[13-18].

Endoscopy is currently the most effective modality for detecting asymptomatic premalignant polyps and CRC and for preventing CRC development by the immediate removal of premalignant lesions[13]. However, this procedure requires prior bowel cleansing and causes postprocedural discomfort, which are both responsible for low participation rates in colonoscopy CRC screening programmes[13]. Moreover, the need for general anaesthesia and the invasiveness of the procedure exposes patients to potentially severe complications[19,20]. Computed tomography (CT) colonography, also known as “virtual colonoscopy”, is a non-invasive alternative to colonoscopy[21-23]. However, CT colonography presents some drawbacks such as its high cost, lack of standardized methods, need for bowel cleansing, poor performance for detecting small or flat lesions, and lack of ability to perform immediate biopsy[14,24]. For these reasons, other non-invasive stool- and blood-based tests are more promising tools for initial CRC screening in two-step approaches.

Non-invasive stool-based early CRC screening modalities involve either detecting occult bleeding in the stool, such as with guaiac fecal occult blood test (gFOBT) and faecal immunohistochemical testing (FIT), or detecting DNA, RNA and protein markers of neoplasia in faeces. The rationale for the use of such tools is that the mucosa of both polyps and CRC is more fragile and vascularized and thus bleeds and desquamates more easily than normal colonic mucosa. The effectiveness of gFOBT and FIT have been established for reducing both CRC incidence and mortality, thus permitting their use in many CRC screening programmes[2,19,21,25-36]. However, these techniques have been criticized for their relatively low specificity[14]. Therefore, new molecular...
Prognostic and predictive factors

Many prognostic factors including many socioeconomic and clinical factors are routinely used to optimize and personalize treatment strategies. However, pathological factors are currently the most powerful tools that are routinely used to adapt therapeutic strategies, especially for conditioning the administration of non-operative treatments, such as chemotherapy and radiotherapy. The TNM international staging system remains the gold-standard for determining both CRC prognosis and treatment indications. Unfortunately, this classification is largely imperfect, which is highlighted by the constant changes it undergoes with the advances in the field of CRC research. Furthermore, it has some limitations. In particular, it is inaccurate for some subgroups of patients whose prognosis and indications for adjuvant treatments remain unclear. Therefore, some other pathological prognostic factors are routinely used, such as the presence of lymphovenous and/or perineural invasion, tumour budding and tumour differentiation grade.

In addition, studies have shown tumour genomics and cellular alterations can supplement clinicopathological factors. Aneuploidy, overexpression of the CEA in tumours, allelic loss in DCC, TP53, APC and MCC genes, TP53 gene mutations, expression of CD44 protein, high levels of thymidylate synthetase, and the detection of MSI and both RAS and BRAF genes mutations all play prominent roles in determining CRC prognosis and in personalizing treatments. They are independent strong prognostic factors and are also predictive of the response to some widely used chemotherapy regimens, including targeted therapies and immunotherapy.

Serum prognostic factors [e.g., C-reactive protein (CRP), carcinoembryonic antigen (CEA), and cfDNA] can also be very useful tools. Both CRP and serum albumin levels are associated with a worse prognosis but are not routinely used as prognostic factors for CRC. The CEA level is a well-recognized and routinely used independent prognostic factor for CRC. Persistently elevated CEA levels after treatment are correlated to both a higher risk of recurrence and decreased survival. Similarly, normal CEA levels after treatment are associated with improved recurrence-free and overall survival. In addition to CRC screening, investigations evaluating cfDNA have recently gained attention in the field of CRC research. The detection and the quantification of specific gene mutations (RAS and BRAF genes) or DNA methylation alterations are correlated with CRC prognosis, independent of the tumour stage. In the future, monitoring of cfDNA levels may guide therapeutic strategies and improve patient follow-up, allowing the earlier detection of recurrence.

As we previously mentioned, all these screening strategies and prognostic markers present many limitations. Therefore, there is still a need to develop other effective non-invasive screening and prognostic tools to replace or supplement those currently in use, with the aim of decreasing CRC mortality and optimizing treatment strategies. The study of intestinal microbiota provides new leads to identify powerful biomarkers.

MICROBIAL SIGNATURES IN CRC

Functions of intestinal microbiota on gut homeostasis

The intestinal microbiota and its host have coevolved to establish a symbiotic relationship leading to digestive system homeostasis. Many metabolic, immunological, structural and protection functions that are essential to the health of the host are performed by the intestinal microbiota. Thus, one of the major functions of the gut microbiota is to ferment complex carbohydrates in order to produce a large amount of metabolites (short chain fatty acids, bile acids, choline, essential vitamins, etc.), some of which represent energetic and nutrient substrates for intestinal cells. Others are involved in the regulation of cellular processes or participate in regulating hepatic lipid and glucose homeostasis. Another function of the gut microbiota is the maintenance of epithelial homeostasis since commensal bacteria can promote epithelial integrity. They can induce structural and functional maturation of the colonic epithelium with strong modulation of mucus cells. The intestinal microbiota also plays a fundamental role in the development, maturation and function of the host immune system and also confers resistance to colonization by pathogenic microorganisms by competing both for nutrients and attachment sites to the intestinal epithelium, by producing and secreting antimicrobial compounds, such as bacteriocins, and by strengthening epithelial tight junctions.

CRC-associated gut microbiota dysbiosis

Approaches used to study the gut microbiome: The gut microbiota has a very large diversity of microbial populations such as bacteria, archaea, eukaryotes and viruses. Because the colonic microbiota is mostly represented by bacteria, the bacterial community is the most studied. Both tissue and faecal samples
provide information about the structure of bacterial populations. The analysis of tissues is more relevant for evaluating the mechanisms of microbiota involvement in physiopathology. However, sampling of mucosa-associated bacteria by endoscopic biopsy of intestinal tissue is an invasive procedure. It is more difficult to implement than collection of a faecal sample, especially in healthy donors. Therefore, the vast majority of studies on the diversity and richness of the gut microbiota are conducted on stool samples, which are more useful for identifying diagnostic or prognostic markers for pathologies such as CRC.

The first data concerning the gut microbiota were provided by cultural approaches. However, these approaches allow only limited evaluation of this ecosystem, since less than 30% of intestinal bacteria have been cultivated to date. Since the 1990s, the advent of molecular tools targeting the bacterial 16S ribosomal RNA (rRNA) gene has revolutionized our knowledge of the gut microbiota from both faeces and tissues. Thus, genetic fingerprinting techniques (terminal restriction fragment length polymorphism (T-RFLP), denaturing gradient gel electrophoresis (DGGE), hybridization approaches (fluorescence in situ hybridization (FISH), microarrays) and clone library analysis have been applied to provide a more complete description of its structure\[118-122\]. Currently, quantitative PCR (Q-PCR) and 16S rRNA gene next-generation sequencing (NGS) are the most used methods for describing the composition of the intestinal bacterial community and therefore for comparing the gut microbiota of healthy individual to that of patients with diseases. Thus, modifications in the microbiota composition, also called dysbiosis, in faeces and tissues have been described in CRC\[123-130\]. These different approaches allow only taxonomic identification of microbial communities. However, the gut is a complex and variable ecosystem. Understanding its function and the specific role of the different bacterial populations is essential. A metagenomics approach (shotgun sequencing) makes taxa identification possible and allows exploration of the metabolic potential of the intestinal microbiota\[126,131,132\]. Thus, the American Human Microbiome Project (HMP) and the European "MetaHIT" (Metagenomics of the Human Intestinal Tract) consortia were formed in order to generate a catalogue of genes carried by the gut microbiota\[133\]. Intestinal microbial dysbiosis was recently investigated in patients with CRC\[125,126,131,132\]. Combining the metagenomics data available in databases with the application of other high-throughput meta-omics approaches, it is now possible to gain a better insight into the microbial function and to reveal the link between genetic potential and functionality in the gut microbiota.

Cancer-associated microbiome alterations: To date, there is no consensus in terms of the microbiota modifications observed in CRC using high-throughput sequencing of the bacterial 16S rRNA gene and metagenomics both in faecal and tissue samples compared with samples from healthy individuals. The complexity of the gut microbial ecosystem associated with other technical or biological parameters, such as the geographical location of the studied populations, the lifestyle, the diet, the sample type (faeces, mucosa, tumour), the location of the tissue sampled (i.e., right or left colon), the age of the individual, the stage of the disease, the number of patients studied, the molecular approach used (metagenomics, 16S rRNA gene NGS), the regions of the 16S rRNA targeted (V1-V2, V3, V3-V5, V4-V6 regions) for the NGS approach, and the taxonomic level considered, contribute to the lack of consensus on the definition of the composition of dysbiotic microbiota in patients with CRC.

Alpha diversity metrics, which represent the specific richness and the distribution of individuals within the species were found to be increased in adenoma and adenocarcinoma tissues compared with in normal mucosa\[134-136\]. However, these parameters were unchanged in faecal samples\[125,134,137,138\], suggesting that the increase in diversity was limited to the neoplastic lesions and did not extend to the entire mucosa.

Significant shifts can be observed in the composition of the bacterial community at the phylum taxonomic level but also on a finer phylogenetic scale in CRC samples. However, these modifications vary depending on the analytic techniques used and the sample location (mucosa, adenocarcinoma, stool). It is accepted that Bacteroidetes, Fusobacteria and Proteobacteria are enriched, while Firmicutes is decreased in CRC patients\[10,119\] (Figure 1). A similar dysbiosis has been described in CRC animal models\[126,134,140\]. In a taxonomy-based analyses conducted in 2013, Ahn et al\[143\] described a lower relative abundance of Clostridia and increased Fusobacterium and Porphyromonas in case subjects compared with control subjects. Additionally, the genera Bacteroides and Prevotella, Bacteroides fragilis, Enterococcus faecalis, Streptococcus bovis, Streptococcus gallolyticus and Escherichia coli were overrepresented, while Bacteroides vulgatus, Lactobacillus and Faecalibacterium prausnitzii were decreased in the faeces or tumours of patients with CRC\[123-125,131,143-146\]. Other studies found that Parvimonas micra, Solobacterium moorei, Peptostreptococcus stomatis were consistently enriched in the microbiomes of the stools of CRC patients\[126,131\]. Many studies showed an enrichment of Fusobacterium nucleatum in stool and tissue samples from patients with colorectal carcinoma, which were confirmed by quantitative polymerase chain reaction (PCR)\[126,131,147-149\] (Figure 1). In addition to taxonomic studies, a metagenomics analysis showed an enrichment of virulence-associated bacterial genes in the tumour microenvironment\[122\]. Many studies reported a high prevalence of toxin genes expressed by some species described above. Most of enterotoxigenic B. fragilis (ETBF) strains detected in mucosal samples from patients with CRC harbour the bft gene, which encodes the bacterial toxin B. fragilis toxin (BFT)\[149\] (Figure 1). ETBF and BFT carrying strains were detected more...
often in stool samples from CRC patients compared with samples from controls [150]. Many studies reported that mucosa-associated BFT-producing *B. fragilis* was more prevalent in late-stage CRC [130,149]. Other toxins named cyclomodulins (cytolethal distending toxin (CDT), colibactin, cytotoxic necrotizing factor and cycle inhibiting factor), which are carried by *E. coli*, were more prevalent in CRC samples [145,146,151,152]. In particular, the colibactin toxin, which is synthesised from the pks island, was preferentially detected in strains isolated from CRC patients [145,146,152] (Figure 1). These toxin genes could be targets of interest for developing new prognostic or diagnostic factors.

The majority of published data were obtained in the adenocarcinoma stage. However, the intestinal mucosa might present different ecological niches according to the stage of colorectal tumour progression. Due to difficulties in sampling patients during the early stages of the disease, studies focusing on the temporal evolution of microbiota dysbiosis during colorectal carcinogenesis are scarce. Nevertheless, metacommunities can be defined from the bacterial community configurations by Nakatsu et al [153]. In early stages of CRC, bacteria belonging to *Fusobacterium*, *Parvimonas*, *Gemella* and *Leptotrichia* genera were more abundant [153]. Moreover, the occurrence of *Escherichia coli*, *Pseudomonas* veronii and *Lactococcus* were specific to adenomatous lesions [135,153], whereas *Fusobacterium* enrichment was observed in carcinomas [153]. Additionally, the abundance of *Bacteroides fragilis* and *Granulicatella* were progressively increased during the adenoma–carcinoma sequence. Lu et al [135] showed that there was a clear increase in both *Proteobacteria* and *Bacteroidetes* in tissues from patients with colorectal adenoma compared with tissues from healthy volunteers. Kostic et al [147] found that *Fusobacterium* spp. were enriched in human colonic adenomas relative to the surrounding tissues and in stool samples from patients with colorectal adenoma compared with samples from healthy subjects. In addition, *Lactococcus* and *Pseudomonas* were enriched in preneoplastic tissue, whereas *Enterococcus*, *Bacillus*, and *Solibacillus* were reduced. A high frequency of colibactin-producing *E. coli* and enterotoxigenic *Bacteroides fragilis* were also observed in adenomas from patients with
Familial Adenomatous Polyposis (FAP). Recently, Eklof et al. showed that the prevalence of colibactin-producing *E. coli* were progressively increased in the adenoma-carcinoma sequence.

Finally, most of these studies do not take into account the heterogeneity of this pathology. As an example, *Prevotella* and *Firmicutes* populations were more abundant in individuals with proximal tumours, whereas *Bacteroidetes* populations were overrepresented in distal cancers. More recently, a drastic difference in the microbial composition has been observed in the mucosa of colitis-associated CRC patients, with an increase in the *Enterobacteriaceae* family and the *Sphingomonas* genus and a decrease in *Fusobacterium* and *Ruminococcus* genera, when compared with the mucosa of sporadic CRC patients. Similarly, we observed that colibactin-producing *E. coli* were more frequently identified in microsatellite stable (MSS) CRC. In the same way, high colonization by *F. nucleatum* and negative-colibactin *E. coli* bacteria were detected in patients with MSI (microsatellite instability) compared with the MSS phenotype. All these data suggest that the dysbiosis signature of CRC might be different according to the tumour phenotype and/or molecular alterations.

In conclusion, all these data show that there is an adeno- and a CRC-associated signature in the microbiome (Figure 1). Metagenomics studies showed a redundancy in the metabolic functions of the intestinal microbiota. Therefore, studies of metabolites might reflect the CRC-associated microbiota dysbiosis with a lower heterogeneity than taxonomic determination.

### Profiling metabolome alterations in CRC

**Metabolite analysis procedures:** The metabolome has been defined as the complete complement of all small molecule (< 1500 Da) metabolites found in a specific cell, organ or organism. Global metabolomics alterations reflect changes due to environmental factors, diet, drugs, exercise, genetic variation and regulation, changes in the gut microflora, and altered kinetic activity or levels of enzymes. The most widely used analytical high-throughput tools for metabolomics studies are nuclear magnetic resonance (NMR) and mass spectrometry (MS), and they both can provide complementary snapshots of the metabolome of body fluids. Sample preparation is arguably the most critical step in metabolomics, but there is no universal, robust and repeatable approach defined. NMR spectroscopy benefits from being highly reproducible and offers the potential to precisely quantify compounds in complex mixtures without complex sample preparation requirements. Gas chromatography coupled with MS (GC-MS), which is used for the analysis of volatile low molecular weight metabolites, and liquid chromatography coupled with MS (LC-MS), which is used for the analysis of non-volatile components, have a higher sensitivity than NMR, but the sample preparation must be more fastidious. For more details regarding sample preparation in metabolomics, several authors previously described methods for NMR, GC-MS and LC-MS analysis.

During the last decade, using NMR and/or MS, metabolic profiles of colorectal tumour biopsies and matched colon mucosa samples revealed potential biomarkers in carbohydrate metabolism, short-chain fatty acid and other lipid metabolism, nucleotide biosynthesis, and secondary bile and steroid metabolism. The repertoire of metabolic derangements derived from cancer cells, their microenvironment and the gut microbiota composition and activity may culminate in a distinct metabolic phenotype (also called metabolotype) that characterizes the pathology of CRC. Metabolome analysis of blood and urine samples from CRC patients were performed, and several studies reported relevant biomarkers for CRC. However, the transient localization of faeces in the colon and rectum makes it an ideal matrix for gaining information about the health and pathological state of the colon and rectum, and thus for colorectal cancer diagnosis. Faecal samples consist of a diverse range of endogenous metabolites including tumour cell metabolites and gut microbiota metabolites.

**Microbial influence on the faecal metabolome:** When ingested, the majority of food nutrients are absorbed in the small intestine. The residues that reach the colon are complex nutrients, generally fibres (complex carbohydrates), but also protein residues and host products such as mucin and bile acids secreted by the liver in response to fat ingestion. These compounds are metabolized by the microbial populations into a diversity of metabolites that can be used by the host colonic cells. The gut microbiota makes an important contribution to human metabolism by providing enzymatic functions that are not encoded by the human genome, such as the breakdown of polysaccharides and polyphenols and the synthesis of vitamins by fermentation and anaerobic degradation.

Polysaccharides and non-digestible carbohydrates are the primary substrates for microbial fermentation. Saccharolytic fermentation of carbohydrates produces beneficial metabolites such as short-chain fatty acids. However, if there is limited carbohydrate intake, bacteria use alternative energy sources, resulting in the production of other metabolites that may be more detrimental to human health. The three SCFAs acetate, propionate and butyrate, which are considered beneficial metabolites, are mainly produced from polysaccharides and oligosaccharides through their metabolism into pyruvate. Propionate and acetate can be produced from lactate, while butyrate and acetate can be produced from acetyl CoA. Butyrate is an important energy source for gut epithelial cells, while propionate and acetate are mostly metabolized in the liver for cholesterol synthesis, gluconeogenesis or lipogenesis. For more information, the complete synthesis and functions of SCFAs was extensively and previously described. Microbial use of bile salts is a classic example of the relationship between the gut...
microbiota metabolism and the host metabolism. The primary bile acids cholic acid and chenodeoxycholic acid are produced in the liver from cholesterol and are conjugated to glycine or taurine. They are secreted by the liver following high fat intakes[176]. The intestinal microbiota can affect the biotransformation of bile acids via deconjugation, dehydrogenation, epimerization, and 7α/β-dehydroxylation of primary bile acids to generate secondary bile acids[171]. A number of gut bacteria possess bile salt hydrolase enzymes that are capable of hydrolysing the amide bond between the bile acid and its conjugated amino acid, producing the secondary bile acids deoxycholate and lithocholate (Figure 2). Bile salt hydrolase genes were identified in Bacteroides, Clostridium, Lactobacillus, and Listeria[178]. Deconjugation of bile acids by microbial hydrolases leads to an increased level of secondary bile acids, especially deoxycholic acid, which is also known to exert carcinogenic activity[179]. Polyphenols from fruits and vegetables are poorly absorbed in the intestine, and 90-95% are metabolized by the gut microbiota[180]. The diversity of polyphenol compounds is wide and 90­95% are metabolized by the gut microbiota[177]. Several gut microbial enzymes can transform plant polyphenols, such as phenolic acids, flavonoids, and lignans into simple phenolic metabolites[172,181,182]. These microbial phenolic metabolites are mostly known for their antioxidant activity and have been shown to inhibit the production of pro-inflammatory mediators[183]. Microbial protein fermentation generates potentially toxic and pro-carcinogenic metabolites involved in CRC, such as phenols, sulfides, ammonia and nitrosamines[171]. A subset of bacteria, including Bacteroidetes and Firmicutes species, produce potentially bioactive substances via the degradation of amino acids, especially nitrogenous compounds that can exert carcinogenic effects[184­186]. Moreover, sulfides produced in the gut by the bacterial reduction of dietary sulfate are enterotoxic[187]. Microbial metabolism may contribute to the toxicity of alcohol, especially in the gastrointestinal tract where aerobic and facultative anaerobic bacteria convert ethanol to acetaldehyde, which is known to be a highly toxic and pro-carcinogenic compound[188]. A specific role for bacterial biofilms in colonic metabolome was also reported[189]. With organic acids, gas is the major products of microbial fermentation in anaerobic systems. Anaerobic bacterial fermentation can lead to the synthesis of hydrogen, carbon dioxide, or methane for the majority. Some minor gases, such as NH₃, N₂, volatile amines, indolic and phenolic compounds, that are mainly associated with peptide fermentation, are also produced[190,191] (Figure 2). Among them, H₂S is produced by sulfate-reducing bacteria (Desulfovibrio spp.) by sulfur amino acids and taurine metabolism.

Given the key role of the gut microbiota in dietary metabolism and the growing evidence of its potential impact on host health, studies were performed to elucidate the metabolic functions of the gut micro-

Figure 2  Metabolism of food dietary components by the gut microbiota. Dietary residues are in blue, intermediate metabolites in green, SCFAs in red, and other metabolites in grey. Adapted from Louis et al[23], Russel et al[34], Rowland et al[172], Flint et al[171]. SCFAs: Short chain fatty acids.
biota. These studies mostly used metagenomic, metatranscriptomic, or metaproteomic approaches, but using metabolomics to study the gut microbiota appears to be a promising method. Indeed, metabolomics allow the measurement of both host and microbial metabolites, including metabolites in host samples derived from the microbiota and host metabolites processed by the microbial population.

**Metabolome alterations in CRC:** The analysis of the CRC-associated metabolome has highlighted differences in the biochemical composition of CRC patients’ stools. Few animal studies have investigated the metabolomic profile of faeces in CRC models. Significant alterations were observed in fatty acid metabolites and metabolites associated with bile acids (hypoxanthine xanthine, taurine) in faeces of Hmga1 mice as analysed by ultra-performance liquid chromatography UPLC/MS/MS and GC-MS[192]. Increased levels of short, arginine-enriched, tetra-peptide fragments were reported in the transgenic mouse faeces[192]. Similarly, CRC patients were found to have higher levels of some amino acids and alterations in the levels of some SCFA in their stools compared with controls[193]. Butyrate and several butyrate-producing bacteria were depleted in CRC patients’ stools[194]. Chen et al.[195] observed that butyrate and butyrate-producing bacteria were more prevalent in healthy controls than in advanced colorectal adenoma patients, with a lower prevalence of Clostridium, Roseburia, and Eubacterium spp. and a higher prevalence of Enterococcus and Streptococcus spp. Moreover, faecal β-glucuronidase activity is increased in patients with CRC compared with healthy controls[195]. High-fat diets, which are positively correlated with the incidence of CRC, lead to an increase in bile secretion, and increased faecal bile acid concentrations were reported in patients with CRC[196]. High faecal bile concentrations were positively correlated with the incidence of CRC, with potential tumour-promoting effects of bile acids themselves[197]. Furthermore, strong antimicrobial bile acid activities lead to significant changes in the gut microbiota composition, with a relative increase in some Gammaproteobacteria and Bacteroidetes species that are associated with CRC[197]. The dietary intake of preformed N-nitroso compounds is positively correlated with CRC in European populations[185], but these compounds can also be formed endogenously via acid-driven nitrosation in the stomach and by the nitrosation of amines that are derived from the microbial fermentation of protein in the large intestine[198]. Higher levels of H:O:O: were detected in CRC patients than those in healthy controls, but no increases in Desulfovibrio spp. were observed in faecal samples from patients with CRC[199]. Using metabolome screening, a metabolic influence of microbial biofilms on colonic tissues and the related occurrence of cancer has been reported. It was found that host cancer and specific bacterial biofilm structures contributed to the polyamine metabolite pool (N³,N³-diacylserpentine), which might have an impact on cancer development and progression[199].

All these data suggest that faecal metabolomics may not only be used for diagnostic purposes but may also serve as a prognostic tool for CRC treatment in the future.

### MICROBIAL-RELATED BIOMARKERS FOR CRC DIAGNOSIS AND CLASSIFICATION

#### Validation of multi-bacteria models for early detection and/or prognosis determination

Even if a precise microbial signature associated with CRC and a precancerous lesion have not been exactly defined, different works have shown that dysbiosis may provide new significant biomarkers for CRC diagnosis or prognostic determination.

Several studies were able to determine a correlation between faecal microbial dysbiosis and CRC prognosis/diagnosis. In 2014, Zackular et al.[140] used faecal samples from 30 CRC patients, 30 colonic adenoma patients and 30 healthy controls to establish a classification model. The authors observed significant differences in the gut microbiome of healthy controls when compared with those with adenoma or CRC, and faecal microbial screening had an accuracy of 0.798 AUC for predicting CRC. By combining their gut microbiome data with known clinical risk factors (body mass index, age, race), they were able to significantly improve the ability of faecal microbial screening to discriminate between the three clinical groups when compared with risk factors alone. Zeller et al.[131], with 53 CRC patients, 42 adenoma patients and 61 controls, established a 16S-based classifier for CRC detection and validated their model with an accuracy of 0.82. More recently, Ai et al.[200] sequenced the 16S rRNA gene from the faeces of 141 Chinese participants and evaluated the performance of different classifiers for predicting CRC based on faecal microbiota. The authors reported a strong variation in the CRC prediction between the different models tested, with the Bayes Net algorithm displaying the best performance. Their prediction test, with 0.93 AUC, was more accurate than the FOBT, while combining both tests improved the accuracy. This list of studies that assessed the use of changes in the faecal microbiota for colorectal adenoma and CRC screening is non-exhaustive. Among these studies, while an association between CRC and microbiota is clear, there is limited agreement in the taxa reported.

For this reason, Shah et al.[201] conducted a microbiome-based meta-analysis for CRC, using the results from 9 studies and including 79 colorectal adenomas patients, 195 CRC patients and 235 controls, in order to identify a common microbial marker in stool samples. In addition to the previously reported taxa, they highlighted a significant increase in Parvimonas micra ATCC 33270, Streptococcus anginosus, Parabacteroides distasonis and other members of Proteobacteria. Their microbiome-based CRC versus control classification produced an area under the receiver operating characteristic (AUROC)
curve of 80.3%, which was 91.3% when combined with clinical markers from the bioinformatics pipeline. Similarly, Amitay et al [202] conducted a systematic review of 19 studies (including most of the studies cited above) examining the differences in the gut microbiome in faecal samples from people diagnosed with CRC or adenomas and from people without colorectal neoplasia to generate faecal microbiota profiles for early detection of adenomas and CRC. Overall, they concluded that there was limited but encouraging evidence that the differences in the faecal gut microbiome between people diagnosed with CRC or adenoma and healthy controls could be used to identify new faecal tests. The authors suggested that future research should focus on developing unified documented and reproducible protocols for studying the human gut microbiome from faecal samples for more comparable results between cohorts [203].

In many studies, patients had a corresponding decrease in the Firmicutes/Bacteroidetes ratio, which might be an important marker for intestinal dysbiosis in colorectal cancer. F. nucleatum, BFT-producing B. fragilis, and colibactin-producing E. coli are known to be associated with a more aggressive TNM stage [130,145,149,158]. Their detection by PCR in tissues or faeces might allow the development of rapid tests for new prognostic factors. With a 16S rRNA gene sequencing approach, F. nucleatum, B. fragilis, and Faecalibacterium prausnitzii were identified as useful prognostic biomarkers for CRC [203]. Indeed, in this study, F. nucleatum and B. fragilis were more abundant in the groups with a worse prognosis, while Faecalibacterium prausnitzii was more abundant in the survival group. Similarly, Mirza et al [159] showed that the amount of F. nucleatum DNA in colorectal cancer tissue was associated with shorter survival. Another study demonstrated that the detection of F. nucleatum associated with the non-invasive screening FIT could be a promising marker for detecting neoplastic lesions [204]. Recently, Eklöf et al [151] explored the use of microbial markers for bacteria harbouring the pks island, which codes for colibactin synthesis, and F. nucleatum in stool as potential screening markers for CRC. Authors suggested that the presence of the pks island and F. nucleatum detection could predict cancer with a specificity of 63.1% and a sensitivity of 84.6%, suggesting the potential value of these microbial parameters as putative non-invasive biomarkers for CRC detection [151]. Multicentric clinical studies need to be performed to validate all these promising results in a larger cohort. Moreover, these bacterial markers might provide a CRC-associated microbiome risk profile that might aid in the early identification of individuals who are at risk and require close surveillance.

**Screening faecal microbial-related metabolites for CRC detection**

Screening the faecal metabolome is a promising non-invasive procedure for obtaining a unique metabolic fingerprint to diagnose or determine the prognosis of CRC. To our knowledge, only a few studies with different metabolomics methods have shown the diagnostic potential of faecal samples for human colorectal cancer. They are summarized in Table 1. In 2009, Bezabeh et al [205] used 1H NMR to detect colorectal neoplasia in 111 CRC human faecal samples and compared them with samples from 412 healthy controls. NMR-based metabolomic profiling of faecal water extracts from patients with colorectal cancer and healthy individuals was able to identify potential diagnostic markers, such as SCFA (acetate and butyrate) and amino acids (proline and cysteine) [206]. More recently, the analysis of lyophilized stools by HPLC-GC/MS-MS resulted in the identification of 41 relevant faecal metabolites, such as xenobiotics, heme, peptides/amino acids, vitamins and co-factors, that had increased or decreased concentrations in CRC samples [207]. Volatile metabolome profiling of faecal samples using GC/TOF-MS also had diagnostic potential for detecting colorectal cancer. Phua et al [208] evaluated a small cohort and identified three specific markers (fructose, linoleic acid, and nicotinic acid) that were found at lower levels in the faecal volatile metabolome of CRC patients than in that of healthy subjects. A similar study using an electronic nose on samples from 40 CRC patients, 60 advanced adenoma patients, and 57 healthy controls was able to strongly distinguish the faecal volatile organic compounds (VOC) profile of patients with CRC and advanced adenoma from controls [209]. Given the volatility of the analysed compounds, improper sample collection and storage can drastically reduce the repeatability and robustness of the method. Batty et al [210] proposed the use of selected ion flow tube mass spectrometry (SIFT-MS) to classify 62 human faecal samples with a positive result on FOBT. Indeed, their method showed a 75% correct discrimination between CRC samples and low-risk samples [220]. More recently, Sinha et al [211] intended to establish a microbe-metabolome correlation through the analysis of lyophilized faecal samples of CRC patients and 89 healthy controls using HPLC-MS-MS. They reported that CRC was independently associated with lower levels of Clostridia, Lachnospiraceae, p-aminobenzoate and conjugated linoleate and with higher levels of Fusobacterium, Porphyromonas, p-hydroxy-benzaldehyde, and palmitoyl-sphingomelin. The authors identified a strong microbe-metabolite correlation in CRC patients. In conclusion, the diagnostic potential of metabolic profiling of faeces is strongly supported by several human studies, but these analytical techniques require standard procedures in order to obtain comparable and robust high-quality data.

**Prediction of treatment response with microbial-associated markers**

It has long been recognized that the gut microbiota can modify the pharmacokinetics of various drugs including anticancer therapies thus influencing therapeutic outcomes [212-217] and/or side effects [218-220]. Irinotecan, which is a first-line chemotherapeutic agent for metastatic CRC, causes adverse and dose-limiting effects that are
largely influenced by bacterial β-glucuronidase. Metagenomic and metabolomic profiling of patients’ gut microbiota could thus be informative before choosing this drug to predict side effects. Moreover, specific members of the gut microbiota might also drive chemoresistance to treatment. Indeed, Fusobacterium nucleatum, causes resistance to oxaliplatin and 5-FU through autophagy induction in colorectal cancer cell lines in vitro and in vivo in mouse models of colorectal xenografts. Conversely, the sensitivity of cancer cells to many chemotherapeutic agents has been shown to be modulated by intra-tumoural microbial agents has been shown to be modulated by intra-tumoural microbial agents. More specifically, the gut microbiota has a long-term impact on both the efficacy and toxicity of anticancer therapies. In particular, it can modulate the concomitant anti-tumour immune response. Indeed, various pre-clinical mouse models of subcutaneous tumours (melanoma, lung cancer, sarcoma) have suggested that specific crosstalk between bacteria and immune cells may affect responses to anticancer chemotherapies. In terms of CRC in particular, the efficacy of oxaliplatin, which is a drug routinely used for CRC treatment, has been shown to be strongly reinforced by gut microbiota in mice that were transplanted with MC38 colorectal syngeneic tumours, due to myeloid cells with increased anti-tumour functions. In addition, it has been described that the gut microbiota contributes to anti-tumour functions of adoptively transferred T cells in several models, including colorectal tumours implanted in mice. This suggests that the patients’ microbiota should also be taken into account in future clinical studies involving infusion of autologous anticancer T cells.

Table 1  Fecal metabolic profiling studies in colorectal cancer

<table>
<thead>
<tr>
<th>Matrix</th>
<th>Cohort</th>
<th>Study observations</th>
<th>Analytical method</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aqueous dispersion of stools</td>
<td>111 CRC, 412 healthy controls</td>
<td>Potential to detect colorectal neoplasia</td>
<td>One-dimensional 1H magnetic resonance spectroscopy</td>
<td>[205]</td>
</tr>
<tr>
<td>Fecal water extract</td>
<td>21 CRC, 11 healthy controls</td>
<td>Reproducible and effective method for detecting colorectal cancer markers.</td>
<td>HPLC-GC/MS-MS</td>
<td>[207]</td>
</tr>
<tr>
<td>Lyophilized human faeces</td>
<td>10 healthy controls</td>
<td>((\gamma)) SCFA (acetate, butyrate) appears to be the most effective marker in CRC.</td>
<td>SIFT-MS</td>
<td>[210]</td>
</tr>
<tr>
<td>Volatile organic compounds in the headspace of lyophilized stool samples</td>
<td>40 CRC, 60 advanced adenomas, 57 healthy controls</td>
<td>Discrimination of fecal VOC profiles of patients with adenomas and CRC.</td>
<td>Electronic nose</td>
<td>[209]</td>
</tr>
<tr>
<td>Lyophilized human faeces</td>
<td>48 CRC, 102 healthy controls</td>
<td>41 metabolites significantly associated with CRC ((\gamma)) xenobiotics ((\gamma)) heme, peptides/amino acids, vitamins, co-factors, other CRC associated molecules.</td>
<td>HPLC-GC/MS-MS</td>
<td>[207]</td>
</tr>
<tr>
<td>Human faeces</td>
<td>31 CRC, 31 controls with positive FOBT</td>
<td>Discrimination of CRC samples with better specificity and sensitivity than FOBT.</td>
<td>SIFT-MS</td>
<td>[210]</td>
</tr>
<tr>
<td>Lyophilized human faeces</td>
<td>42 CRC, 89 healthy controls</td>
<td>Microbe-metabolite correlation in CRC: ((\gamma)) Clostridia, Lactinospiraceae, (\gamma)-aminobenzoate and conjugated linoleate, ((\gamma)) Fusobacterium, Porphyromonas, (\beta)-hydroxy-bernaldehyde, and palmitoyl-sphingomelin.</td>
<td>HPLC-GC/MS-MS</td>
<td>[211]</td>
</tr>
<tr>
<td>Human faeces</td>
<td>13 CRC</td>
<td>Discrimination of CRC samples with 7 metabolites: aldehydes, isovaleric acid, N1-methyl-2-pyridine-5-carboxamide, 7-ketodeoxycholate, deoxycholate, valerate, and tryptophan.</td>
<td>UPLC-MS/MS; GC/MS</td>
<td>[165]</td>
</tr>
</tbody>
</table>

CRC: Colorectal cancer; CT: Computed tomography; gFOBT: Guaiac fecal occult blood test; NMR: Nuclear magnetic resonance; MS: Mass spectrometry; GC-MS: Gas chromatography coupled with MS; LC-MS: Liquid chromatography coupled with MS; UPLC: Ultra-performance liquid chromatography; HPLC: High-performance liquid chromatography; TOF: Time of flight; VOC: Volatile organic compounds; SIFT-MS: Selected ion flow tube MS.
Investigations are thus needed to elucidate the potential role of the microbiota on immunotherapy efficacy and toxicity in preclinical models and patients. Gut microbiota profiling of CRC patients might be useful in the future to predict treatment responses and/or side effects of cancer chemotherapies and immunotherapies. All these data could lead to “personalized medicine 2.0” that is being developed in an ongoing study investigating the impact of the intestinal microbiome on treatment responses and the toxicity of capecitabine or TAS-102 chemotherapies in patients with metastatic and/or resectable CRC.

Perioperative antibiotic prophylaxis strategies targeting the gut microbiota are widely used in colorectal surgery. The aim of such approaches is to prevent both surgical site and wound infections, protecting against potential contamination by the microorganisms that are spread from surgically induced alterations of the intestinal epithelial barrier. However, these antibiotic prophylaxis strategies may also impair the beneficial impact of the gut microbiota on intestinal immune function and healing (for review[246]). Few studies have investigated the impact that bowel preparation prior to colorectal surgery may have on the mucosa-associated and luminal colonic microbiota[247]. After the suppression of beneficial bacteria, the host may lose its colonization resistance to pathogens and may have decreased local stimulation of systemic immunity and the healing process[248,249]. After surgery, if the commensal microbiota does not repopulate the tract, the situation can lead to the development of pathogenic bacteria such as C. difficile[250]. Moreover, data from animal and human studies have suggested that some of the bacteria present at the anastomotic site may respond to surgical stress by activating virulence pathways, which results in alterations in the healing process. An investigation of the faecal microbiota of patients before and after colorectal surgery for CRC revealed that Enterococcus faecalis and Pseudomonas aeruginosa were the dominant pathogens present postoperatively in the stools, with a several log-fold increase during the postoperative recovery period[251]. Postoperative sepsis (e.g., anastomotic leakage, perianastomotic abscess) is the most feared complication after colorectal surgery and is responsible for potentially deadly complications and significant alterations in the patient’s quality of life. The critical role of the intestinal microbiome in sepsis has been illustrated by studies on germ-free animals, which demonstrated improved survival in these animals compared with conventional animals[252]. Therefore, the restoration and/or maintenance of a microbiota favouring intestinal healing and preventing surgical site infections after colorectal surgery could be a promising approach for the development of new therapeutic strategies, thus targeting the gut microbiota to improve surgical outcomes. Moreover, the parallel development of tools, such as the “personalized microbiota composition analysis”, to be performed pre- and postoperatively to evaluate the clinical relevance of gut microbiota modulation to positively influence the clinical outcomes and to optimize the perioperative strategies appears promising.

**Limitations of microbial markers and future challenges and directions**

The complexity of the microbiome turns the need for microbial marker-based diagnosis techniques into a real challenge. Numerous studies have reported associations between microbial markers, such as F. nucleatum, or colibactin-producing E. coli, and CRC, but to date, there is no universal microbial marker defined for CRC detection. Several limitations should be taken into consideration for the future development of new tests. First, the very high variability of the microbiota composition between individuals due to sex, age, diet, lifestyle, genetic background, medication use, ethnicity, or geographical location make finding a universal microbial marker almost impossible. Antibiotic therapy greatly influences the expression of microbial markers and is a critical limitation to microbial marker use. Moreover, standardization in terms sample collection and storage, DNA or metabolite extraction, sample analysis, and data processing is essential to compare studies. If studies that screen for microbiota composition use the same method of 16S rRNA sequencing, the use of metabolomics technologies will require exhaustive standardization between studies. Stary et al.[253] shared a very interesting and detailed list of recommendations for using microbial markers for CRC screening. (1) Any studies should be prepared carefully, taking into account the recommendations and limitations of techniques published previously. (2) Validation of CRC screening markers on specific populations should be encouraged because differences in the gut microbiome are observed in different geographical locations or in different racial/ethnic groups. (3) Whenever possible, conventional culture should be used to confirm the findings from sequencing studies. Particularly, the candidate marker status of species or genes revealed by molecular techniques should be confirmed or refuted by culture and vice versa. Systematic high-throughput culturomics should be developed because cultivation represents an approach that is economical. (4) Screening techniques for CRC risk should evaluate all the known candidate markers, combining particular species, genotoxin production and possibly further strain characteristics whenever relevant. And (5) the potential for practical detection should always be considered. For example, tumour screening, which requires colonoscopy, is costly, uncomfortable for patient and is the gold standard tool for CRC screening. Faecal samples are better non-invasive specimens for developing these microbial-associated-markers. However, faeces are likely to contain a large number of microbial species unrelated to the disease site, which may introduce noise in the detection of potential biomarkers of the disease. Because some bacteria that are associated with CRC, such as F. nucleatum, are indigenous to the human oral cavity, analysing the oral microbiome may be an alternative screening method.
for CRC. Saliva is a biological fluid that may be suitable for biomarker detection. The oral microbial compositions may theoretically reflect the oral and general health status. Flemer et al. found that several oral taxa (such as Streptococcus and Prevotella spp.) were differentially abundant in CRC patients versus controls. Moreover, they developed a classification model based on the oral swab microbiota that distinguished individuals with CRC or polyps from controls (sensitivity: 53% CRC, 67% polyps; specificity: 96%). Importantly, when data from both the faecal and oral swab microbiota were considered in this model, the sensitivity increased to 76% for CRC and 88% for polyps. In addition to Stary’s recommendations, we showed that microbial markers might be different depending on the tumour CRC phenotype. Moreover, a deeper understanding of the gut microbiota structure and function may help to identify several bacteria that when combined may provide a real CRC-associated microbial signature.

Finally, the therapeutic efficacy of anticancer drugs could also be improved by active modulation of the gut microbiota through the use of probiotics, prebiotics or specific inhibitors. This perspective is supported by the fact that immunotherapy resistance observed in germ-free mice, in antibiotic-treated mice, and in those that have previously received faecal microbiota transplantation (FMT) from non-responder patients, can be reversed by FMT from responder patients. Moreover, it has been highlighted that specific bacterial species, such as Lactobacillus johnsonii, Enterococcus hirae, Barnesiella intestinohominis, Akkermansia muciniphila, Bacteroides, and Bifidobacterium species, are beneficial in this context, which may lead the way to innovative "oncobiotics" strategies that combine anticancer and microbiota-targeting agents.

CONCLUSION

In this review, we discuss both predictive and prognostic microbial-associated markers identified in CRC. The faecal-associated microbiota may be dynamically linked to colon cancer, which, in turn, may offer evidence for microflora-associated diagnostic, preventive, and prognostic approaches for CRC. However, it is clear that additional clinical studies are necessary to validate these parameters to improve the diagnosis and therapeutic management of CRC.

REFERENCES

23. Sali L, Mascali E, Falchini M, Ventura L, Carozzi F, Castiglione G, Delsanto S, Mallardi B, Mantellini P, Milani S, Zappa M, ...
be improved by follow-up? Prospective study of 909 patients.
Chirurgia 1989; 115: 228-36; discussion 236-7 [PMID: 2805932]

DOI: 10.1056/NEJM19900230226002]


58 Villéger R et al. Microbial markers for colon cancer screening


69 Richman SD, Seymour MT, Chambers P, Elliott F, Daly CL, Meade AM, Taylor G, Barrett JH, Quirke P. KRAS and BRAF mutations in advanced colorectal cancer are associated with poor prognosis but do not preclude benefit from oxaliplatin or irinotecan: results from the MRC FOCUS trial. J Clin Oncol 2009; 27: 5931-5937 [PMID: 19884549 DOI: 10.1200/JCO.2009.22.4295]


Microbiota, Microbiome, and Cancer: The Gut Microbiome--Masters of Host Development and Physiology


June 14, 2018


cancer reveal the presence of different bacterial and archaeal biomarkers. *J Gastroenterol* 2015; **50**: 167-179 [PMID: 24811328 DOI: 10.1007/s00535-014-0963-x]


156 Dumas ME, Maibaum EC, Teague C, Ushima H, Zhou B, Lindon JC, Nicholson JK, Stabler J, Elliott P, Chan Q, Holmes E. Assessment of analytical reproducibility of 1H NMR spectroscopy...


WJG | www.wjgnet.com

2344
June 14, 2018 | Volume 24 | Issue 22 |


Nakayama H, Kinouchi T, Kataoka K, Akimoto S, Matsuda Y, Ohnishi Y. Intestinal anaerobic bacteria hydrolyse sorbinivide, producing the high blood concentration of 5-(E)- (2-bromovinyl)uracil that increases the level and toxicity of the antitumor drug 5-fluorouracil. Pharmacogenetics 1997; 7: 35-43 [PMID: 910360]


Villéger R et al. Microbial markers for colon cancer screening

24264989 DOI: 10.1126/science.1240527


Sassone-Corsi M, Raffatellu M. No vacancy: how beneficial microbes cooperate with immunity to provide colonization resistance to pathogens. *J Immunol* 2015; 194: 4081-4087 [PMID: 25888704 DOI: 10.4049/jimmunol.1403169]


