



**HAL**  
open science

## Virulence Factors in Colorectal Cancer Metagenomes and Association of Microbial Siderophores with Advanced Stages

Nour El Houda Mathlouthi, Aicha Kriaa, Leila Ammar Keskes, Moez Rhimi,  
Radhouane Gdoura

► **To cite this version:**

Nour El Houda Mathlouthi, Aicha Kriaa, Leila Ammar Keskes, Moez Rhimi, Radhouane Gdoura. Virulence Factors in Colorectal Cancer Metagenomes and Association of Microbial Siderophores with Advanced Stages. *Microorganisms*, 2022, 10 (12), pp.2365. 10.3390/microorganisms10122365 . hal-04090485

**HAL Id: hal-04090485**

**<https://hal.inrae.fr/hal-04090485>**

Submitted on 5 May 2023

**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 4.0 International License



## Article

# Virulence Factors in Colorectal Cancer Metagenomes and Association of Microbial Siderophores with Advanced Stages

Nour El Houda Mathlouthi <sup>1</sup>, Aicha Kriaa <sup>2</sup>, Leila Ammar Keskes <sup>3</sup>, Moez Rhimi <sup>2</sup> and Radhouane Gdoura <sup>1,\*</sup>

<sup>1</sup> Laboratoire de Recherche Toxicologie Microbiologie Environnementale et Santé (LR17ES06), Faculté des Sciences de Sfax, Université de Sfax, Sfax 3000, Tunisia

<sup>2</sup> Microbiota Interaction with Human and Animal Team (MIHA), Micalis Institute-UMR1319, Ag-roParisTech, Université Paris-Saclay, INRAE, F-78350 Jouy-en-Josas, France

<sup>3</sup> Laboratoire de Recherche de Génétique Moléculaire Humaine, Faculté de Médecine de Sfax, Avenue Majida BOULILA, Université de Sfax, Sfax 3029, Tunisia

\* Correspondence: radhouane.gdoura@fss.usf.tn

**Abstract:** Colorectal cancer (CRC) is a growing public health challenge, featuring a multifactorial etiology and complex host–environment interactions. Recently, increasing evidence has pointed to the role of the gut microbiota in CRC development and progression. To explore the role of gut microbes in CRC, we retrieved metagenomic data from 156 stools from the European Nucleotide Archive database and mapped them against the VFDB database for virulence factors (VFs). GO annotations of VFs and KEGG pathways were then performed to predict the microbial functions and define functional pathways enriched in the tumor-associated microbiota. Interestingly, 306 VFs were detected in the metagenomic data. We revealed the enrichment of adenomas with VFs involved in cell adhesion, whereas in the early stages of CRC they were enriched in both adhesins and isochorismatase. Advanced stages of CRC were enriched with microbial siderophores, especially enterobactin, which was significantly associated with isochorismate synthase. We highlighted higher abundances of porins and transporters involved in antibiotic resistance and the development of biofilm in advanced stages of CRC. Most VFs detected in CRC, particularly in advanced stages, were shown to be included in siderophore biosynthesis pathways. This enrichment of predicted VFs supports the key role of the gut microbiota in the disease.

**Keywords:** colorectal cancer; gut microbiota; pathogens; virulence factors; siderophores; prognosis



**Citation:** Mathlouthi, N.E.H.; Kriaa, A.; Keskes, L.A.; Rhimi, M.; Gdoura, R. Virulence Factors in Colorectal Cancer Metagenomes and Association of Microbial Siderophores with Advanced Stages. *Microorganisms* **2022**, *10*, 2365. <https://doi.org/10.3390/microorganisms10122365>

Academic Editors: Peng Chen, Xingyin Liu and Wenke Feng

Received: 2 October 2022

Accepted: 26 October 2022

Published: 30 November 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

In 2020, the International Agency for Research on Cancer recorded 935,173 deaths caused by colorectal cancer (CRC), making it the third most common cancer [1]. The etiology of CRC involves both genetic and environmental factors [2]. Inherited cancers account for only 5% of all CRC cases and could be classified into polyposis (familial adenomatous polyposis (FAP)) and hereditary nonpolyposis colorectal cancer (HNPCC) [3]. Unlike these cancers which are driven by inherited mutations [4], sporadic cancers stemming from point mutations in the lifetime, target several loci linked to the initiation of the adenoma formation and the development and progression of carcinoma [5]. In addition, personal medical history of inflammatory bowel disease (IBD) could promote carcinogenesis [6]. Another important risk factor was strongly associated with CRC is lifestyle, including dietary patterns (a diet high in meat and low in fiber [7]), obesity [8], physical inactivity [9], and cigarette smoking [10].

The role of the gut microbiota in CRC is increasingly recognized [11]. Accordingly, we have recently witnessed the emergence of the holobiont concept to understand the host-microbiota interplay and its relevance in the disease [12]. Next-generation sequencing technologies have advanced our knowledge of the human microbiota as they have enabled the characterization of noncultivable microbes, among others, in health and disease and

the prediction of their function [13]. Bioinformatic analysis presents one of the most important steps in metagenomic research; it translates a large amount of biological data into interpretable results. One of the many crucial analyses that we can cite is profiling the taxonomic composition of microbial communities and mapping it against a database such as the Virulence Factor Database (VFDB), which is a reference database for virulence factors produced by bacterial pathogens [14].

Indeed, pathogenic bacteria can produce virulence factors (VFs) that could affect different signaling pathways and induce colorectal carcinogenesis [15]. VFs could be categorized into several classes based on their functions: adherence and colonization factors, invasion factors, capsules and other surface components, and endotoxins [16]. It has been previously shown that the virulence factor of enterotoxigenic *Bacteroides fragilis* (ETBF) and *Fusobacterium adhesin A* (*FadA*) can bind to intestinal epithelial cell receptors and cleave the extracellular domain of E-cadherin (anchoring junction protein), causing a loss of cell–cell contact and leading to the activation of the NF- $\kappa$ B pathway [17,18]. Other virulence factors such as colibactin and cytolethal distending toxins can trigger host cell double-strand DNA breaks. These factors could alter the immune response of the host and induce the production of pro-inflammatory cytokines [19]. The cytotoxic necrotizing factor 1 (CNF1) produced by *Escherichia coli* leads to cellular motility [20] and enhances proliferation due to the perturbation of the host cellular cycle [21].

In the present study, we aim to uncover the association of virulence factors produced by microbial pathogens in CRC patients at different stages in comparison to healthy subjects by analyzing published metagenomic data generated by whole-genome sequencing from stools of healthy subjects and CRC patients. Next, we will perform a functional profiling of the virulence factors of interest as well as pathway analysis. The overall objective is not only to understand the role of pathogens in colorectal carcinogenesis but also to reveal the importance of the virulence factors in CRC.

## 2. Materials and Methods

### 2.1. Published Fecal Metagenomes

We retrieved gut metagenomic data of the project PRJEB6070 from the European nucleotide archive. This project involves 156 fecal samples from the French population, including 53 from CRC patients (15 from stage I, 7 from stage II, 10 from stage III, and 21 from stage IV), 42 samples from adenoma patients (15 from large adenoma and 27 from small adenoma), 61 samples from randomly selected controls, and 43 samples from the German population (38 fecal samples from CRC patients and 5 samples from healthy subjects). These 199 samples were sequenced using two different techniques: whole-genome shotgun and 16S rRNA gene sequencing. We only analyzed the whole metagenomic data from 156 French samples [22]. (Supplementary Data S1). Datasets of published fecal metagenomes are available from the European Nucleotide Archive (ENA) database (<https://www.ebi.ac.uk/ena/browser/home>, accessed on 25 October 2022), accession number: ERP005534.

### 2.2. Detection of Virulence Factors (VFs)

In order to detect the virulence factors of bacterial pathogens, we mapped the reads against the virulence factor database (VFDB) [14]. The mapping was realized using tools from PATRIC website (<https://patricbrc.org/app/MetagenomicReadMapping>) accessed from 9 December 2021 to 6 January 2022 [23,24].

### 2.3. Statistical Analysis

The Chi-square test and Spearman correlation coefficient were calculated to estimate the correlations between cancer progression and the detected VFs.

Multivariate analysis was performed to investigate the linear relationship between Isochorismate synthase as a dependent variable and type 1 Fimbriae regulatory protein (FimB), type 1 Fimbriae regulatory protein FimE, type 1 Fimbrial protein A chain pre-

cursor, iron-enterobactin ABC transporter permease, phosphopantetheinyl transferase component of enterobactin synthase multienzyme complex, FimG protein precursor, ferrienterobactin ABC transporter periplasmic binding protein, *E. coli* common pilus usher EcpC, enterobactin/ferric enterobactin, esterase, enterobactin exporter (iron-regulated), general secretion pathway protein D, general secretion pathway protein K, and aerobactin synthesis protein IucB as covariates.

Statistical analyses were performed using statistical software package SPSS 20.0. (SPSS 20.0 for Windows; SPSS Inc., Chicago, IL, USA). A two-tailed  $p$  value below 0.05 was considered statistically significant.

#### 2.4. Functional Profiling of VFs, Pathway Analysis, and Association of VF's Functions with Cancer Progression

Correlated VFs with cancer progression were mapped against the Uniprot–Gene Ontology (GO) database ([https://www.uniprot.org/help/gene\\_ontology](https://www.uniprot.org/help/gene_ontology)) accessed from 9 December 2021 to 6 January 2022, and the functional annotation pathway mapping of VFs was realized using KEGG pathway (<https://www.kegg.jp/kegg/pathway.html>) they were accessed from accessed from 20 December 2021 to 10 January 2022. A Venn diagram was used to display the most detected functions of VFs correlated with each CRC stage compared to healthy state (<http://bioinformatics.psb.ugent.be/webtools/Venn/>) accessed on 25 January 2022.

### 3. Results

#### 3.1. Published Fecal Metagenomes

By mapping the reads against the Virulence Factor Database (VFDB), we detected the presence of 306 VFs (Supplementary Data S2). These factors are produced by pathogenic microbial species that belong to different phyla: Firmicutes (*Clostridium perfringens*, *Enterococcus faecalis*, *Enterococcus faecium*, *Streptococcus agalactiae* and *Streptococcus pneumoniae*), Actinobacteria (*Corynebacterium diphtheriae*), and Proteobacteria (*Escherichia coli*, *Haemophilus influenzae*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa*, *Salmonella enterica*, *Shigella dysenteriae*, *Shigella flexneri* and *Yersinia pestis*).

#### 3.2. VFs as Markers of Cancer Progression: Bivariate Analysis

Correlation analysis was performed to explore the virulome of different CRC stages and investigate the link between the VFs from pathogenic features and the disease stage (see Supplementary Data S3).

The comparison of different stages of CRC with the normal state revealed an enrichment of *K. pneumoniae* VFs in small adenoma, and *E. coli* VFs and *K. pneumoniae* in large adenoma ( $p < 0.05$ ). *H. influenzae* VFs and *K. pneumoniae* VFs were more abundant in the early stages of CRC. The advanced stages of CRC were found to be more enriched with *E. coli*, *K. pneumoniae*, *S. flexneri* and *Y. pestis* VFs (Supplementary Data S3). The comparison of VF levels at different stages did not reveal any significant correlation between the virulomes of small adenoma vs. large adenoma and large adenoma vs. early-stage CRC. On the contrary, a comparison of adenoma (large and small) with early-stage CRC revealed a significant enrichment with *K. pneumoniae* VFs, *E. coli* VFs and *E. faecalis* VFs in early-stage CRC. We also compared early-stage and advanced-stage CRC and we noticed that stages III and IV were more enriched with *E. coli* VFs, *K. pneumoniae* VFs especially colibactin, *Y. pestis* VFs and *E. faecalis* VF (Supplementary Data S3).

#### 3.3. Multivariate Analysis

Multivariate analysis revealed that isochorismate synthase 1 was positively correlated with iron-enterobactin ABC transporter permease ( $p < 0.001$ ), FimG protein precursor ( $p = 0.030$ ), *E. coli* common pilus usher EcpC ( $p < 0.001$ ), enterobactin exporter (iron-regulated) ( $p < 0.001$ ), and general secretion pathway protein D ( $p = 0.030$ ), and negatively correlated with type-1 Fimbrial protein, A chain precursor ( $p = 0.014$ ), the phos-

phosphopantetheinyl transferase component of enterobactin synthase multienzyme complex ( $p = 0.036$ ), and general secretion pathway protein K ( $p = 0.002$ ) (Table 1).

**Table 1.** Multivariate analysis of virulence factors detected in advanced stages of CRC.

Model	Coefficients <sup>a</sup>			T	Sig.
	Unstandardized Coefficients		Standardized Coefficients		
	B	Std. Error	Beta		
(Constant)	−0.005	0.008		−0.550	0.584
Type 1 Fimbriae Regulatory protein FimB	0.014	0.057	0.014	0.239	0.812
Type 1 Fimbriae Regulatory protein FimE	0.017	0.040	0.017	0.425	0.672
Type-1 Fimbrial protein, A chain precursor	−0.081	0.032	−0.080	−2.510	0.014
iron-enterobactin ABC transporter permease	0.501	0.053	0.505	9.498	0.000
phosphopantetheinyl transferase component of enterobactin synthase multienzyme complex	−0.072	0.034	−0.075	−2.139	0.036
FimG protein precursor	0.067	0.030	0.069	2.211	0.030
ferrienterobactin ABC transporter periplasmic binding protein	0.041	0.038	0.042	1.092	0.278
<i>E. coli</i> common pilus usher EcpC	0.354	0.060	0.351	5.881	0.000
enterobactin/ferric enterobactin esterase	−0.038	0.081	−0.039	−0.468	0.641
enterobactin exporter, iron-regulated	0.213	0.054	0.216	3.910	0.000
general secretion pathway protein D	0.078	0.035	0.074	2.207	0.030
general secretion pathway protein K	−0.111	0.035	−0.109	−3.146	0.002
aerobactin synthesis protein <i>LucB</i>	0.020	0.022	0.015	0.879	0.382

<sup>a</sup>. Dependent Variable: isochorismate synthase 1.

### 3.4. Functional Annotation of VFs

We found that only VFs involved in cell adhesion were enriched in small adenomas. Large adenoma was enriched with VFs involved in DNA integration, DNA recombination, the cellular response to DNA damage stimulus, cell adhesion, regulator activity, biosynthetic processes, and metabolic processes. As for early-stage and advanced-CRC, the analysis revealed that the fecal microbiome contains VFs that seem to be a part of different processes and activities (Table 2).

To identify relevant functions in CRC stages, we used a Venn diagram (Figure 1). VFs involved in cell adhesion were detected in all stages of CRC, and those involved in the biosynthetic and metabolic processes were found in large adenoma and all stages of CRC. Large adenoma and advanced-stage CRC were enriched in VFs involved in DNA integration, DNA recombination, the cellular response to DNA damage stimulus, and regulator activity. All CRC samples exhibited VFs that play a role in the binding, signal transduction, the regulation of transcription, porin activity, and transportation. Pilus assembly was the only function included exclusively in early-stage CRC. Ligase activity, the NAD metabolic process, protein phosphopantetheinylation, biofilm development, host cell recognition, mechanosensory behavior, the homeostasis process, the detection of virus, viral entry into host cell, response to antibiotics, oxidoreductase activity, and signaling receptor activity were mostly associated with VFs found in advanced-stage CRC.

Functions of targeted VFs detected:

- Cell adhesion in large adenoma, small adenoma, CRC I-II, and CRC III-IV;
- Metabolic process and biosynthetic process in large adenoma, CRC I-II, and CRC III-IV;
- DNA recombination, DNA integration, regulator activity, and cellular response to DNA damage stimulus in large adenoma and CRC III-IV;
- Binding activity, the regulation of transcription, porin activity, transporters, and signal transduction in CRC I-II and CRC III-IV;
- Pilus assembly in CRC I-II;
- Response to an antibiotics, biofilm development, NAD metabolic process, viral entry into the host cell, homeostasis process, oxidoreductase activity, detection of virus,

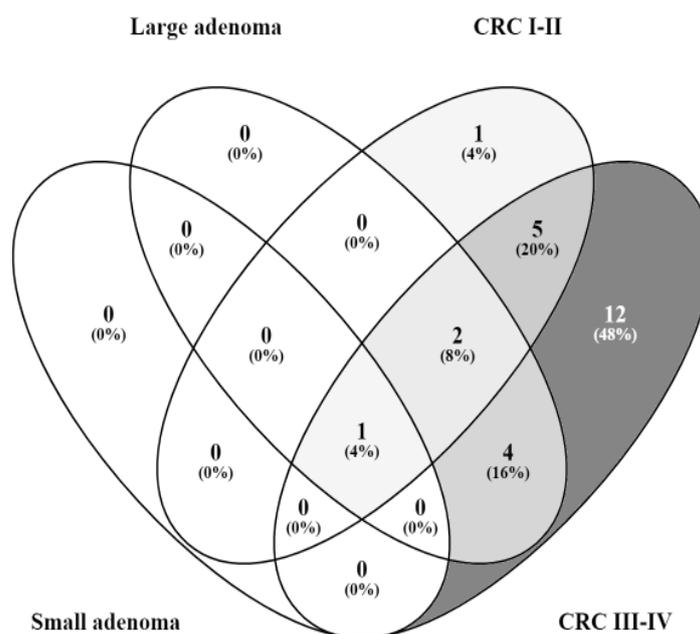
signaling receptor activity, mechanosensory behavior, ligase activity, protein phosphopantetheinylation, and host cell recognition only in CRC III-IV.

**Table 2.** Functions and pathway analysis of the VFs associated with different stages of CRC.

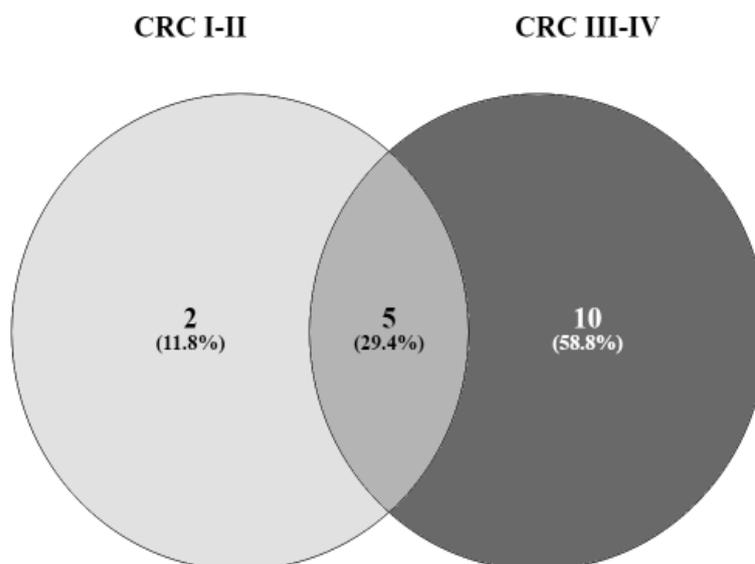
	Early Stage CRC	Advance Stage CRC
Functions of Virulence factors	<ul style="list-style-type: none"> <li>• Biosynthetic Process</li> <li>• Metabolic Process</li> <li>• Binding Activity</li> <li>• Signal Transduction</li> <li>• Regulation of Transcription</li> <li>• Cell Adhesion</li> <li>• Porin Activity</li> <li>• Pilus Assembly</li> <li>• Transportation</li> </ul>	<ul style="list-style-type: none"> <li>• Metabolic process</li> <li>• Ligase activity</li> <li>• Binding activity</li> <li>• Biosynthetic process</li> <li>• Dna integration</li> <li>• Dna recombination</li> <li>• Cellular response to dna damage stimulus</li> <li>• Cell adhesion</li> <li>• Nad metabolic process</li> <li>• Signal transduction</li> <li>• Regulator activity</li> <li>• Protein phosphopantetheinylation</li> <li>• Biofilm development</li> <li>• Host cell recognition</li> <li>• Mechanosensory behavior</li> <li>• Transportation</li> <li>• Hemeostasis process</li> <li>• Porin activity</li> <li>• Detection of virus</li> <li>• Viral entry into the host cell</li> <li>• Response to an antibiotic oxidoreductase activity</li> <li>• Regulation of transcription</li> <li>• Signaling receptor activity</li> </ul>
Pathway of virulence factor	<ul style="list-style-type: none"> <li>• Streptomycin biosynthesis</li> <li>• O-antigen nucleotide sugar biosynthesis</li> <li>• Metabolic pathways</li> <li>• Biosynthesis of secondary metabolites</li> <li>• Biosynthesis of nucleotide sugars</li> <li>• Biosynthesis of siderophore group</li> <li>• Nonribosomal peptides</li> <li>• Biosynthesis of secondary metabolites two-component system</li> </ul>	<ul style="list-style-type: none"> <li>• Biosynthesis of secondary metabolites</li> <li>• Biosynthesis of siderophore group nonribosomal peptides</li> <li>• Abc transporters</li> <li>• Ubiquinone and terpenoid-quinone biosynthesis</li> <li>• Metabolic pathway</li> <li>• Bacterial secretion system</li> <li>• Pentose and glucuronate interconversions ascorbate and aldarate metabolism</li> <li>• Amino sugar and nucleotide sugar metabolism</li> <li>• o-antigen nucleotide sugar biosynthesis</li> <li>• biosynthesis of cofactors</li> <li>• Biosynthesis of nucleotide sugars</li> <li>• Pentose and glucuronate interconversions galactose metabolism</li> <li>• Starch and sucrose metabolism</li> </ul>

### 3.5. Pathway annotation of VFs

The pathway analysis was conducted on the VFs considered to be associated with each stage of CRC (Figure 2). The predominantly detected VFs in small adenoma and large adenoma vs. normal ( $p < 0.05$ ) were not included in KEGG pathway analysis. We found that VFs associated with advanced- and early-stages CRC were involved in many pathways (Table 2).



**Figure 1.** Venn diagram showing the number of shared and unique functions of targeted VFes detected in different stages of CRC.



**Figure 2.** Venn diagram showing the number of shared and unique pathways that include targeted VFes detected at different stages of CRC.

A Venn diagram was used at this level to explore the relevant pathways in CRC stages. For early CRC, streptomycin biosynthesis and a two-component system were mostly revealed. As for advanced CRC, many pathways were identified, including ABC transporters, ubiquinone and terpenoid-quinone biosynthesis, a bacterial secretion system, pentose and glucuronate interconversions, ascorbate and aldarate metabolism, amino sugar and nucleotide sugar metabolism, the biosynthesis of cofactors, galactose metabolism, starch and sucrose metabolism, and microbial metabolism in diverse environments.

- Pathways that include targeted VFes detected in all CRC stages: O-antigen nucleotide sugar biosynthesis, metabolic pathways, biosynthesis of secondary metabolites, biosynthesis of nucleotide sugars, biosynthesis of siderophore group nonribosomal peptides;
- Pathways that include targeted VFes detected in CRC I-II: streptomycin biosynthesis and two-component system;

- Pathways that include targeted VFs detected in CRC III-IV: ABC transporters, ubiquinone and terpenoid-quinone biosynthesis, a bacterial secretion system, pentose and glucuronate interconversions, ascorbate and aldarate metabolism, amino sugar and nucleotide sugar metabolism, the biosynthesis of cofactors, galactose metabolism, starch and sucrose metabolism, and microbial metabolism in diverse environments.

#### 4. Discussion

Several studies have highlighted the role pathogenic bacteria and their virulence factors in colorectal carcinogenesis [11,15,22]. Defining a clear set of virulence factors as potential targets of interest and investigating the involved pathways is a promising approach for CRC therapy. The analysis of this association requires cohort studies and comparisons of the results in a large-scale meta-analysis [22]. In the present study, we analyzed shotgun-metagenomic data available on ENA, generated from the sequencing of 156 samples in order to explore the virulence factors present in the intestinal microbiota of patients with CRC. Only Fimbriae virulence factors produced by *K. pneumoniae* mrkD and mrkF were found to be associated with small adenoma. Previous studies revealed their involvement in the surface stability of Fimbriae and facilitation of binding to collagen, respectively [25]. Similarly, large adenoma seemed to be significantly associated with Fimbriae virulence factors produced not only by *K. pneumoniae* (mrkF) but also by *E. coli* (FimE). LPS virulence factor (UTP-glucose-1-phosphate uridylyltransferase) was correlated with large adenoma and it was previously confirmed to be significantly increased after hypoxia, which is a hallmark of the tumor microenvironment [26].

Early-stage CRC was associated with different Fimbriae (mrkD, mrkF, and mrkC) and with LPS (dTDP-glucose 46-dehydratase). At this stage, a significant association was revealed between early-stages CRC and RcsB. This transcriptional regular was earlier related to biofilm formation and cell division [27]. Moreover, we noticed a correlation between stages I and II of CRC and EntB (siderophore). Different types of virulence factors have been found in advanced CRC. In fact, several Fimbriae were detected such as FimE, FimB, FimG, erpC, and FimA, which have been shown to enhance the ability of uropathogenic *E. coli* (UPEC) to elicit symptomatic infection [28]. LPS of the outer membrane A and UTP-glucose-1-phosphate uridylyltransferase subunit GalF were also highly detected in these stages. We noticed a significant association between stages III and IV, and hasB, which is involved in capsular polysaccharide biosynthesis and plays an important role in the formation of hyaluronic acid (HA), thus promoting the cellular growth and migration of physiological cells [29]. Additionally, advanced CRC has been correlated with the immunomodulatory protein TcpC, known to counter the host's innate immune defense by abrogating the function of MyD88 in macrophages that further impairs the involved signaling processes [30]. We also noticed that many types of secretion systems could be involved in the advanced level of adenoma–carcinoma sequence: the type 3 secretion ESPL1 and ESPX1 (type III secretion systems are used by pathogenic species to inject toxins into attacking immune cells [31,32]), and general secretion system D, F, K (general secretion pathway can translocate the small-signal sequence-containing proteins into the periplasm which can be targeted then to ABC transporter/TolC-based export system such as the enterotoxins produced by *E. coli* [33]). We also observed that siderophores were the most detected VFs in early-stage CRC, which is in line with a previous study highlighting the enrichment of entC in CRC [34].

Enterobactin (EntF, FepC, FepG, EntD, EntE, FepB, FepD, Fes, EntS), which was highly detected in advanced-stage CRC stool, has been previously studied and shown to play a role in promoting a pro-inflammatory response, as well as reducing basal reactive oxygen species (ROS) generation in Intestinal epithelial cells [35]. Similarly, yersiniabactin (ybtE and fyuA), which is a major bacterial iron-uptake system, has been shown to promote oxidative stress [36] by triggering three signaling pathways that have been reported to be closely linked to CRC development [37]: the Janus kinases (JAK)/signal transducer and activator

of transcription proteins (STAT), Wnt/ $\beta$ -catenin, and PI3K/AKT pathways. However, no previous study has confirmed an association between aerobactin (*iucB*) and CRC.

While the comparison of the different CRC stages with healthy stools was for prognostic purposes, we compared the different stages with each other to study the variation and the evolution of the virulome throughout the adenoma–carcinoma sequence. No significant differences were revealed between the VFs in small adenoma vs. large adenoma and large adenoma vs. early CRC. However, when we compared adenomas (large and small) with early-stage CRC, we noticed a significant correlation with VFs, such as Fimbriae (*SfaG*, *SfaS*), siderophore (*EntB*), transcriptional regulator (*rccb*), immunomodulatory protein (*TcpC*), a type VI secretion system (which plays an important role in delivering effectors into cells using a bacteriophage-like puncture mechanism) [38], CNF1 (toxins induce the activation of Rho GTPases that triggers different cellular events such as NF- $\kappa$ B activation (via the Akt/I $\kappa$ B kinase pathway) and provide protection against apoptosis) [39], and polysaccharide lyase produced by *E. faecalis*, involved in malignant transformation, invasion and metastasis [40].

We also compared the virulomes of early- and advanced-stage CRC, and found that stages III and IV were more enriched with Fimbriae (*FimB*, *FimE*, *fumH*, *FimI* and *FimG*), transcriptional regulator (*Ybta*), outer-membrane A (*Lps*) and siderophores (aerobactin (*iucb*), enterobactin (*fepC* and *fepG*) and yersiniabactin (*Ybtp*, *Irp1*, and *Irp2*).

A significant association between colibactin (*clbK*, *clbF*, *clbI*, and *clbL*) and advanced-stage CRC was observed. Colibactin is a genotoxic metabolite that causes DNA double-strand breaks in mammalian cells, leading to chromosomal abnormalities, cell cycle arrest, and senescence [41]. It has been demonstrated that colibactin-producing bacteria can promote the emergence of senescent cells by regulating the tumor microenvironment, which can also influence tumor progression and cancer progression through the secretion of growth factors [42]. In addition, *E. faecalis* produces a fibrinogen-binding protein (*FSS2*) at this level involved in the bacteria–host interaction and which is necessary in the early stages of infection [43].

The VFs involved in cell adhesion were detected in all stages of adenoma and carcinoma; however, binding activity, signal transduction, transcription regulation, and the activity of porins and transporters involved in antibiotic resistance were not found in the adenoma. Chuanfa et al. showed that the gut microbiota of CRC patients had higher antibiotic-resistance genes than those of healthy subjects [44]. The VFs associated with advanced-stage CRC showed a unique functional profile characterized by oxidoreductase activity and Nad metabolism. In addition, biofilm development, previously shown to improve epithelial permeability, promotes the inflammation of pro-carcinogenic tissues via the direct access of bacterial antigens/mutagens to an unprotected epithelial surface [45]. As for pathway analysis, many VFs correlated with advanced-stage CRC are involved in the biosynthesis of nonribosomal siderophore group peptides that could drive colitic and pro-carcinogenic responses [46].

Here, we investigated the network pathogenicity of VFs associated with advanced-stage CRC via multivariate analysis, and noticed that isochorismate synthase was highly correlated with enterobactin exporter. This link was previously described, and isochorismate was considered as a precursor for the biosynthesis of the siderophore enterobactin [47]. In fact, pathogenic bacteria producing isochorismate can uptake iron from other living organisms even when iron concentrations are extremely low [48]. Since iron is required to maintain immune functions, its deficiency may increase oncogenicity by altering neoplastic changes in immunosurveillance and potentially remodeling the tumor immune microenvironment [49]. This could trigger other pathways involved in cell cycle arrest and apoptosis [50].

## 5. Conclusions

In this study, we analyzed the virulome using metagenomic data available in ENA and identified a significant enrichment of predicted VFs in CRC patients' microbiomes. These

VFs were produced by pathogens such as pathogenic *E. coli*, *K. pneumoniae*, *Enterococcus sp.*, *Shigella sp.*, etc.

Colibactin and siderophores were remarkably more abundant in advanced stages of CRC compared to healthy microbiomes and early stages of CRC. Using this combined approach, our work offers a holistic exploration of the predicted pathogenic potential of the microbiome and reveals specific tax/functions as potential targets in this disease. Further functional analyses are needed to understand the role of VFs in oncogenicity.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms10122365/s1>. Supplementary Data S1: Participant characterization; Supplementary Data S2: VFDB\_virulence factor; Supplementary Data S3: GO annotation.

**Author Contributions:** Conceptualization: N.E.H.M., A.K., L.A.K., M.R. and R.G.; methodology: N.E.H.M. and R.G.; investigation: N.E.H.M., A.K., M.R., L.A.K. and R.G.; data curation: N.E.H.M. and R.G.; writing—original draft preparation: N.E.H.M.; writing—review and editing: A.K., M.R., L.A.K. and R.G.; visualization: M.R., LK and R.G.; supervision: R.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work received funding from the Twinning European project: 952583—MICAfrica—H2020-WIDESPREAD-2018-2020/H2020-WIDESPREAD-2020-5.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the analyses or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

- International Agency for Research on Cancer. Colorectal Cancer. 2020. Available online: [https://gco.iarc.fr/today/data/factsheets/cancers/10\\_8\\_9-Colorectum-fact-sheet.pdf](https://gco.iarc.fr/today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf) (accessed on 1 February 2022).
- Louis, P.; Hold, G.L.; Flint, H.J. The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* **2014**, *12*, 661–672. [[CrossRef](#)] [[PubMed](#)]
- Mármol, I.; Sánchez-De-Diego, C.; Pradilla Dieste, A.; Cerrada, E.; Rodriguez Yoldi, M. Colorectal carcinoma: A general overview and future perspectives in colorectal cancer. *Int. J. Mol. Sci.* **2017**, *1*, 197. [[CrossRef](#)] [[PubMed](#)]
- Al-Sohaily, S.; Biankin, A.; Leong, R.; Kohonen-Corish, M.; Warusavitarne, J. Molecular pathways in colorectal cancer. *J. Gastroenterol. Hepatol.* **2012**, *27*, 1423–1431. [[CrossRef](#)] [[PubMed](#)]
- Kheirleisid, E.A.; Miller, N.; Kerin, M. Molecular biology of colorectal cancer: Review of the literature. *Am. J. Mol. Biol.* **2013**, *3*, 30688. [[CrossRef](#)]
- Amersi, F.; Agustín, M.; Ko, C.Y. Colorectal cancer: Epidemiology, risk factors, and health services. *Clin. Colon Rectal Surg.* **2005**, *18*, 133–140. [[CrossRef](#)]
- Giovannucci, E.; Rimm, E.B.; Stampfer, M.J.; Colditz, G.; Ascherio, A.; Willett, W.C. Intake of fat, meat, and fiber in relation to risk of colon cancer in men. *Cancer Res.* **1994**, *54*, 2390–2397.
- Jochem, C.; Leitzmann, M. Obesity and colorectal cancer. *Obes. Cancer* **2016**, *208*, 17–41. [[CrossRef](#)]
- Tayyem, R.F.; Shehadeh, I.N.; AbuMweis, S.S.; Bawadi, H.A.; Hammad, S.S.; Bani-Hani, K.E.; Al-Jaberi, T.M.; Alnusair, M.M. Physical inactivity, water intake and constipation as risk factors for colorectal cancer among adults in Jordan. *APJCP* **2013**, *14*, 5207–5212. [[CrossRef](#)]
- Liang, P.S.; Chen, T.Y.; Giovannucci, E. Cigarette smoking and colorectal cancer incidence and mortality: Systematic review and meta-analysis. *IJC* **2009**, *124*, 2406–2415. [[CrossRef](#)]
- Cheng, Y.; Ling, Z.; Li, L. The intestinal microbiota and colorectal cancer. *Front. Immunol.* **2020**, *11*, 615056. [[CrossRef](#)]
- Margulis, L. Words as battle cries: Symbiogenesis and the new field of endocytobiology. *Bioscience* **1990**, *40*, 673–677. [[CrossRef](#)]
- Cason, C.; D’Accolti, M.; Soffritti, I.; Mazzacane, S.; Comar, M.; Caselli, E. NGS and PCR technologies in monitoring the hospital microbiome and its drug resistance. *Front. Microbiol.* **2022**, *13*, 969863. [[CrossRef](#)] [[PubMed](#)]
- Chen, L.; Yang, J.; Yu, J.; Yao, Z.; Sun, L.; Shen, Y.; Jin, Q. VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* **2005**, *33* (Suppl. S1), D325–D328. [[CrossRef](#)] [[PubMed](#)]
- Sun, J. Impact of bacterial infection and intestinal microbiome on colorectal cancer development. *Chin. Med. J.* **2022**, *135*, 400–408. [[CrossRef](#)] [[PubMed](#)]
- Baron, S. *Medical Microbiology*; University of Texas Medical Branch at Galveston: Galveston, TX, USA, 1996.

17. Rubinstein, M.R.; Wang, X.; Liu, W.; Hao, Y.; Cai, G.; Han, Y.W. Fusobacterium nucleatum promotes colorectal carcinogenesis by modulating E-cadherin/ $\beta$ -catenin signaling via its FadA adhesin. *Cell Host Microbe* **2013**, *14*, 195–206. [[CrossRef](#)] [[PubMed](#)]
18. Wu, S.; Shin, J.; Zhang, G.; Cohen, M.; Franco, A.; Sears, C.L. The Bacteroides fragilis toxin binds to a specific intestinal epithelial cell receptor. *Infect. Immun.* **2006**, *74*, 5382–5390. [[CrossRef](#)]
19. Fiorentini, C.; Carlini, F.; Germinario, E.A.P.; Maroccia, Z.; Travaglione, S.; Fabbri, A. Gut microbiota and Colon Cancer: A role for bacterial protein toxins? *Int. J. Mol. Sci.* **2020**, *21*, 6201. [[CrossRef](#)]
20. Doye, A.; Mettouchi, A.; Bossis, G.; Clément, R.; Buisson-Touati, C.; Flatau, G.; Gagnoux, L.; Piechaczyk, M.; Boquet, P.; Lemichez, E. CNF1 exploits the ubiquitin-proteasome machinery to restrict Rho GTPase activation for bacterial host cell invasion. *Cell* **2002**, *111*, 553–564. [[CrossRef](#)]
21. El-Aouar Filho, R.A.; Nicolas, A.; Castro, T.L.d.; Deplanche, M.; Azevedo, V.A.d.; Goossens, P.L.; Taieb, F.; Lina, G.; le Loir, Y.; Berkova, N. Heterogeneous family of cyclomodulins: Smart weapons that allow bacteria to hijack the eukaryotic cell cycle and promote infections. *Front. Cell. Infect.* **2017**, *7*, 208. [[CrossRef](#)]
22. Zeller, G.; Tap, J.; Voigt, A.Y.; Sunagawa, S.; Kultima, J.R.; Costea, P.I.; Amiot, A.; Böhm, J.; Brunetti, F.; Habermann, N.; et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol. Syst. Biol.* **2014**, *10*, 766. [[CrossRef](#)]
23. Davis, J.J.; Wattam, A.R.; Aziz, R.K.; Brettin, T.; Butler, R.; Butler, R.M.; Stevens, R. The PATRIC Bioinformatics Resource Center: Expanding data and analysis capabilities. *Nucleic Acids Res.* **2020**, *48*, D606–D612. [[CrossRef](#)] [[PubMed](#)]
24. Yan, J.; Liao, C.; Taylor, B.P.; Fontana, E.; Amoretti, L.A.; Wright, R.J.; Xavier, J.B. A compilation of fecal microbiome shotgun metagenomics from hematopoietic cell transplantation patients. *Sci. Data* **2022**, *9*, 1–11. [[CrossRef](#)] [[PubMed](#)]
25. Paczosa, M.K.; Mecsas, J. Klebsiella pneumoniae: Going on the offense with a strong defense. *MMBR* **2016**, *80*, 629–661. [[CrossRef](#)] [[PubMed](#)]
26. Zois, C.E.; Harris, A.L. Glycogen metabolism has a key role in the cancer microenvironment and provides new targets for cancer therapy. *J. Mol. Med.* **2016**, *94*, 137–154. [[CrossRef](#)]
27. Pannen, D.; Fabisch, M.; Gausling, L.; Schnetz, K. Interaction of the RcsB response regulator with auxiliary transcription regulators in *Escherichia coli*. *JBC* **2016**, *291*, 2357–2370. [[CrossRef](#)]
28. Bronowski, C.; Smith, S.L.; Yokota, K.; Corkill, J.E.; Martin, H.M.; Campbell, B.J.; Rhodes, J.M.; Hart, C.A.; Winstanley, C. A subset of mucosa-associated *Escherichia coli* isolates from patients with colon cancer, but not Crohn’s disease, share pathogenicity islands with urinary pathogenic *E. coli*. *Microbiology* **2008**, *154*, 571–583. [[CrossRef](#)]
29. Saha, S.; Yao, S.; Elakad, O.; Lois, A.; Henric-Petri, H.; Buentzel, J.; Hinterthaler, M.; Danner, B.C.; Ströbel, P.; Emmert, A.; et al. UDP-glucose 6-dehydrogenase expression as a predictor of survival in patients with pulmonary adenocarcinoma. *IJS Oncol.* **2020**, *5*, e85. [[CrossRef](#)]
30. Pasi, S.; Kant, R.; Surolia, A. Toll/interleukin-1 receptor domain derived from TcpC (TIR-TcpC) ameliorates experimental autoimmune arthritis by down-modulating Th17 cell response. *JBC* **2016**, *291*, 12358–12369. [[CrossRef](#)]
31. Marsden, A.E.; Intile, P.J.; Schulmeyer, K.H.; Simmons-Patterson, E.R.; Urbanowski, M.L.; Wolfgang, M.C.; Yahr, T.L. Vfr directly activates exsA transcription to regulate expression of the Pseudomonas aeruginosa type III secretion system. *J. Bacteriol. Res.* **2016**, *198*, 1442–1450. [[CrossRef](#)]
32. Bulir, D.C.; Waltho, D.A.; Stone, C.B.; Mwawasi, K.A.; Nelson, J.C.; Mahony, J.B. Chlamydia pneumoniae CopD translocator protein plays a critical role in type III secretion (T3S) and infection. *PLoS ONE* **2014**, *9*, e99315. [[CrossRef](#)]
33. Wilson, M.M.; Anderson, D.E.; Bernstein, H.D. Analysis of the outer membrane proteome and secretome of Bacteroides fragilis reveals a multiplicity of secretion mechanisms. *PLoS ONE* **2015**, *10*, e0117732. [[CrossRef](#)]
34. Allali, I.; Boukhatem, N.; Bouguenouch, L.; Hardi, H.; Boudouaya, H.A.; Cadenas, M.B.; Ouldim, K.; Amzazi, S.; Azcarate-Peril, M.A.; Ghazal, H. Gut microbiome of Moroccan colorectal cancer patients. *Med. Microbiol.* **2018**, *207*, 211–225. [[CrossRef](#)]
35. Saha, P.; Yeoh, B.S.; Xiao, X.; Golonka, R.M.; Abokor, A.A.; Wenceslau, C.F.; Shah, Y.M.; Joe, B.; Vijay-Kumar, M. Enterobactin induces the chemokine, interleukin-8, from intestinal epithelia by chelating intracellular iron. *Gut Microbes* **2020**, *12*, 1841548. [[CrossRef](#)]
36. Galardini, M.; Clermont, O.; Baron, A.; Busby, B.; Dion, S.; Schubert, S.; Beltrao, P.; Denamur, E. Major role of iron uptake systems in the intrinsic extra-intestinal virulence of the genus Escherichia revealed by a genome-wide association study. *PLoS Genet.* **2020**, *16*, e1009065. [[CrossRef](#)]
37. Basak, D.; Uddin, M.N.; Hancock, J. The role of oxidative stress and its counteractive utility in colorectal cancer (CRC). *Cancers* **2020**, *12*, 3336. [[CrossRef](#)]
38. Allsopp, L.P.; Bernal, P.; Nolan, L.M.; Filloux, A. Causalities of war: The connection between type VI secretion system and microbiota. *Cell. Microbiol.* **2020**, *22*, e13153. [[CrossRef](#)]
39. Fabbri, A.; Travaglione, S.; Rosadi, F.; Ballan, G.; Maroccia, Z.; Giambenedetti, M.; Guidotti, M.; Ødum, N.; Krejsgaard, T.; Fiorentini, C. The *Escherichia coli* protein toxin cytotoxic necrotizing factor 1 induces epithelial mesenchymal transition. *Cell. Microbiol.* **2020**, *22*, e13138. [[CrossRef](#)]
40. Bernacki, R.J.; Niedbala, M.J.; Korytynk, W. Glycosidases in cancer and invasion. *Cancer Metastasis Rev.* **1985**, *4*, 81–101. [[CrossRef](#)]
41. Oliero, M.; Calvé, A.; Fragoso, G.; Cuisiniere, T.; Hajjar, R.; Dobrindt, U.; Santos, M.M. Oligosaccharides increase the genotoxic effect of colibactin produced by pks+ *Escherichia coli* strains. *BMC Cancer* **2021**, *21*, 1–10. [[CrossRef](#)]
42. Dalmaso, G.; Cougnoux, A.; Delmas, J.; Darfeuille-Michaud, A.; Bonnet, R. The bacterial genotoxin colibactin promotes colon tumor growth by modifying the tumor microenvironment. *Gut Microbes* **2014**, *5*, 675–680. [[CrossRef](#)]

43. Jamet, A.; Dervyn, R.; Lapaque, N.; Bugli, F.; Perez-Cortez, N.G.; Blottière, H.M.; Twizere, J.-C.; Sanguinetti, M.; Posteraro, B.; Serror, P.; et al. The Enterococcus faecalis virulence factor ElrA interacts with the human Four-and-a-Half LIM Domains Protein 2. *Sci. Rep.* **2017**, *7*, 1–13. [[CrossRef](#)]
44. Liu, C.; Li, Z.; Ding, J.; Zhen, H.; Fang, M.; Nie, C. Species-Level Analysis of the Human Gut Microbiome Shows Antibiotic Resistance Genes Associated with Colorectal Cancer. *Front. Microbiol.* **2021**, *12*, 765291. [[CrossRef](#)] [[PubMed](#)]
45. Dejea, C.M.; Wick, E.C.; Hechenbleikner, E.M.; White, J.R.; Welch, J.L.M.; Rossetti, B.J.; Peterson, S.N.; Snesrud, E.C.; Borisy, G.G.; Lazarev, M.; et al. Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18321–18326. [[CrossRef](#)] [[PubMed](#)]
46. Ellermann, M.; Arthur, J.C. Siderophore-mediated iron acquisition and modulation of host-bacterial interactions. *Free. Radic. Biol. Med.* **2017**, *105*, 68–78. [[CrossRef](#)] [[PubMed](#)]
47. Kwon, O.; Hudspeth, M.; Meganathan, R. Anaerobic biosynthesis of enterobactin *Escherichia coli*: Regulation of entC gene expression and evidence against its involvement in menaquinone (vitamin K2) biosynthesis. *J. Bacteriol.* **1996**, *178*, 3252–3259. [[CrossRef](#)]
48. Golonka, R.; Yeoh, B.S.; Vijay-Kumar, M. The iron tug-of-war between bacterial siderophores and innate immunity. *J. Innate Immun.* **2019**, *11*, 249–262. [[CrossRef](#)]
49. Aksan, A.; Farrag, K.; Aksan, S.; Schroeder, O.; Stein, J. Flipside of the coin: Iron deficiency and colorectal cancer. *Front. Immunol.* **2021**, *12*, 644. [[CrossRef](#)]
50. Fu, D.; Richardson, D.R. Iron chelation and regulation of the cell cycle: 2 mechanisms of posttranscriptional regulation of the universal cyclin-dependent kinase inhibitor p21CIP1/WAF1 by iron depletion. *Am. J. Hematol.* **2007**, *110*, 752–761. [[CrossRef](#)]