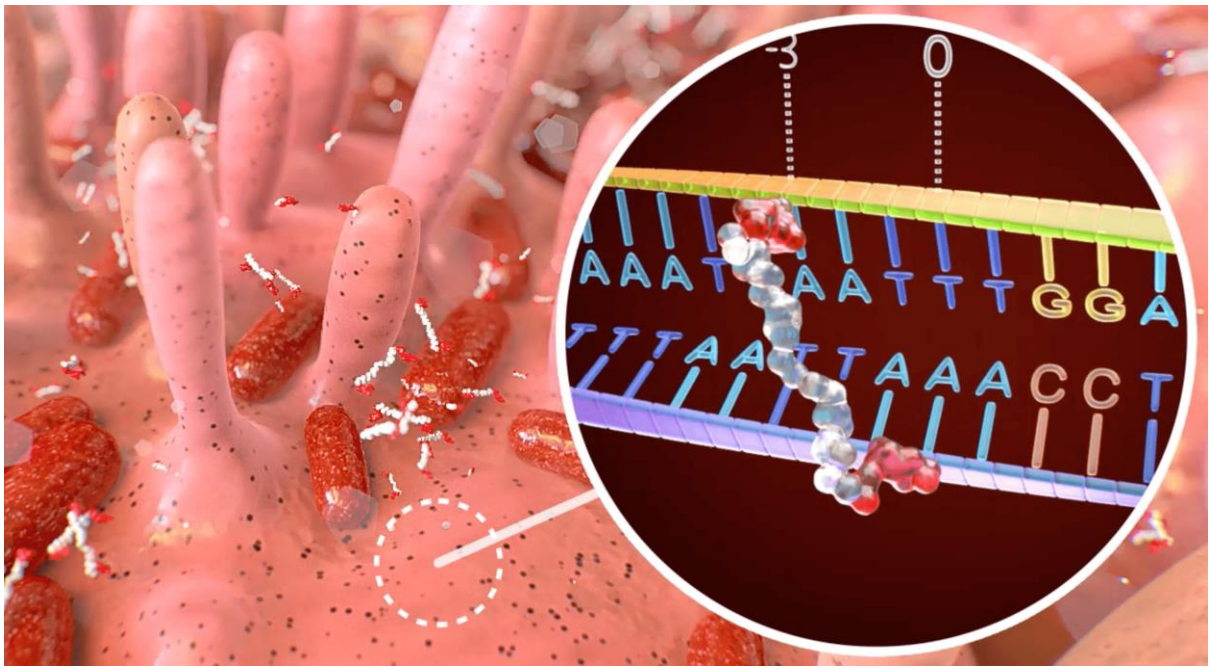
university of groningen / faculty of science and engineering

Bachelor's thesis

Molecular mechanisms through which pathogenic *E. coli* can contribute to colorectal cancer development.

Author: Jesper Post, s4111419

Thesis supervisor: Dr. ir. H.J.M. Harmsen



Colibactin from E. coli crosslinks DNA and may cause colorectal cancer (derived from S. Ktori et al. 2023).

June 26, 2023

Abstract

In this review the current knowledge is shown on multiple aspects of colorectal cancer (CRC). The incidence of CRC is on the rise, especially for patients below fifty years of age. This predicts a significant increase in the healthcare burden of the disease, both for the individual and for society. CRC is predominantly a disease with a spontaneous genetic onset, in which genes like *apc* and *TP53* are mutated. Often these mutations increase proliferative signalling of the colonic epithelial cells. Evidence amounts for the involvement of pathogenic strains of *E. coli* which have a pathogenic *pks+* virulence island. This codes for a bacterial toxin called colibactin, which has a genotoxic effect because it can form crosslinks between DNA strands. In attempts by the host cell to repair this damage, mutations occur which can induce the onset of colorectal cancer. Furthermore, DNA repair systems are also undermined by an *E. coli* effector protein, *EspF*, which makes this a two-hit system. Research has also challenged healthy benefits of Nissle *E. coli* strains and oligosaccharide probiotics by connecting these to colorectal carcinogenesis. This knowledge can be applied to diagnostics, prognostics, and therapeutics. By combining standard screening with new molecular diagnostics, we can look for oncobacteria in samples of both at-risk (a)symptomatic individuals and patients. The genotoxin colibactin could be a target, by either blocking its active site or reducing its transcription. Finally, by combining early detection with therapeutics, disease outcome can be improved and treatment burden reduced.

Table of contents

Introduction

Colorectal cancer (CRC) is the third biggest cancer type worldwide with only lung and breast cancer affecting more people. This makes the disease one of the biggest health challenges in the 21st century (WCRF international, 2022). In the past decades, the prevalence of the disease increased, especially in those under the age of fifty. Even though the disease outcome has improved, this is a worrying trend. The cause of the increase in prevalence is not yet completely understood, but it could be linked to a sedentary lifestyle and Western diet.

Despite advances in treatment, some cases of CRC are hard to treat. Especially advanced stages of the cancer and relapsed disease patients have a poor prognosis. A lot regarding the development of the disease remains unknown. With an increasing amount of literature on the subject, it has been established that the gut microbiome (GMB) plays a key role in both the development and the treatment of the disease (Candela et al., 2014).

The GMB is a key determinant of health in humans (Cresci et al., 2015). The GMB consists of a complicated ecosystem of microbial species, with a remarkable difference across populations and individuals. One of the functions of the healthy gut microbiome is protection against colonization by invasive pathogens, this is called colonization resistance. However, the gut flora can also turn against its host, in a process that is called dysbiosis (Ducarmon et al., 2019).

Research investigating the role of microbial gut species in colorectal carcinogenesis has been accelerated by advances in sequencing techniques like 16S RNA, high-throughput and whole genome sequencing (Rifaie et al., 2022). These have shown multiple species to be connected to CRC, with some studies linking *Fusobacterium nucleatus* and *Bacteroitetes fragilis* to various stages in the disease (Dai et al., 2018). Next to these candidates also a relatively more well-known gut inhabitant has been suspected to be also involved: *Escherichia coli* (Sarshar et al., 2017).

E. coli is a model organism and a common inhabitant of the human gut. The bacteria usually live in good health with the human host. This is called commensalism. Pathogenic strains of *E. coli* are linked to enteric disease, which is thought to kill up to 2 million humans each year (Tenaillon et al., 2010). The pathogenic effect does not only cause diarrhea since certain *E. coli* strains can contribute to colorectal cancer development through the secretion of bacterial toxins (Wang et al., 2023).

This review will try to answer the question ‘Through which mechanisms can pathogenic *E. coli* contribute to the development of colorectal cancer?’. First, the current situation regarding the disease will be summarized, followed by data which supports the link to pathogenic *E. coli*. After some relevant information regarding pathogenicity of different strains has been discussed, the

review will continue by investigating the effect on cellular processes of *E. coli*-produced metabolites. Finally, some implications of this information will be linked to future perspectives for diagnosis, additionally prevention and treatment of CRC will be discussed.

Colorectal cancer: current impact and disease profile

Incidence

As already mentioned, CRC is the third biggest cancer type worldwide. Data from 2019 showed a prevalence of 36 in 100.000 in the United States for CRC, to which this year 150.000 new colorectal cancer cases are expected. Age is a main risk factor for disease. This is reflected in the fact that under the age of 50, prevalence between each 5-year group showed an increase between 80 and 100 percent. CRC is slightly more prevalent in men than in women (Siegel et al., 2023).

Incidence of CRC has decreased in older adults, but in young adults the trend is opposite, with rates between ages 20 to 39 increasing since the 1980s. Occurrence of CRC below the age of fifty is called early-onset colorectal cancer (EO-CRC). *Figure 1* shows the worldwide prevalence in EO-CRC (Pan et al., 2023). Furthermore, this EO-CRC is the leading cause of death in men below the age of 50. The increase in EO-CRC offers an opportunity to investigate the effect of the contemporary factors, since most CRC cases in older age groups are the effect of long-term effects accumulating during in life. (Siegel et al., 2023).

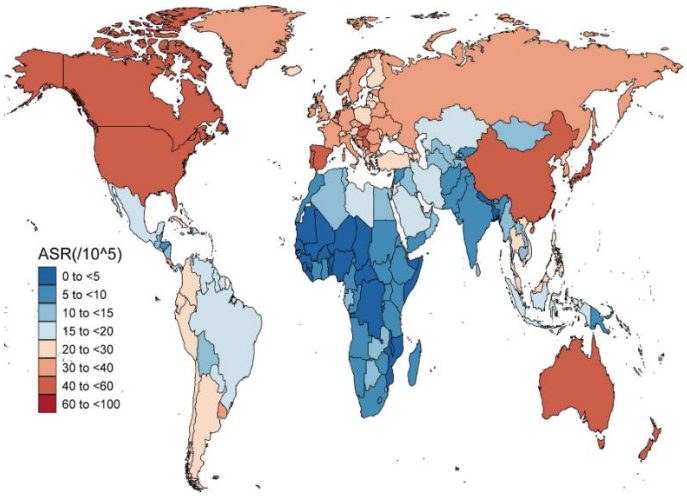


Figure 1: Age-standardized prevalence of early-onset CRC in 2019. Adapted from Pan et al., 2023.

These trends in the US fit in with worldwide trends of CRC statistics. The most recent data is from 2020, which had 1.9 million new cases of CRC. The Netherlands registers as country with the fourth highest incidence rate, with 41 individuals per 100.000, compared to a worldwide rate of 19.5 per 100.000 (WCRF international, 2023). This fits within the observation that CRC is a disease that strikes mainly people in the Western world, as these rates are comparable to many European and developed Asian nations.

Increased screening caused the stage at which the disease was being diagnosed to decline for years. Unfortunately, this trend has been reversed and the stage at which cancer is diagnosed has been increasing for the past few years, especially for EO- CRC (Pan et al., 2023). At more advanced stages the CRC will be harder to treat efficiently and eradicate. (E Osterman et al., 2018). To conclude, the overall risk of developing CRC during lifetime is 3-5 %. (Mármol et al., 2017).

Mortality

A death toll of approximately 50.000 is predicted for the US in 2023. Compared to previous years this is a steep decline, even though the annual decline in mortality decreased from 4% to 2%. Opposite to this, the mortality rate among individuals under the age of 50 has increased by roughly one percent per year (Siegel et al., 2023).

5-year survival rates of CRC cancer in developed countries range from 45 to 76 percent, with large variability between the sexes, geographical region and age. Less developed countries have a lower survival rate (Jiang et al., 2022). The disease state at time of diagnosis is the most important predictor of survival, with local disease having a 91 % 5-year survival rate and distant disease 14%. It is also interesting to mention that differences between different ethnicities in the Unites States are observed, with a striking difference between Alaska Native individuals (35.7 per 100.000) and white individuals (13.1 per 100.000) (Siegel et al., 2023). This offers interesting avenues for research, which can enable researchers to investigate differences in GMB composition, lifestyle and/or metabolic factors.

A reason for the increase in cases can be the increased screening for the disease. In the 1970s the first screening recommendations were made, like colonoscopy. Later sigmoidoscopy was also added to this. Better diagnostic techniques and increased fecal occult blood sampling contributed as well (Chu et al., 1994; Sears et al., 2023). A main drawback to colonoscopy is the invasiveness. CT scans can give a ‘digital’ colonoscopy, without having to perform the invasive part of the treatment (Nasseri et al., 2017). Moreover, in the Netherlands a large population research is currently ongoing in which the stool of participants, which are all 50 years or older, is analysed for the presence of blood. Bloody stool can be an indicator for CRC (Baer et al., 2017).

Etiology

Most colorectal cancer cases (70 %) develop sporadically. This indicates involvement of environmental, personal risk factors like alcohol consumption, diet, microbiome composition and/or a sedentary lifestyle. (Mármol et al., 2017). Other risk factors include the presence of other diseases.

People suffering from an inflammatory bowel disease like Ulcerative Colitis or Crohn's disease also have an increased chance of developing CRC (Ekbom et al., 1990; Rutter et al., 2006).

In 30 % of all CRC cases a genetic factor can be determined, with 25 % having a familial and 5 % a hereditary origin (Mármol et al., 2017). For familial CRC it is hard to define a single genetic cause. Classifying this is complicated by the complex gene-environment interactions. Suspected factors are epigenetic alterations of onco- and tumor-suppressor genes and post-transcriptional gene regulation by micro RNAs (Stoffel et al., 2014). The presence of a hereditary factor in these patients is likely, since many have a relative who suffered from the disease.

In hereditary colon cancer a somatic mutation is often carried on one allele, with a mutation in the second corresponding allele being enough to lose the entire gene in the colonic epithelial cell. An example of this is a mutation in alleles coding for DNA mismatch repair proteins like MLH1 and MSH2 in Lynch syndrome (Lynch et al., 2015). This allows benign polyps to grow, with somatic mutations being acquired as the cells divide, ultimately resulting in a carcinoma. In addition, familial adenomatous polyposis (FAP) predisposes individuals to the development of CRC through the increased formation of potentially malignant polyps in the colorectal area.

Mutations in CRC

Three types of mutational instabilities can be classified in CRC, with chromosomal instabilities being most common followed by microsatellite and epigenetic instability.

A common chromosomal instability mutation in spontaneous-onset CRC are germline point mutations in the adenomatous polyposis coli (*apc*) gene, which cause the gene to lose its function. This tumor-suppressor gene is multi-functional and codes for the APC protein. One of the functions of this APC protein is controlling beta-catenin concentrations. In a cell with a functional APC gene, beta-catenin is phosphorylated and ubiquitinated. These protein modifications prevent translocation to the nucleus, where the beta catenin acts as a transcription factor for proto-oncogenes. (Goss et al., 2000; Su et al., 1993). Furthermore, catenins activate the Wnt-signaling pathway. This pathway signals cellular survival (Fodde, 2002; Fodde et al., 2001). This corresponds to the 'sustaining proliferative signaling' hallmark of cancer (Hanahan et al., 2000; Hanahan, 2011). A mutation in the *apc* gene causes the formation of a polyp, which is also called a non-malignant adenoma.

Other key mutations can occur, for example in KRAS and PI3K genes, which are involved in MAPK signaling, which increases proliferation signalling. The TP53 gene is also often mutated, which causes a loss of control over cell cycle entrance by the loss of the p53 checkpoint protein (Mármol et al., 2017).

Already in 1990, a model was established which associated mutations in the *TP53*, *KRAS* and *DCC* genes. This model hypothesized that a benign adenoma progresses to a malignant carcinoma over a long period of time as mutations accumulate (Fearon et al., 1990).

Microsatellite instabilities involve the loss of repair mechanisms, which causes a hypermutable phenotype. This results in the inability to repair tandem repeats which can cause loss of genes due to frameshifts. Common epigenetic instability is characterized by silencing of (promotors of) tumor suppressor genes and activation of oncogenes (Mármol et al., 2017).

The clinical picture

Onset of CRC

CRC often starts with benign polyps in various places in the intestine. As described above, after certain intracellular events these can spontaneously progress to a malignant tumor. Once the tumor has reached a certain size the first symptoms arise. These include pain due to obstruction of the colon and bloody stool. This is often the moment when patients first seek medical assistance.

The cancer is divided into multiple stages based on metastasis, involvement of lymph nodes and the tumor depth. Major distinctions between stages I to IV are the depth of penetration (stage I and II) the spread to the lymph nodes (stage III) and other organs (stage IV) (Mahmoud, 2022).

Current treatment of CRC

Currently, colorectal cancer is treated in different ways, divided in curative and palliative treatment. Often the cancer is removed from by operating, with chemo- and radiotherapy acting as support. In recent years immunotherapy has also become standard practice for cancers which are harder to treat.

For stage I and II CRC an operation is often performed, cancer size can be decreased by first using chemotherapy. Nearby lymph nodes are analysed to exclude further spread of cancer. After the operation patients sometimes need a colostomy. For chemotherapy multiple agents are available, the intensity and the duration of the cure depends on the stage of the cancer and the treatment burden. Radiotherapy is not the preferred option since the bowel is particularly radiation sensitive. Regular check-ups are required to detect relapses early (Cunningham et al., 2010).

The chance that the cancer recurs is dependent on the stage of the initial cancer. Population-wide research in Sweden has shown that the chance of a relapse strongly increases from stage I to stage III CRC, from 5 to 33 % (E Osterman et al., 2018). Stage IV CRC is often considered terminal, and the

focus is on alleviating disease symptoms and prolonging the life of the patients. Metastasis in the liver or lungs is common.

E. coli in the gut microbiome

E. coli is a universal gut inhabitant.

Escherichia coli is a gram-negative, non-sporulating aerobe (Tenaillon et al., 2010), which can also respirate anaerobically (Jones et al., 2006). It can be found across warm-blooded animals and reptiles, inhabiting the gut microbiome of 90% of humans (Mitsuoka et al., 1973). *E. coli* is predominant aerobe bacteria in the gut microbiome but outnumbered by anaerobes 1 to 100-10.000. The relationship with the host can be described as commensal or symbiotic. Discussion persists if the relationship is profitable for both host and micro-organism or only for the micro-organism (Tenaillon et al., 2010). For example, bio-engineered and natural *E. coli* Nissle strains have been shown to have beneficial probiotic properties in hosts suffering from diarrhea and inflammatory bowel diseases. This effect is thought to be due to antagonistic effects towards other gut inhabitants and a stimulating effect on intestinal barrier immunomodulation and reinforcement (Schultz, 2008; Sonnenborn, 2016). Contradictory, the beneficial probiotic effect of the Nissle strain has been linked to a mutational signature which is also found in CRC patients. Questions have been raised if this probiotic is purely commensal, since genotoxic effects of the bacteria have been observed which are linked to a toxin named colibactin (Nougayrède et al., 2021).

The population structure of *E. coli* in the gut is divided into multiple lineages, determined by factors like the host and the gut micro-environment. Additionally, acquired virulent factors can help different strains of *E. coli* proliferate in the gut environment. A couple main phylogenetic groups have been determined: A, B1, B2, D and F (Herzer et al., 1990). Some of these phylogroups have been associated with certain virulence factors and CRC. This will be elucidated in the following chapters.

In an individual one strain is often dominant, amounting to more than 50 % of the strain type of the samples. In addition a second strain, called the resident strain, is present for prolonged periods (Tenaillon et al., 2020). This corresponds with the fact that between the human population and pet animals crossovers between E. coli populations can occur, with different phylotypes inhabiting the gut at various moments in time. Population wide group A is found most often across individuals, followed by B2. Dominant *E. coli* strains also differ between and within continents. Within the phylogroups, group B2 has been associated with presence in polyps and CRC lesions (Raisch et al., 2014).

E. coli that does not carry any virulence factors is considered commensal. However, not all *E. coli* are commensal. Infection with an enterotoxic strain of the bacterium can cause severe diarrhea. Multiple different types of pathogenic *E. coli* can be identified based on the biotype, with two main culprits in

CRC being Adhesive-Invasive *E. coli* (AIEC) and enteropathogenic *E. coli* (EPEC) (Hérnandez-Luna et al., 2019).

Along with the different phylogenetic groups, multiple genes to the pathogenicity of *E. coli* can be identified. These genes are often carried on virulence islands and can be horizontally transmitted between bacteria. The genes can offer a selective advantage, allowing the *E. coli* to outgrow competing *E. coli* strains. Examples are genes for adhesins, iron capture, protectins and toxins (Nowrouzian et al., 2006; Tenaillon et al., 2010).

Discovering the link between colorectal cance and *E. coli.*

Presence of E. coli is correlated to colorectal cancer

It is important to determine if the presence of *E. coli* in CRC is a cause or a consequence of the disease. Studies should be conducted looking at various stages in the disease. This includes studies looking at local colonization at the tumor site compared to the healthy gut and fecal samples between large numbers of patients and healthy volunteers. Furthermore, for a proper interpretation of a large portion of research it is important to note that many of the studies have been conducted either *in vitro* or *in vivo* in mouse models.

Already in 1998 associations were made between *E. coli* and colorectal cancer. A study compared asymptomatic and symptomatic colonoscopic biopsies with biopsies from adenomas and carcinomas of CRC patients. In 3 % of asymptomatic biopsies and 31 % of symptomatic biopsies *E. coli* could be found, and this increased to 90 and 93 % of adenomas and carcinomas. Furthermore, in 87 % of adenoma and carcinoma samples the bacteria were also found intracellular (Swidsinski et al., 1998).

Studies have shown elevated presence of *E. coli* expressing the *afa-1* operon for epithelial cell adhesion. The same study also found that the combination of *afa-1+* and *pks+* bacteria was significantly elevated in CRC patients (Prorok-Hamon et al., 2014). *Afa-1* codes for an adhesin and the *pks+* island codes for the genotoxin colibactin. This was also shown by the authors in a previous study, which showed that mucosa-associated *pks+ E. coli* was found in a significant amount of CRC patients (Arthur et al., 2012). Significant levels of *pks+ E. coli* were also found in cohorts of healthy volunteers in Canada and Japan (Oliero et al., 2022; Simpoh et al., 2017), however other studies contradict this (Nouri et al., 2021). Possible explanations could be mother-child transmission and environmental factors like diet.

Further strengthening these finding, a connection has been found between CRC biopsies of 48 patients and the pathogenic B2 *E. coli* phylogroup, which were positive for *pks* or *pks-cnf* (Raisch et al., 2014). These *pks+* strains are also more prevalent in patients with advanced CRC stages, with increased numbers of the pathogenic bacteria being found in patients with stage II or III CRC compared to stage I CRC patients (Bonnet et al., 2014).

E. coli B2 strains have been found to be elevated in polyp lesions. Other studies also found phylogenetic group D to be associated (Kohoutova, 2014). These samples were found to have high genetic diversity, showing an array of different virulent genes. Within the biopsies a gradation of *E. coli* presence was found, with increasingly less bacteria being found next to the polyp compared to on the polyp, and with none being found in healthy parts of the gut (Sarshar et al., 2017).

However, not all studies found a strong link between *E. coli* and CRC. A meta-analysis of 526 samples found a core of seven bacterial species enriched, these did not include *E. coli* (Dai et al., 2018). This does not exclude an effect of *E. coli* but could indicate a smaller role. Contradictory, another meta-analysis of 768 samples did include *pks+ E. coli* in a core set of 29 species (Wirbel et al., 2019).

In conclusion, not only the presence of certain pathogenic strains can predispose to CRC, but also a dysbiosis of the gut. A study by Couturier-Maillard et al showed that dysbiosis predisposes an individual to CRC in a murine model. Furthermore, this dysbiosis was transferred vertically from mother to child. Researchers achieved these results by knocking out the NO2 receptor in mice, which introduces a pro-inflammatory micro-environment (Couturier-Maillard et al., 2013). Another study in a mouse model showed that colibactin produced by pathogens could induce CRC in mice which showed no genetic susceptibility and were not exposed to a carcinogen, implicating an involvement of the bacteria in spontaneous disease onset. The bacteria were derived from a CRC patient (Salesse et al., 2021). Hypotheses in which the entire gut microbiome can contribute to colorectal carcinogenesis are described, in the following paragraph.

Proposed models of bacterial-induced CRC

Two main hypotheses for bacterial-induced CRC are currently proposed. The first theory is called the alpha bug hypothesis and is proposed by Sears and colleagues. The second one, which is most cited, is the passenger-driver model proposed by Tjalsma and his team.

The alpha-bug hypothesis focuses on two aspects, the first of which is pro-oncogenic properties of an intestinal bacterial species. Additionally, it also assumes a role of said bacteria in the modulation of the gut microbiome. Together both factors create an environment which has a carcinogenic effect on colonic epithelial cells.

The hypothesis as originally proposed is focussed on enterotoxic *Bacteroides fragilis* (ETBF), since their research was focused on this pathogen at the time. Some characteristics can be extended to the effects *E. coli* is known to have in colorectal carcinogenesis. These include upregulation of pathways like β-catenin-, Wnt-, and NF-kB-signaling. As established in the first chapter, the *apc* gene is often lost early in CRC development. Sears and his colleagues argue that bacterial species which proliferate because of the alpha bug can induce loss of the *apc* gene on one of both alleles, contributing to the onset of CRC. Next to *E. coli* and *B. fragilis, S. galloltyicus* is also suggested as a potential alpha bug (Sears et al., 2011).

The driver-passenger model has become the main theory on how CRC develops in an intestinal dysbiosis. The model is an analogue to the model for CRC in which (epi)genetic mutations cause an

adenomatous polyp to progress to a carcinoma over time (Fearon et al., 1990). Tjalsma discusses two types of bacteria. The driver bacteria are like the alpha bug and induce the first mutations in the colonic epithelial cells. They also induce continuous inflammation of the gut epithelial wall. During the time that the adenoma progresses to a carcinoma the composition of the gut changes. The driver bacteria can be replaced by commensals or opportunistic bacteria. These opportunists are the passengers which can further contribute to carcinogenesis. This aspect is the main distinction between both hypotheses: the alpha bug persists during tumor development, whereas driver bacteria are replaced by passenger bacteria. A key aspect of the driver-passenger hypothesis is that different composition of the microbiome exists during the progression to the carcinoma. Some possible passenger bacteria would be *Fusobacterium* or *Streptococcus* species (Tjalsma et al., 2012).

Proposed mechanisms of oncogenicity

As described, pathogenic *E. coli* are found regularly in CRC patients. These strains possess often virulent genes which can give them an advantage within the microbiome. This allows them to proliferate and infect the healthy colonic epithelial layer. Furthermore, pathogenic *E. coli* can produce multiple toxins which affect the colonic epithelial cell layer.

The genes for these toxins are encoded on virulence islands. They code for colibactin, cytotoxic disbanding toxin (CDT), cytotoxic necrotizing factors (CNFs) and effector proteins (EspFs). These genes and their transcription products have various effects on epithelial cells and the micro-environment, some of which have been shown to induce DNA damage, inflammation and therefore possibly contributing to CRC development.

Colibactin: an elusive molecule

For a long time, colibactin was hard to isolate due to its instability. Colibactin is a secondary metabolite encoded on a polyketide synthase (*pks*) island which can be acquired by bacteria through horizontal gene transfer. It is often found in *E. coli* of the phylogenetic subgroup B2 (Johnson et al., 2008). Recently, colibactin and its precursors have been isolated and some potential structures have been determined. (Zhou et al., 2021). Colibactin is often found in AIEC, which can enter epithelial cells via the CEACAM6 receptor (Barnich et al., 2007).

Colibactin has multiple identical variants called diastereomers and isomers. Colibactin is not produced in a bio-active form, first a linear precursor is formed. During activation of the pre-colibactin the blue motifs are cleaved off, triggering cyclization. This can be seen in *Figure 2*. The proposed structure of colibactin consists of two almost identical subunits, both with a so-called ‘warhead’, which is represented by the small triangles in *Figure 2*. The name for this chemical group is cyclopropane (Brotherton et al., 2013).

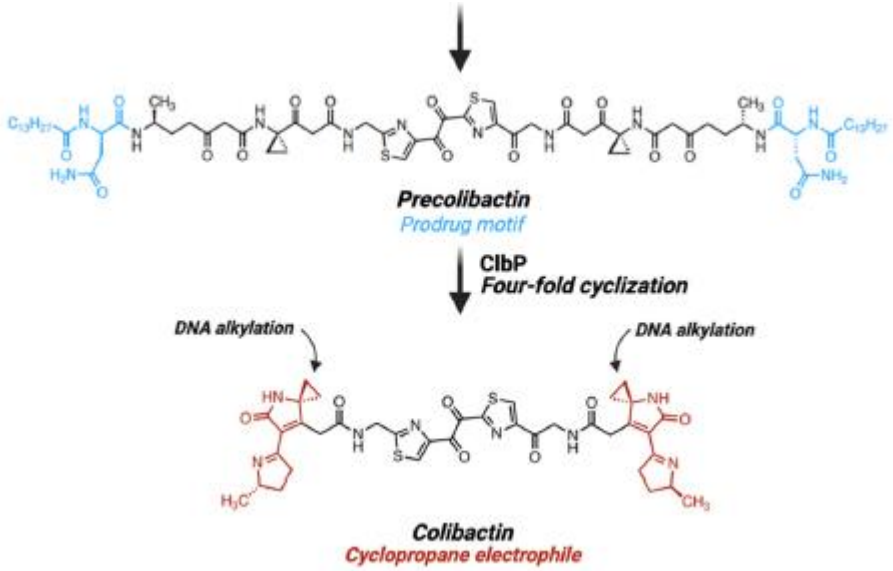


Figure 2: *The molecular structure of pre-colibactin and colibactin. The blue parts are the pro-drug motifs, the red parts the active groups of the bio-activated version (adapted from Dougherty et al., 2023)*

The genotoxic effect of colibactin is thanks to the cyclopropane groups. These groups are electrophilic. This means that they like to undergo a chemical reaction with a group which is, for example, negatively charged. In this case, the negatively charged group is the phosphate backbone of the DNA which performs a nucleophilic attack on the cyclic propane, which results in a covalent bond. The first step is depicted in *Figure 3*. This first reaction creates a second electrophilic site for the DNA to attack, due to the movement of electrons within the colibactin molecule. This second step, with DNA again as 'attacker' results in a second covalent bond and therefore a crosslink. This is called a DNA adduct. (Healy et al., 2016; Wilson et al., 2019).

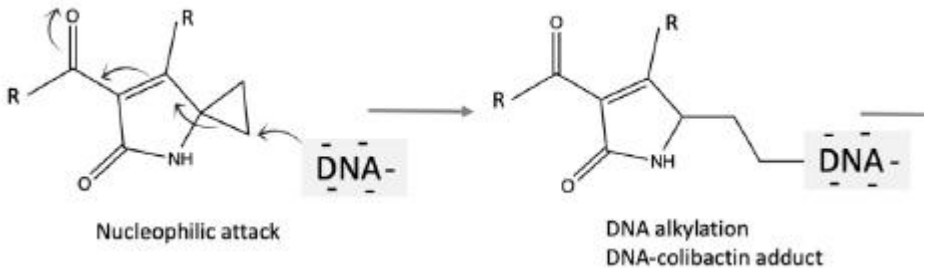


Figure 3: *Nucleophilic attack by a DNA molecule on colibactin (adapted from Mousa, 2022).*

Crosslinking induces general DNA instability and associated damage induces traditional DNA repair pathways. Often, double-strand break repair mechanisms are required. These mechanisms are known to be inaccurate, sometimes inducing a deletion or insertion of a base, also duplications are

possible (Heidenreich et al., 2003). The effect of this is that the gene is longer being transcribed into mRNA and thus the protein it encodes is no longer transcribed. This indicates that colibactin mainly damages DNA indirectly. Furthermore, certain *E. coli* can downregulate mismatch repair proteins, which further contributes to the induction of mutations (Maddocks et al., 2009).

The mechanism described above is one of the possible ways in which colibactin can induce DNA damage in colonic epithelial cells. Other ways which have been shown in mice are chromosomal rearrangements and cell cycle arrest (Wang et al., 2023).

The effects of colibactin are not only found in eukaryotic cells. The final gene on the *pks* island codes for the ClbS protein which induces self-resistance of the bacterium to colibactin. The ClbS protein exerts its protective function by destroying the cyclopropane 'warheads', essentially preventing DNA crosslinking (Bossuet-Greif et al., 2016). Next to that, the *pks* island induces prophages, which allow horizontal gene transfer between bacteria. It is suggested that bacteria which do not have the protective ClbS protein are susceptible to prophage-induced lytic developments. This indicates that the *pks+ E. coli* could shape the gut micro-environment by outcompeting other bacterial species (Silpe et al., 2022).

The EspF effector protein increases mutational frequency

EPEC are suspected to have multiple carcinogenic properties. These include inhibition of macrophages, inducing inflammation through a MAP-kinase pathway and damage to DNA repair systems. This is mediated by an effector protein: EspF. Effector proteins are secreted by pathogenic bacteria into host cells and act to enhance pathogen invasion, immune system suppression and overall pathogen survival. In colonic epithelial cells EspF depletes mismatch repair proteins by a posttranslational mechanism. A study showed an increase in mutational frequency upon infection with EPEC (Maddocks et al, 2013).

Other toxins

Cytotoxic disbanding toxins (CDTs) and cytotoxic necrotizing factors (CNFs) are other toxins produced by pathogenic *E. coli* strains.

CDT is found more often in CRC tissue compared to healthy gut tissue (Buc et al., 2013). CDT is a tetramer that enters the cell via a membrane receptor. The exact mechanisms of this process are not yet discovered. Inside the cell a subunit, CdtB, translocates to the nucleus. This protein has a high sequence homology to some DNases and thus it is suspected that it cleaves DNA, inducing mutations. Specifically, this happens in the S-phase of the cell cycle. (Elwell et al., 2000; Wang et al., 2023).

CNF is mainly known in the context of extraintestinal infections, for example meningitis. It binds to targets on the host cell surface and is after endocytosis intracellularly cleaved to form the active product. A main target is RhoGTPase which is involved in the inflammatory NF-kB pathway. It also causes increased bcl-2 levels, which is an anti-apoptotic protein (Fabbri et al., 2013). This corresponds to two different hallmarks of cancer, namely 'tumor-promoting inflammation' and 'resisting cell death' (Hanahan, 2011).

Toxins like CDT and CNF are not unique to pathogenic *E. coli* and therefore are not very relevant when considering mechanisms through which *E. coli* specifically could induce cancerous growth in colonic epithelial cells.

Therapeutic perspectives

Multiple treatment strategies can be considered, targeting different aspects of the complicated biology that was detailed in the previous chapters. In CRC treatment a high treatment burden is experienced by patients, with especially chemo- and radiotherapy reducing the quality of life. Replacing these treatments by less straining procedures could therefore bring a large benefit for individuals. Earlier detection could improve outcome as well because five-year survival rates for local disease are around 90 % and distal disease around 15 % (Siegel et al. 2023).

Screening options

We can screen for biomarkers in fecal blood samples, looking for metabolites that correspond to pathogenic strains. An example of this is screening for the *clb* gene. Compared to a colonoscopy this is advantageous both with respect to costs and invasiveness. Currently, tests for hemoglobin are already available but they are dependent on blood in the stool, which is not always the case in early-stage CRC or asymptomatic patients. For tests it is important that they have a low false positive rate due to the heavy conclusions following a diagnosis. Next to that, high sensitivity is a prerequisite.

As described in the second chapter, a lot of research has been conducted using tests which sample feces of colorectal cancer patients and controls. This has been extended to a wider array of oncogenic bacteria like *Fusobacterium nucleatum, Bacteroides fragilis* and *Streptococcus gallolyticus* (Veziant et al., 2021). Molecular diagnostic techniques for this include whole-genome shotgun sequencing and 16S rRNA sequencing. Actual implementation of this as a diagnostic tool is complicated because no conclusion has been reached as of what an oncogenic composition of the microbiota is. Another option would be combining this with the standard fecal immunochemical test. The detection of oncogenic species and the depletion of beneficial species was found to increase the accuracy of these tests (Baxter et al., 2016).

Knowledge about the composition of the microbiome can also be used to predict chemotherapy and immunotherapy efficacy and thus to an extent disease outcome. *E. coli* is one of the species enriched in patients with poor response to treatment (Veziant et al., 2021).

Targeting E. coli and its metabolites

The pathogen itself could also be targeted with an antibiotic. However, the use of antibiotics in the gut would not be very feasible due to side-effects on commensal bacteria. Therefore, discussing antibiotic use for prevention of CRC is not very relevant. This treatment could also be implemented after the first diagnosis of CRC to prevent the disease from further progressing, however also in this

case negative-side effects on commensal microbes would probably be worse than the protective effect. And again, more knowledge would be required on which bacteria can induce CRC.

Targeting colibactin itself might prove beneficial. In an article published in 2014 researchers found that the DNA damaging effects of colibactin could be reduced by using a small-molecule inhibitor for the active site. This effect was observed both *in vitro* and *in vivo* (Cougnoux et al., 2014). Also, D-serine was found to have a reductive effect on the genotoxicity of *pks+ E. coli* in an in vitro model. D-serine was suspected to exert its influence by downregulation the *clb* gene (Hallam et al., 2023). To conclude this therapeutic option, a high supply of iron was also shown to decrease colibactin synthesis through a ferric uptake regulator (Fur) and a small regulatory non-coding RNA (RyhB). The exact mechanism of this has yet to be uncovered (Tronnet et al., 2017).

To finish this section also another unconventional approach can be used. As shortly mentioned in the chapter on *E. coli*, questions regarding the safety of using the Nissle *E. coli* strain as probiotic have been raised, since it is suspected to increase likelihood of colorectal tumor formation (Nougayàrede et al., 2021). However, it has been found that by pasteurizing these bacteria may offer a solution. This forms something that is called a parabiotic. Due to the pasteurization the instable colibactin is neutralized, while the beneficial effects of the Nissle strain are retained. The beneficial effects include the outcompetition of pathogens by the Nissle *E. coli.* This parabiotic has in addition also anti-tumorigenic effect, which makes it a promising therapeutic (Oliero et al., 2023).

Focus on a symbiotic gut micro-environment

However, if we want to get rid of pathogens, we can also induce growth of commensal strains and rely on the dynamics of the gut ecosystem to return to a symbiotic state. This could be done with dietary interventions. A meta-analysis compared 26 studies which aimed to improve gut health by administering fermentable non-digestible carbohydrates (Fratila et al., 2023). Examples include fruit-oligosaccharides and other inulin-type probiotics. Most of these studies succeeded in decreasing the levels of pathogenic bacteria, but they all failed to record levels of pathogenic *E. coli*, instead focussing on better known pathogens like *Bifidobacteria* and *Fusobacteria*.

To extend on this topic, recently a correlation was found between oligosaccharides and an increase in the genotoxic effect of colibactin. Generally, oligosaccharides are considered healthy and a prebiotic for commensal species, which this research article contradicts (Oliero et al., 2021). A well-known prebiotic is inulin, although its effects can vary among individuals. Oliero et al found that inulin can increase colonization by *pks+ E. coli* in a $Apc^{Min/+}$ mice model (Oliero et al., 2023). Furthermore, the same research found that this same colonization resulted in an increased number of intestinal

tumors. Since the presence of *E. coli* with the ability to produce colibactin can differ between individuals (Nouri et al., 2021; Oliero et al. 2022; Simpoh et al., 2017), administration of this prebiotic should not be done without analysis in advance.

I want to conclude this chapter by repeating the key role decisions regarding lifestyle play in health and disease, regardless of which disease an individual suffers from. In colorectal cancer this has also been shown (Song et al., 2019). Therefore, lifestyle modulation should be a core aspect of each treatment option.

Discussion

The evidence that exists for the oncogenic activity of pathogenic *E. coli* strains seems very strong, however, the facts fit within a bigger, more complicated picture.

First, the question arises if *E. coli* is a cause or an effect of the disease, like Trudy Wassenaar argued in her 2018 review (Wassenaar, 2018). Due to the enormous complexity of the colonic microbiome, it is hard or even impossible to pinpoint a single species as being the dominant cause of the formation of a colorectal tumor. Microbiome composition tends to fluctuate over time due to a variety of reasons. Thinking along the lines of the 'driver-passenger' hypothesis of Tjalsma et al, pathogenic AIEC and EPEC could play an important role in the development of the cancer as a driver bacterium. Adherent invasive *E. coli* can invade epithelial cells and induce mutations via a colibactin-driven mechanism, which becomes a two-hit system when combined with the effector protein EspF secreted by enteropathogenic *E. coli*. Furthermore, the bactericidal properties of colibactin in prokaryotes provide a selective advantage, suggesting that pathogenic *E. coli* could outcompete other (pathogenic) bacteria.

To confirm this hypothesis, two types of studies need to be conducted. First, a long-term study containing a large cohort of patients in various disease stages, at-risk non-symptomatic individuals and healthy controls should be tracked for many years to determine changes in microbiome composition over time. This study should take regular fecal samples to have a reliable indication of the colonic micro-environment. This requires cutting-edge sequencing techniques to make the study feasible, like deep metagenomic sequencing. This type of sequencing combines high specificity with high sensitivity and allows for detection of low-abundance species in the gut microbiome. This is possible because larger parts of the bacterial genome are analysed, however this also offers a computational challenge since way more data is put in (Jin et al., 2022). When a suspected oncobacterium is detected, a species-specific growth medium can be used to confirm the read-out. The study described above will most likely point to multiple pathogens being involved, and with this information and relative abundance of these pathogens a partial conclusion could be drawn about the role of *E. coli* in carcinogenesis.

A study like the one described above could determine which pathogens have the strongest association with the onset of CRC. Next to that the bactericidal effect of colibactin should be studied and how this can shape the gut niche. It would be relevant to construct multiple *in vitro* models, in which common gut inhabitants and pathogens are cocultured with *clb+ E. coli*. These studies could elucidate if harbouring the gene for colibactin gives a selective advantage. If this study shows that *clb+ E. coli* significantly models the gut microbiome along with a strong association with CRC as

described above, we could conclude that pathogenic *E. coli* plays a central role in the development of colorectal cancer.

It is however very likely that pathogenic *E. coli* is not the only factor that influences the development of a colorectal carcinoma. Numerous studies can be found that link other bacterial species to the disease, for example species like *Fusobacterium nucleatum* and enterotoxic *Bacteroides fragilis* (ETBF) in more advanced stages of CRC (Viljoen et al., 2015). Furthermore, decreased levels of commensal bacteria can increase an individual's susceptibility to CRC these bacteria prevent colonization of the gut mucosa in a healthy state. Significant decrease in species like *F. prausnitzii, Barnesiella intestinihominis, Alistipes finegoldii, Bacteroides eggerthii* and *Eubacterium siraeum* were found in CRC patients (Touchefeu et al., 2021). This raises the question if the focus in treatment should be on repelling pathogens, or increasing commensals which would in theory outcompete the intruders. This seems plausible at first, but since the exact composition of the gut at any time is unknown and quick action can be required, it would be better to have tools with which we can repel these pathogens. Like mentioned in the chapter on therapeutics, pasteurized Nissle *E.* coli might offer a possibility.

From the same chapter we can conclude that the use of pathogenic *E. coli* as a biomarker is very promising. On a side note, the bacteria must be present at the time of testing, which is in theory not always the case following the 'driver-passenger' hypothesis. To develop a proper toolkit for screening the scope should be increased to include a larger number of species, so that the specificity will increase. Also as already mentioned, deep-metagenomic sequencing can be used to increase the sensitivity. These kinds of test could be used in the future as a replacement for the relatively costly and invasive colonoscopy. Tests indicating possible CRC could then be followed up with a colonoscopy. This could possibly revolutionize the screening process for CRC, widely increasing its availability and accessibility. Of course, this can only be done if the two study possibilities that were described above yield conclusive results. Recently it has also been shown that humans can be roughly divided into three so-called enterotypes (Arumugan, 2011). These enterotypes should be taken into consideration when designing said tests.

These tests will be a step in the direction of personalized medicine, as advances in screening technology will make it cheaper and accessible for more people. My vision is that once screening becomes so ubiquitous it can be done on massive scale, an individual's microbiome can be modelled accurately, predicting development of a CRC-inducing dysbiosis. Using this information will require an integration of biology, data science and statistics. Once we are accurately able to predict the composition of the gut microbiome composition, we can start to administer personalized treatment

to at-risk individuals and CRC patients. This can take a preventive shape which could reduce the treatment burden chemo- or radiotherapy otherwise could have given. Recent developments in the field also point in the direction of a molecular & pathogenic epidemiology field, which would combine data science with microbiology and pathology (Ogino et al., 2019). Successful application of this principle stretch way further than CRC only, since the gut microbiome is involved in a wide array of disease states (Cresci et al., 2015).

This would not be the only innovative route to detect CRC in early stages. Research groups have been investigating the 'smell' of cancer, which is caused by a group of molecules called volatile organic compounds (VOCs). VOCs show great potential as a biomarker for CRC. Multiple routes of detection are investigated, with one being the feces (de Boer et al., 2014). The use of the fecal gas as a sample has been confirmed in a proof-of-concept study comparing patients with advanced CRC and healthy controls (de Meij et al., 2014).

Possible targets of treatment could include colibactin. Multiple studies have shown different targets of colibactin biology and if a general overview of its associated pathways is known, even more therapeutic agents can be designed. However, any treatment should not be limited to only colibactin since pathogenic *E. coli* and other pathogenic species can induce DNA damage in with a wide array of metabolites and toxins. Furthermore, also non-biological agents can induce CRC with an example being heterocyclic amines which are often ingested with red meat consumption (Marchand, 2021). Therefore, any treatment focused on colibactin is never a complete prevention of CRC, and responsible prevention management should also focus on modulation of unhealthy lifestyle habits.

Treatments should however be designed with care. While it is desired that we eradicate pathogens from the gut, any side-effects should be well-studied so that we don't disturb any beneficial effects from the commensal microflora.

In this review I hope to have shone a light on a small piece in the enormously complicated world of the gut microbiome and its role in the third most prevalent cancer type worldwide. Although knowledge now is not enough to design treatment and prevention strategies, more research and advances in the field make this microbiome-based view of colorectal cancer an interesting avenue to explore. Earlier detection of CRC can reduce mortality and reduce treatment burden, which makes this a relevant topic for the thousands of individuals suffering from CRC worldwide. I have summarized the findings and implications of this literature review in a graphical way, which can be found in Figure 4 below.

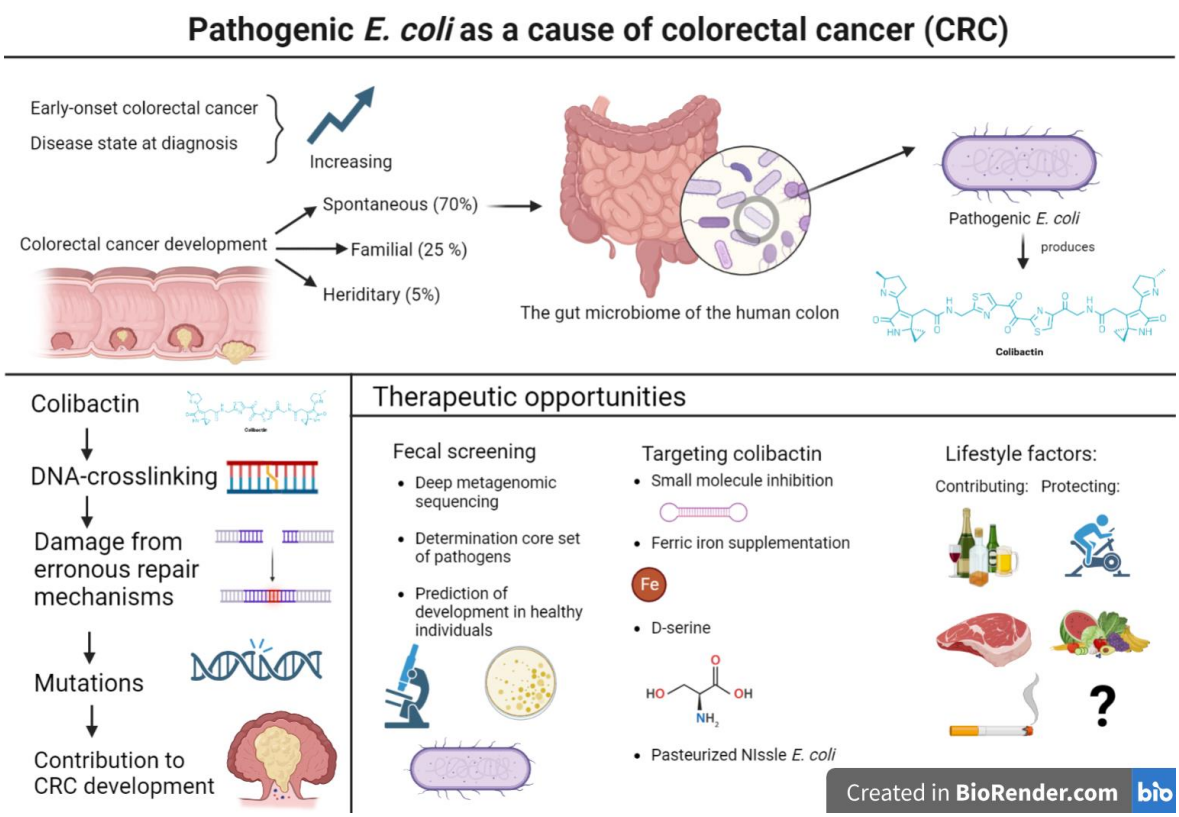


Figure 4: Molecular mechanisms through which pathogenic E. coli can contribute to colorectal cancer development and therapeutic opportunities (made with BioRender)

References

Arthur, J. C., Perez-Chanona, E., Mühlbauer, M., Tomkovich, S., Uronis, J. M., Fan, T., Campbell, B. J., Abujamel, T. S., Dogan, B., Rogers, A. B., Rhodes, J. M., Stintzi, A., Simpson, K. W., Hansen, J. J., Keku, T. O., Fodor, A. A., & Jobin, C. (2012). Intestinal Inflammation Targets Cancer-Inducing Activity of the Microbiota. Science, 338(6103), 120–123. https://doi.org/10.1126/science.1224820

Arumugam, M., Raes, J., Pelletier, E., Paslier, D. L., Yamada, T., Mende, D. R., Fernandes, G. R., Tap, J., Bruls, T., Batto, J., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., . . . Bork, P. (2011a). Enterotypes of the human gut microbiome. *Nature*, *473*(7346), 174–180. https://doi.org/10.1038/nature09944

Baer, C., Menon, R., Bastawrous, S., & Bastawrous, A. L. (2017). Emergency Presentations of Colorectal Cancer. *Surgical Clinics of North America*, *97*(3), 529–545. https://doi.org/10.1016/j.suc.2017.01.004

Barnich, N., Carvalho, F. A., Glasser, A., Darcha, C., Jantscheff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J., & Darfeuille-Michaud, A. (2007). CEACAM6 acts as a receptor for adherent-invasive E. coli, supporting ileal mucosa colonization in Crohn disease. *Journal of Clinical Investigation*, *117*(6), 1566–1574. https://doi.org/10.1172/jci30504

Baxter, N. N., Ruffin, M. T., Rogers, M. A., & Schloss, P. D. (2016). Microbiota-based model improves the sensitivity of fecal immunochemical test for detecting colonic lesions. *Genome Medicine*, *8*(1). https://doi.org/10.1186/s13073-016-0290-3

Bonnet, M., Buc, E., Sauvanet, P., Darcha, C., Dubois, D., Pereira, B., Déchelotte, P., Bonnet, R., Pezet, D., & Darfeuille-Michaud, A. (2014). Colonization of the Human Gut by *E. coli* and Colorectal Cancer Risk. *Clinical Cancer Research*, *20*(4), 859–867. https://doi.org/10.1158/1078-0432.ccr-13-1343

Bossuet-Greif, N., Dubois, D., Petit, C., Tronnet, S., Martin, P., Bonnet, R., Oswald, E., & Nougayrède, J. (2016). Escherichia coli ClbS is a colibactin resistance protein. Molecular Microbiology, 99(5), 897–908. https://doi.org/10.1111/mmi.13272

Brotherton, C. A., & Balskus, E. P. (2013). A Prodrug Resistance Mechanism Is Involved in Colibactin Biosynthesis and Cytotoxicity. *Journal of the American Chemical Society*, *135*(9), 3359–3362. https://doi.org/10.1021/ja312154m

Buc, E., Dubois, D., Sauvanet, P., Raisch, J., Delmas, J., Darfeuille-Michaud, A., Pezet, D., & Bonnet, R. (2013). High Prevalence of Mucosa-Associated E. coli Producing Cyclomodulin and Genotoxin in Colon Cancer. PLOS ONE, 8(2), e56964. https://doi.org/10.1371/journal.pone.0056964

Candela, M., Turroni, S., Biagi, E., Carbonero, F., Rampelli, S., Fiorentini, C., & Brigidi, P. (2014). Inflammation and colorectal cancer, when microbiota-host mutualism breaks. *World Journal of Gastroenterology*, *20*(4), 908. https://doi.org/10.3748/wjg.v20.i4.908

Clermont, O., Christenson, J. K., Denamur, E., & Gordon, D. L. (2013). The ClermontEscherichia coliphylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environmental Microbiology Reports, 5(1), 58–65. https://doi.org/10.1111/1758-2229.12019

Cougnoux, A., Delmas, J., Gibold, L., Faïs, T., Romagnoli, C., Robin, F., Cuevas-Ramos, G., Oswald, E., Darfeuille-Michaud, A., Prati, F., Dalmasso, G., & Bonnet, R. (2016). Small-molecule inhibitors prevent the genotoxic and protumoural effects induced by colibactin-producing bacteria. Gut, 65(2), 278–285. https://doi.org/10.1136/gutjnl-2014-307241

Cresci, G. A., & Bawden, E. (2015). Gut Microbiome. *Nutrition in Clinical Practice*, *30*(6), 734–746. https://doi.org/10.1177/0884533615609899

Cunningham, D., Atkin, W., Lenz, H., Lynch, H. T., Minsky, B. D., Nordlinger, B., & Starling, N. (2010). Colorectal cancer. The Lancet, 375(9719), 1030–1047. https://doi.org/10.1016/s0140-6736(10)60353-4

Dai, Z., Coker, O. O., Nakatsu, G., Ng, S. C., Zhao, L., Chen, Z., Chan, F. K., Kristiansen, K., Sung, J. J., Wong, S. H., & Yu, J. (2018). Multi-cohort analysis of colorectal cancer metagenome identified altered bacteria across populations and universal bacterial markers. Microbiome, 6(1). https://doi.org/10.1186/s40168-018-0451-2

De Boer, N. K. H., De Meij, T. G. J., Oort, F. A., Larbi, I. B., Mulder, C. J. J., Van Bodegraven, A. A., & Van Der Schee, M. P. (2014). The Scent of Colorectal Cancer: Detection by Volatile Organic Compound Analysis. *Clinical Gastroenterology and Hepatology*, *12*(7), 1085–1089. https://doi.org/10.1016/j.cgh.2014.05.005

De Meij, T. G. J., Larbi, I. B., Van Der Schee, M. P., Lentferink, Y. E., Paff, T., Droste, J. S. T. S., Mulder, C. J. J., Van Bodegraven, A. A., & De Boer, N. K. H. (2013). Electronic nose can discriminate colorectal carcinoma and advanced adenomas by fecal volatile biomarker analysis: proof of principle study. *International Journal of Cancer*, *134*(5), 1132–1138. https://doi.org/10.1002/ijc.28446

Ducarmon, Q. R., Zwittink, R. D., Hornung, B., Van Schaik, W., Young, V. B., & Kuijper, E. J. (2019). Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. Microbiology and Molecular Biology Reviews, 83(3). https://doi.org/10.1128/mmbr.00007-19

Eckburg, P. B., Bik, E. M., Bernstein, C. N., Purdom, E., Dethlefsen, L., Sargent, M. A., Gill, S. R., Nelson, K. E., & Relman, D. A. (2005). Diversity of the Human Intestinal Microbial Flora. Science, 308(5728), 1635–1638. https://doi.org/10.1126/science.1110591

Ekbom, A., Adami, H., Helmick, C. G., & Zack, M. M. (1990). Increased risk of large-bowel cancer in Crohn's disease with colonic involvement. The Lancet, 336(8711), 357–359. https://doi.org/10.1016/0140-6736(90)91889-i

Elwell, C. A., & Dreyfus, L. A. (2000). DNase I homologous residues in CdtB are critical for cytolethal distending toxin-mediated cell cycle arrest. *Molecular Microbiology*, *37*(4), 952–963. https://doi.org/10.1046/j.1365-2958.2000.02070.x

Fabbri, A., Travaglione, S., Ballan, G., Loizzo, S., & Fiorentini, C. (2013). The Cytotoxic Necrotizing Factor 1 from E. Coli: A Janus Toxin Playing with Cancer Regulators. *Toxins*, *5*(8), 1462–1474. https://doi.org/10.3390/toxins5081462

Fearon, E. R., & Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. Cell, 61(5), 759–767. https://doi.org/10.1016/0092-8674(90)90186-i

Fodde, R. (2002b). The APC gene in colorectal cancer. European Journal of Cancer, 38(7), 867–871. https://doi.org/10.1016/s0959-8049(02)00040-0

Fratila, T., Ismaiel, A., & Dumitrascu, D. L. (2023). Microbiome modulation in the prevention and management of colorectal cancer: a systematic review of clinical interventions. *Medicine and Pharmacy Reports*, *96*(2), 131–145. https://doi.org/10.15386/mpr-2526

Goss, K. H., & Groden, J. (2000). Biology of the Adenomatous Polyposis Coli Tumor Suppressor. Journal of Clinical Oncology, 18(9), 1967–1979. https://doi.org/10.1200/jco.2000.18.9.1967

Hallam, J. C., Sandalli, S., Floria, I., Turner, N. C. A., Tang-Fichaux, M., Oswald, E., O'Boyle, N., & Roe, A. J. (2023). D-Serine reduces the expression of the cytopathic genotoxin colibactin. *Microbial Cell*, *10*(3), 63–77. https://doi.org/10.15698/mic2023.03.793

Hanahan, D., & Weinberg, R. A. (2000). The Hallmarks of Cancer. Cell, 100(1), 57–70. https://doi.org/10.1016/s0092-8674(00)81683-9

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of Cancer: The Next Generation. Cell, 144(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Healy, A. R., Nikolayevskiy, H., Patel, J. C., Crawford, J. M., & Herzon, S. B. (2016). A Mechanistic Model for Colibactin-Induced Genotoxicity. Journal of the American Chemical Society, 138(48), 15563–15570. https://doi.org/10.1021/jacs.6b10354

Heidenreich, E., Novotny, R., Kneidinger, B., Holzmann, V., & Wintersberger, U. (2003). Non-homologous end joining as an important mutagenic process in cell cycle-arrested cells. *The EMBO Journal*, *22*(9), 2274–2283. https://doi.org/10.1093/emboj/cdg203

Herzer, P. J., Inouye, S., Inouye, M., & Whittam, T. S. (1990). Phylogenetic distribution of branched RNA-linked multicopy single-stranded DNA among natural isolates of Escherichia coli. Journal of Bacteriology, 172(11), 6175–6181. https://doi.org/10.1128/jb.172.11.6175-6181.1990

Jiang, Y., Yuan, H., Li, Z., Ji, X., Shen, Q., Tuo, J., Bi, J., Li, H., & Xiang, Y. (2022). Global pattern and trends of colorectal cancer survival: a systematic review of population-based registration data. Cancer Biology and Medicine, 18, 0. https://doi.org/10.20892/j.issn.2095-3941.2020.0634

Jin, H., You, L., Zhao, F., Li, S., Ma, T., Kwok, L., Xu, H., & Sun, Z. (2022). Hybrid, ultra-deep metagenomic sequencing enables genomic and functional characterization of low-abundance species in the human gut microbiome. *Gut Microbes*, *14*(1). https://doi.org/10.1080/19490976.2021.2021790

Johnson, J. R., Johnston, B. T., Kuskowski, M. A., Nougayrède, J., & Oswald, E. (2008). Molecular Epidemiology and Phylogenetic Distribution of the Escherichia coli pks Genomic Island. Journal of Clinical Microbiology, 46(12), 3906–3911. https://doi.org/10.1128/jcm.00949-08

Jones, S. A., Chowdhury, F. Z., Fabich, A. J., Anderson, A. B., Schreiner, D. M., House, A. L., Autieri, S. M., Leatham, M. P., Lins, J. J., Jorgensen, M., Cohen, P. R., & Conway, T. (2007). Respiration of *Escherichia coli* in the Mouse Intestine. *Infection and Immunity*, *75*(10), 4891–4899. https://doi.org/10.1128/iai.00484-07

Kohoutova, D., Šmajs, D., Moravkova, P., Cyrany, J., Moravkova, M., Forstlova, M., Cihak, M., Rejchrt, S., & Bures, J. (2014). Escherichia colistrains of phylogenetic group B2 and D and bacteriocin production are associated with advanced colorectal neoplasia. BMC Infectious Diseases, 14(1). https://doi.org/10.1186/s12879-014-0733-7

Ktori, S., & Ktori, S. (2023). Microbiome Bacteria Toxin Linked Directly to Colorectal Cancer-Driving Mutations. *GEN - Genetic Engineering and Biotechnology News*. https://www.genengnews.com/news/microbiome-bacteria-toxin-linked-directly-to-colorectal-cancer-driving-mutations/

Lynch, H. T., Snyder, C., Shaw, T. G., Heinen, C. D., & Hitchins, M. P. (2015). Milestones of Lynch syndrome: 1895–2015. Nature Reviews Cancer, 15(3), 181–194. https://doi.org/10.1038/nrc3878

Maddocks ODK, Scanlon KM, Donnenberg MS. 2013. An *Escherichia coli* effector protein promotes host mutation via depletion of DNA mismatch repair proteins. mBio 4(3):00152-13. doi:10.1128/mBio.00152-13.

Maddocks, O. D. K., Short, A., Donnenberg, M. S., Bader, S., & Harrison, D. G. (2009). Attaching and effacing Escherichia coli downregulate DNA mismatch repair protein in vitro and are associated with colorectal adenocarcinomas in humans. *PLOS ONE*. https://doi.org/10.1371/journal.pone.0005517

Mahmoud, N. N. (2022). Colorectal Cancer. Surgical Oncology Clinics of North America, 31(2), 127–141. https://doi.org/10.1016/j.soc.2021.12.001

Marchand, L. L. (2021). The role of heterocyclic aromatic amines in colorectal cancer: the evidence from epidemiologic studies. *Genes and Environment*, *43*(1). https://doi.org/10.1186/s41021-021-00197-z

Mármol, I., Sánchez-De-Diego, C., Dieste, A. P., Cerrada, E., & Yoldi, M. J. R. (2017). Colorectal Carcinoma: A General Overview and Future Perspectives in Colorectal Cancer. International Journal of Molecular Sciences, 18(1), 197. https://doi.org/10.3390/ijms18010197

Nasseri, Y., & Langenfeld, S. J. (2017). Imaging for Colorectal Cancer. *Surgical Clinics of North America*, *97*(3), 503–513. https://doi.org/10.1016/j.suc.2017.01.002

Nougayrède, J., Chagneau, C. V., Motta, J., Bossuet-Greif, N., Belloy, M., Taieb, F., Gratadoux, J., Thomas, M., Langella, P., & Oswald, E. (2021). A Toxic Friend: Genotoxic and Mutagenic Activity of the Probiotic Strain Escherichia coli Nissle 1917. mSphere, 6(4). https://doi.org/10.1128/msphere.00624-21

Nouri, R., Hasani, A., Shirazi, K. H., Nouri, M., Sepehri, B., Sotoudeh, S., Hemmati, F., Fattahzadeh, A., Abdinia, B., & Rezaee, M. A. (2021). Mucosa-Associated Escherichia coli in Colorectal Cancer Patients and Control Subjects: Variations in the Prevalence and Attributing Features. *Canadian Journal of Infectious Diseases & Medical Microbiology*, *2021*, 1–8. https://doi.org/10.1155/2021/2131787

Ogino, S., Nowak, J. A., Hamada, T., Milner, D. A., & Nishihara, R. (2019). Insights into Pathogenic Interactions Among Environment, Host, and Tumor at the Crossroads of Molecular Pathology and Epidemiology. *Annual Review of Pathology-mechanisms of Disease*, *14*(1), 83–103. https://doi.org/10.1146/annurev-pathmechdis-012418-012818

Oliero, M., Calvé, A., Fragoso, G., Cuisiniere, T., Hajjar, R., Dobrindt, U., & Santos, M. M. (2021). Oligosaccharides increase the genotoxic effect of colibactin produced by pks+ Escherichia coli strains. BMC Cancer, 21(1). https://doi.org/10.1186/s12885-021-07876-8

Oliero, M., Hajjar, R., Cuisiniere, T., Fragoso, G., Calvé, A., & Santos, M. M. (2023). Inulin impacts tumorigenesis promotion by colibactin-producing Escherichia coli in ApcMin/+ mice. *Frontiers in Microbiology*, *14*, 1067505. https://doi.org/10.3389/fmicb.2023.1067505

Oliero, M., Hajjar, R., Cuisiniere, T., Fragoso, G., Calvé, A., Dagbert, F., Loungnarath, R., Sebajang, H., Schwenter, F., Wassef, R., Ratelle, R., De Broux, É., Richard, C., & Santos, M. M. (2022). Prevalence of pks + bacteria and enterotoxigenic Bacteroides fragilis in patients with colorectal cancer. *Gut Pathogens*, *14*(1). https://doi.org/10.1186/s13099-022-00523-y

Osterman, E., & Glimelius, B. (2018). Recurrence Risk After Up-to-Date Colon Cancer Staging, Surgery, and Pathology: Analysis of the Entire Swedish Population. Diseases of the Colon & Rectum, 61(9), 1016–1025. https://doi.org/10.1097/dcr.0000000000001158

Prorok-Hamon, M., Friswell, M. K., Alswied, A., Roberts, C. A., Wells, J. W., Flanagan, P. G., Knight, P. R., Codling, C. E., Marchesi, J. R., Winstanley, C., Hall, N., Rhodes, J. M., & Campbell, B. J. (2014). Colonic mucosa-associated diffusely adherentafaC+ EscherichiacoliexpressinglpfAandpksare increased in inflammatory bowel disease and colon cancer. Gut, 63(5), 761–770. https://doi.org/10.1136/gutjnl-2013-304739

Rifaie, S., Patil, V. S., & Jangid, K. (2022). Advances in sequencing technology, databases, and analyses tools for the assessment of microbial diversity. In *Elsevier eBooks* (pp. 317–347). https://doi.org/10.1016/b978-0-323-90148-2.00003-1

Rutter, M. D., Saunders, B. P., Wilkinson, K. H., Rumbles, S., Schofield, G., Kamm, M. A., Williams, C., Price, A. B., Talbot, I. C., & Forbes, A. (2006). Thirty-Year Analysis of a Colonoscopic Surveillance Program for Neoplasia in Ulcerative Colitis. Gastroenterology, 130(4), 1030–1038. https://doi.org/10.1053/j.gastro.2005.12.035

Salesse, L., Lucas, C., Hoang, M. H. T., Sauvanet, P., Rezard, A., Rosenstiel, P., Damon-Soubeyrand, C., Barnich, N., Godfraind, C., Dalmasso, G., & Nguyen, H. T. T. (2021). Colibactin-Producing Escherichia coli Induce the Formation of Invasive Carcinomas in a Chronic Inflammation-Associated Mouse Model. *Cancers*, *13*(9), 2060. https://doi.org/10.3390/cancers13092060

Sarshar, M., Scribano, D., Marazzato, M., Ambrosi, C., Aprea, M. C., Aleandri, M., Pronio, A., Longhi, C., Nicoletti, M., Zagaglia, C., Palamara, A. T., & Conte, M. R. (2017b). Genetic diversity, phylogroup distribution and virulence gene profile of pks positive Escherichia coli colonizing human intestinal polyps. Microbial Pathogenesis, 112, 274–278. https://doi.org/10.1016/j.micpath.2017.10.009

Schultz, M. (2008). Clinical use of E. coli Nissle 1917 in inflammatory bowel disease. Inflammatory Bowel Diseases, 14(7), 1012–1018. https://doi.org/10.1002/ibd.20377

Sears, C. L., & Pardoll, D. M. (2011). Perspective: Alpha-Bugs, Their Microbial Partners, and the Link to Colon Cancer. The Journal of Infectious Diseases, 203(3), 306–311. https://doi.org/10.1093/jinfdis/jiq061

Shimpoh, T., Hirata, Y., Ihara, S., Suzuki, N., Kinoshita, H., Hayakawa, Y., Ota, Y., Narita, A., Yoshida, S., Yamada, A., & Koike, K. (2017). Prevalence of pks-positive Escherichia coli in Japanese patients with or without colorectal cancer. *Gut Pathogens*, *9*(1). https://doi.org/10.1186/s13099-017-0185-x

Siegel, R. L., Wagle, N. S., Cercek, A., Smith, R. A., & Jemal, A. (2023). Colorectal cancer statistics, 2023. CA: A Cancer Journal for Clinicians, 73(3), 233–254. https://doi.org/10.3322/caac.21772

Silpe, J. E., Wong, J., Owen, S. V., Baym, M. H., & Balskus, E. P. (2022). The bacterial toxin colibactin triggers prophage induction. *Nature*, *603*(7900), 315–320. https://doi.org/10.1038/s41586-022-04444-3