



HAL
open science

The Archaeome's Role in Colorectal Cancer: Unveiling the DPANN Group and Investigating Archaeal Functional Signatures

Nour El Houda Mathlouthi, Imen Belguith, Mariem Yengui, Hamadou Oumarou Hama, Jean-Christophe Lagier, Leila Ammar Keskes, Ghiles Grine, Radhouane Gdoura

► To cite this version:

Nour El Houda Mathlouthi, Imen Belguith, Mariem Yengui, Hamadou Oumarou Hama, Jean-Christophe Lagier, et al.. The Archaeome's Role in Colorectal Cancer: Unveiling the DPANN Group and Investigating Archaeal Functional Signatures. *Microorganisms*, 2023, 11 (11), pp.2742-10.3390/microorganisms11112742 . hal-04537518

HAL Id: hal-04537518

<https://amu.hal.science/hal-04537518>

Submitted on 8 Apr 2024

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

The Archaeome's Role in Colorectal Cancer: Unveiling the DPANN Group and Investigating Archaeal Functional Signatures

Nour El Houda Mathlouthi ¹, Imen Belguith ², Mariem Yengui ¹, Hamadou Oumarou Hama ³, Jean-Christophe Lagier ³, Leila Ammar Keskes ², Ghiles Grine ^{3,4} and Radhouane Gdoura ^{1,*}

- ¹ Laboratoire de Recherche Toxicologie Microbiologie Environnementale et Santé (LR17ES06), Faculté des Sciences de Sfax, University of Sfax, Sfax 3000, Tunisia; nourhouda0810@gmail.com (N.E.H.M.); yenguiimariem1995@gmail.com (M.Y.)
- ² Laboratoire de Recherche de Génétique Moléculaire Humaine, Faculté de Médecine de Sfax, University of Sfax, Avenue Majida BOULILA, Sfax 3029, Tunisia; mayno87.ib@gmail.com (I.B.); ammarkeskesl@gmail.com (L.A.K.)
- ³ IHU Méditerranée Infection, l'unité de Recherche Microbes, Evolution, Phylogénie et Infection (MEPHI), 19-21, Bd. Jean Moulin, 13005 Marseille, France; hamadouh2o@gmail.com (H.O.H.); jean-christophe.lagier@univ-amu.fr (J.-C.L.); grineghiles@gmail.com (G.G.)
- ⁴ Institut de Recherche pour le Développement (IRD), Aix-Marseille Université, IHU Méditerranée Infection, l'unité de Recherche Microbes, Evolution, Phylogénie et Infection (MEPHI), 13005 Marseille, France
- * Correspondence: radhouane.gdoura@fss.usf.tn



Citation: Mathlouthi, N.E.H.; Belguith, I.; Yengui, M.; Oumarou Hama, H.; Lagier, J.-C.; Ammar Keskes, L.; Grine, G.; Gdoura, R. The Archaeome's Role in Colorectal Cancer: Unveiling the DPANN Group and Investigating Archaeal Functional Signatures. *Microorganisms* **2023**, *11*, 2742. <https://doi.org/10.3390/microorganisms11112742>

Academic Editor: Konstantinos Triantafyllou

Received: 21 August 2023
Revised: 25 September 2023
Accepted: 11 October 2023
Published: 10 November 2023



Copyright: © 2023 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/>).

Abstract: Background and Aims: Gut microbial imbalances are linked to colorectal cancer (CRC), but archaea's role remains underexplored. Here, using previously published metagenomic data from different populations including Austria, Germany, Italy, Japan, China, and India, we performed bioinformatic and statistical analysis to identify archaeal taxonomic and functional signatures related to CRC. Methods: We analyzed published fecal metagenomic data from 390 subjects, comparing the archaeomes of CRC and healthy individuals. We conducted a biostatistical analysis to investigate the relationship between *Candidatus Mancarchaeum acidiphilum* (DPANN superphylum) and other archaeal species associated with CRC. Using the Prokka tool, we annotated the data focusing on archaeal genes, subsequently linking them to CRC and mapping them against UniprotKB and GO databases for specific archaeal gene functions. Results: Our analysis identified enrichment of methanogenic archaea in healthy subjects, with an exception for *Methanobrevibacter smithii*, which correlated with CRC. Notably, CRC showed a strong association with archaeal species, particularly *Natrinema* sp. J7-2, *Ferroglobus placidus*, and *Candidatus Mancarchaeum acidiphilum*. Furthermore, the DPANN archaeon exhibited a significant correlation with other CRC-associated archaea ($p < 0.001$). Functionally, we found a marked association between MvhB-type polyferredoxin and colorectal cancer. We also highlighted the association of archaeal proteins involved in the biosynthesis of leucine and the galactose metabolism process with the healthy phenotype. Conclusions: The archaeomes of CRC patients show identifiable alterations, including a decline in methanogens and an increase in Halobacteria species. MvhB-type polyferredoxin, linked with CRC and species like *Candidatus Mancarchaeum acidiphilum*, *Natrinema* sp. J7-2, and *Ferroglobus placidus* emerge as potential archaeal biomarkers. Archaeal proteins may also offer gut protection, underscoring archaea's role in CRC dynamics.

Keywords: colorectal cancer; gut microbiome; archaea; DPANN group; MvhB-type polyferredoxin; *Ferroglobus placidus*; *Natrinema* sp. J7-2; archaeal biomarker

1. Introduction

Colorectal cancer (CRC) is considered to be one of the deadliest cancers worldwide, with an estimated incidence of 1.9 million cases in 2020 [1]. The gut microbiota appear to

be one key factor in CRC tumorigenesis [2]. It is a complex ecosystem that includes many microorganisms hosted in the digestive tract, mainly bacteria, viruses, fungi, and archaea.

An impairment of the gut microbiota, referred to as dysbiosis, can favor the colonization of the intestinal mucosa by opportunistic microbes [3], which could be involved in carcinogenesis according to a “driver-passenger” model. The first stage consists of colon colonization by pathogenic “driver” bacteria with pro-inflammatory and pro-carcinogenic potential, such as *Bacteroides fragilis* and *Escherichia coli* species, contributing to the initiation of CRC [4]. The second step involves opportunistic “passenger” bacteria, such as *Fusobacterium nucleatum*, leading to tumor growth and its progression via interactions with the host immune system or the production of metabolites [5].

Viromes were also found to be changed when comparing the late and early stages of CRC. Advanced CRC microbiomes were enriched with several bacteriophages such as Betabaculovirus and Mulikevirus [6]. In addition, the abundance of human papillomavirus (HPV) and bocavirus was found to be significantly higher in tumoral tissues compared to adjacent tissues [7]. It was proposed that colon viruses could induce DNA alteration and alter immune homeostasis [6,8].

CRC is also linked to a fecal fungal dysbiosis presented by an abundance of Basidiomycetes and Ascomycetes as well as opportunistic fungi such as *Malassezia*. In contrast, the depletion of Saccharomycetes such as *Saccharomyces cerevisiae* was highlighted. This fungal species was found to exhibit regulatory and anti-inflammatory effects on the host by inducing interleukin-10 production [9,10].

However, due to the difficulty of culturing, detecting, and identifying it, archaea have remained “a forgotten player in the human microbiota” [11,12], and researchers only recently brought into focus the importance of the archaeome in CRC [13].

Archaea are prokaryotic unicellular microorganisms, and their entirety, including their genetic material in a particular environment, is called an archaeome [14]. Some of these microbes are characterized by membranes built with ether/isoprenoid chains of L-glycerol, and their cell walls lack peptidoglycan [15], while others are shared with eukaryotes or bacteria. The archaea were mainly considered extremophiles, residing in environments foreign to the human habitat and the microbial ecosystem. Thanks to the progress of molecular methods, especially the development of DNA sequencing and metagenomics, the presence of archaea in non-extreme environments has been revealed [16].

Indeed, metagenomic analyses highlighted its presence in the human microbiome and its ability to behave as a pathogen. Among its characteristics that may mediate human pathogenicity, we cite the ability to access and colonize the human microbiome and to compete with other endogenous microbes for survival, inducing dysbiosis [16], and the possibility to exchange their genetic material via the horizontal gene transfer process, i.e., exchanging pathogenic genes [17]. It was shown previously that the only groups of archaea involved in the dysbiosis of the human microbiota are the methanogenic archaea (named methanogens within the manuscript), which were found to be reduced in Crohn’s disease, ulcerative colitis, and severe acute malnutrition [18,19]. One of the most studied methanogens was *Methanobrevibacter smithii*, which is so far the only archaea implicated in microbial dysbiosis, including vaginosis and urinary tract infection [20]. Recently, a study performed on a Chinese population showed enrichment of halophilic archaea and depletion of methanogens in CRC compared to healthy subjects [13].

Among the archaea clades detected recently in the human gut microbiota [21] is the DPANN (Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanoarchaeota, and Nanohaloarchaeota) superphylum proposed in 2013 [22], characterized by its small cell size, limited metabolic capacities [23], and unique genetic characteristics [24]. The lack of genes associated with essential metabolic pathways in these microorganisms indicates a potential dependence on symbiotic relationships with other organisms, or even a parasitic mode of existence [25]. While their symbiotic or parasitic lifestyles could increase the chances for horizontal gene transfer with potentially pathogenic organisms, especially given their

prevalence in tumor microenvironments [26], there is currently no established evidence linking them to colorectal carcinogenesis.

A prior study linked specific archaeal metabolites, such as short-chain fatty acids, indoles, and arsenic, to carcinogenesis. Metabolites such as H₂ and trimethylamine (TMA) have been linked to carcinogens, specifically H₂S and trimethylamine N-oxide (TMAO). These carcinogens are implicated in processes like inflammation, DNA damage, genotoxicity, and disrupted signaling pathways [21]. It is worth noting, however, that these metabolites can also be produced by other microorganisms, such as bacteria [27,28]. Thus, the unique functional signature of archaea in the gut microbiome of colorectal cancer remains to be discovered.

In our study, we delved into the relationship between the archaeome and colorectal cancer by examining 390 published fecal metagenomics datasets from both CRC patients and healthy individuals. We employed bivariate and multivariate analyses to explore the associations of the DPANN archaeon (*Candidatus Mancarchaeum acidiphilum*) with other detected archaeal species. Furthermore, we conducted a functional analysis to pinpoint the unique archaeal functional signatures present in the gut of CRC patients.

2. Materials and Methods

2.1. Published Fecal Metagenomic Data

Published metagenomes of 390 fecal specimens were examined, including samples from 192 CRC patients and 198 randomly selected controls (Table 1). Datasets of published fecal metagenomes are available from the European Nucleotide Archive (ENA) database “<https://www.ebi.ac.uk/ena/browser/home>” (accessed on 20–25 April 2022). Bioproject accession numbers are presented in Table 1. The clinical characteristics are shown in Supplementary Data S1.

Table 1. Summary of study populations.

Project n°	CRC	Healthy	Country	Reference
PRJEB7774	20	24	Austria	[29]
PRJEB27928	42	39	Germany	[30]
PRJNA447983	22	22	Italy	[31]
PRJDB4176	33	35	Japan	[32]
PRJEB10878	46	48	China	[33]
PRJNA397112	29	30	India	[34]

2.2. Bioinformatic Processing and Biostatistical Analysis

2.2.1. Taxonomic Analysis

Archaea Identification

The assignment of taxonomic labels to the published metagenomic data of 390 fecal samples was performed with Kraken, which is a taxonomic sequence classifier that labels short DNA reads with taxonomic labels [35]. This was performed by looking at the k-mers in a read and then querying a database with those k-mers. In a taxonomic tree of all genomes containing that k-mer, every k-mer in Kraken’s genomic library was mapped to its lowest common ancestor (LCA). After that, the set of LCA taxa corresponding to a read’s k-mers was analyzed to produce a single taxonomic label for the read, which could be any node in the taxonomic tree [35]. The taxonomic analysis of the metagenomic data was performed using the Galaxy Europe platform and its tool “Kraken” by mapping them against the archaea database “archaea 2020”; we used default parameters. The outputs of Kraken were then assembled into one single output “Kraken abundance table” (Supplementary Data S2), using the “Kraken taxonomic report” tool “<https://usegalaxy.eu/>” (accessed on 20 April to 2 May 2022).

Data Filtering and Archaeal Species Patterns in Colorectal Cancer

Data filtering was applied to remove low-quality features from the data, as very small numbers in very few samples were likely due to sequencing errors. To reduce false positives resulting from sequencing errors or low-level contamination, taxonomic characteristics were filtered based on their minimum count (set at four) and prevalence (at least 20% of samples containing four or more counts). Additionally, features that remained constant throughout the experiment were removed by filtering based on low variance (10% of features, determined using the inter-quantile range). Following filtering, library size differences were adjusted using cumulative sum scaling (CSS) normalization. Data filtering and normalization were performed using microbiomeAnalysit "<https://www.microbiomeanalyst.ca/MicrobiomeAnalyst/home.xhtml>" (accessed on 3–5 May 2022). Spearman correlation was used to calculate the correlation between archaeal species, CRC, and healthy phenotypes. A p -value less than 0.05 (typically ≤ 0.05) was statistically significant. Statistical analyses were performed using SPSS for Windows 20.0 (SPSS 20.0 for Windows; SPSS Inc., Chicago, IL, USA).

Correlation Analysis of DPANN Archaea and Other Archaeal Species

The Spearman's correlation coefficient was calculated to estimate the possible associations between the DPANN archaeon *Candidatus Mancarchaeum acidiphilum* and the different archaeal species detected in the gut microbiomes of healthy subjects and CRC patients (Supplementary Data S3). A p -value less than 0.05 (typically ≤ 0.05) was statistically significant. Statistical analyses were performed using SPSS for Windows 20.0 (SPSS 20.0 for Windows; SPSS Inc., Chicago, IL, USA).

Multivariate Analysis

A multivariate analysis was performed to identify the association between *Candidatus Mancarchaeum acidiphilum*, considered an independent factor, and other archaeal species significantly associated with CRC or healthy phenotypes and found to be strongly correlated with the DPANN archaeon following bivariate analysis, which are *Methanocella conradii*, *Natrinema* sp. J7-2, *Ferroglobus placidus*, *Methanobrevibacter olleyae*, *Sulfolobus* sp. E5-1-F, *Pyrolobus fumarii*, and *Thermococcus chitonophagus* (Supplementary Data S4). Statistical analyses were performed using SPSS for Windows 20.0 (SPSS 20.0 for Windows; SPSS Inc., Chicago, IL, USA). A two-sided p -value of less than 0.05 was considered statistically significant.

2.2.2. Data Assembly

De novo assembly of the metagenomic data was performed using the MEGAHIT tool in the Galaxy Europe platform <https://usegalaxy.eu/> (accessed on 1–6 June 2023). MEGAHIT is a computational tool for ultra-fast assembly of large and complex metagenomics sequencing data [36]. The adjustment of assembly parameters was carried out following the instructions of the tutorial provided by the Galaxy Training Network [37]. Briefly, in the basic assembly options, the minimum k-mer size was adjusted to 21, the maximum k-mer size was adjusted to 91, and the increment of k-mer size for each iteration was fixed at 12.

2.2.3. Functional Analysis

Genome Annotation

A functional study was then carried out using the contigs generated by the MEGAHIT tool. For this investigation, we utilized Prokka, a rapid prokaryotic genome annotation tool hosted on the Galaxy Europe platform. Prokka is used to predict and annotate coding sequences, rRNAs, tRNAs, and other genomic features. It annotates microbial genomes by integrating several bioinformatics tools, including Prodigal, RNAmmer, Aragorn, SignalP, and Infernal [38]. To map the contigs against only archaeal genes, we adjusted the "kingdom" option to "Archaea" "<https://usegalaxy.eu/>" (accessed on 5–7 June 2023).

Association of Archaeal Genes with Colorectal Cancer

The set of archaeal genes found in the studied metagenomic data was then analyzed using Spearman's correlation test in Python [39] to establish their relationship with the CRC phenotype. The genes found to be correlated with CRC were then mapped against the UniprotKB database (<https://www.uniprot.org/>), accessed on 8–10 June 2023, to identify the genes that are mostly or only found in archaeal species. The identified genes were mapped against Uniprot-Gene Ontology (GO) database (https://www.uniprot.org/help/gene_ontology), accessed on 12–14 June 2023, to find their biological function.

3. Results

3.1. Study Characterization and Data Processing

We analyzed published fecal metagenomic data generated from metagenomic shotgun sequencing of fecal samples collected from 192 CRC patients and 198 healthy controls (Table 1, Supplementary Data S1). The pan-archaeome included 7 archaeal phyla, 11 archaeal classes, and 294 archaeal species found in both CRC and healthy samples (Supplementary Data S2). A total of 62 low-abundance features (archaeal species) were removed based on prevalence, and 25 low-variance features were removed based on an inter-quantile range (IQR); 216 features remained after data filtering (Supplementary Data S3). For the archaeal phyla and archaeal classes, the filtering step eliminated one and two low-variance features, respectively (Supplementary Data S3).

3.2. Archaeal Communities Associated with CRC Microbiomes

The taxonomic analysis revealed the presence of six phyla in the gut microbiomes of healthy and CRC subjects, which are Euryarchaeota, Thermoproteota, Nitrososphaerota, Candidatus Geothermarchaeota, Candidatus Micrarchaeota, and Candidatus Lokiarchaeota. A strong correlation between Candidatus Micrarchaeota, a phylum belonging to the superphylum DPANN, and CRC patients was highlighted ($p < 0.001$). Euryarchaeota was also correlated with the CRC phenotype, unlike Candidatus Geothermarchaeota, which was found to be associated with healthy subjects (Table 2).

Table 2. Correlation analysis of archaeal phyla and CRC phenotype.

Phylum	Spearman's Correlation Coefficient (r)	p -Value
Candidatus Micrarchaeota	0.199	0.000
Candidatus Geothermarchaeota	−0.147	0.004
Euryarchaeota	0.118	0.020
Thermoproteota	−0.046	0.369
Candidatus Lokiarchaeota	0.005	0.916
Nitrososphaerota	−0.004	0.933

In our study, we detected the presence of nine classes after the filtering step of the data presented by Archaeoglobi, Thermoplasmata, Methanococci, Halobacteria, Methanobacteria, Thermoprotei, Nitrososphaeria, Methanomicrobia, and Thermococci. We found a significant positive correlation between Archaeoglobi and CRC patients ($p = 0.017$) (Table 3).

The Spearman correlation showed a significant association between the different *Halobacteria* sp. and CRC phenotype. *Halapricum salinum*, *Halomicrobium* sp. LC1Hm, *Halosimplex rubrum*, *Salinibaculum litoreum*, *Haloferax gibbonsii*, *Halopenitus persicus*, *Halorubrum* sp. RHB-C, *Halorubrum trapanicum*, *Halostagnicola larsenii*, and *Natrinema* sp. J7-2 were positively and highly correlated with CRC gut microbiomes ($p < 0.001$) (Supplementary Data S4). As for *Haloferax alexandrines* and *Halorubrum lacusprofundi*, these halophilic archaea were highly correlated with healthy subjects ($p < 0.001$). Other than halophilic archaea, *Methanoculleus bourgensis*, *Methanocella conradii*, *Methanosphaerula palustris*, and *Thermococcus* sp. were correlated with CRC patients ($p < 0.001$). *Methanobacteria* sp. and *Methanomicrobia* sp. were more abundant in healthy subjects' gut microbiomes than in CRC patients. From

the Methanobacteriaceae family, we highlighted a strong correlation ($p < 0.001$) between healthy subjects and *Methanothermobacter wolfeii*, *Methanobrevibacter olleyae*, *Methanobacterium paludism*, and *Methanobacterium* sp. MZ-A1. In contrast, *Methanobrevibacter smithii* showed a high correlation with CRC patients ($p < 0.001$). Regarding *Methanomicrobia* sp., correlation analysis revealed a strong and significant association ($p < 0.001$) between healthy gut microbiomes and *Methanosarcina barkeri*, *Methanosarcina flavescens*, and *Methanosarcina siciliae*. Two archaeal species belonging to the Archaeoglobi class showed different correlations with the CRC phenotype; *Ferroglobus placidus* was found to be highly correlated with CRC patients ($p < 0.001$), and *Archaeoglobus veneficus* was correlated with healthy subjects ($p = 0.025$). Our analysis revealed the detection of a DPANN archaeon, presented by *Candidatus Mancarchaeum acidiphilum*, and its correlation with the CRC gut microbiome ($p = 0.002$) (Supplementary Data S4).

Table 3. Correlation analysis of archaeal classes and CRC phenotype.

Class	Spearman's Correlation Coefficient (r)	p -Value
Archaeoglobi	0.1199	0.017844
Thermoplasmata	0.088356	0.081387
Methanococci	−0.08369	0.09889
Halobacteria	−0.05394	0.28799
Methanobacteria	0.042504	0.40256
Thermoprotei	−0.03954	0.43616
Nitrososphaeria	0.03813	0.45273
Methanomicrobia	−0.02483	0.62497
Thermococci	−0.02424	0.63325

3.3. DPANN Archaeon Association with the Different Archaeal Species Detected in Healthy and CRC Subjects

Bivariate analysis showed a significant positive association ($p < 0.001$) between *Candidatus Mancarchaeum acidiphilum* and archaeal species positively correlated with CRC gut microbiomes (Section 3.2): *Methanocella conradii*, *Methanoculleus bourgensis*, *Natrinema* sp. J7-2, *F. placidus*, and *Methanohalophilus halophilus*. However, the DPANN archaeon was negatively and significantly correlated ($p < 0.001$) with archaea previously revealed to be associated with healthy subjects (Section 3.2): *Sulfolobus* sp. E5-1-F, *Pyrolobus fumarii*, *Thermococcus chitonophagus*, *Methanobrevibacter olleyae*, and *Methanobacterium* sp. MZ-A1 (Supplementary Data S4). The multivariate analysis showed a highly significant correlation between *Candidatus Mancarchaeum acidiphilum* and both *Natrinema* sp. J7-2 and *F. placidus* ($p < 0.001$) (Table 4).

Table 4. Multivariate analysis of DPANN archaeon and its association with other archaeal species detected in the gut microbiome.

Model	Coefficients ^a					
	Unstandardized Coefficients		Standardized Coefficients	Sig.	95.0% Confidence Interval for B	
	B	Std. Error	Beta		Lower Bound	Upper Bound
<i>Thermococcus chitonophagus</i>	0.186	0.405	0.032	0.647	−0.610	0.981
<i>Ferroglobus placidus</i>	1.868	0.288	0.278	0.000	1.301	2.435
<i>Natrinema</i> sp. J7-2	0.872	0.081	0.461	0.000	0.713	1.031
<i>Methanocella conradii</i>	0.076	0.109	0.030	0.487	−0.138	0.289
<i>Sulfolobus</i> sp. E5-1-F	−0.199	0.298	−0.355	0.505	−0.784	0.387
<i>Pyrolobus fumarii</i>	0.269	0.522	0.273	0.607	−0.758	1.296
<i>Methanobrevibacter olleyae</i>	0.073	0.095	0.081	0.444	−0.114	0.259

^a Dependent variable: *Candidatus Mancarchaeum acidiphilum*.

3.4. Archaeal Proteins Identified in the Gut Microbiomes of Healthy Individuals and CRC Patients

Following annotation using Prokka, we identified 2427 archaeal proteins belonging to different functional categories, including enzymes (e.g., ribosomal RNA small subunit methyltransferase H, protein-arginine kinase, and glycine dehydrogenase subunits 1 and 2), structural proteins (e.g., DNA-binding protein and membrane-bound lytic murein transglycosylase F), and transport proteins (e.g., molybdate/tungstate-binding protein WtpA). Additionally, we noted the presence of regulatory proteins such as protein-arginine kinase and DNA mismatch repair protein MutL, as well as DNA/RNA-related proteins like DNA polymerase PolB subunit 25 and error-prone DNA polymerase. Proteins involved in metabolic processes, such as formate-tetrahydrofolate ligase and galactose-1-phosphate uridylyltransferase, were also detected, along with proteins involved in protein modification, including the NAD-dependent protein deacetylase sirtuin-5, mitochondrial (Supplementary Data S5).

3.5. Association of Proteins Exclusively Detected in Archaea with Colorectal Cancer

To understand the potential implications of the detected archaeal proteins in CRC, we performed a Spearman correlation analysis. We found that 704 of the proteins exhibited either a positive or negative correlation with CRC (Supplementary Data S6). The genes showing a significant association with the CRC phenotype (Supplementary Data S6) and having relative abundance in the studied metagenomic data (Supplementary Data S5) were mapped against the UniProtKB database to identify genes predominantly or exclusively detected in archaea (Supplementary Data S7). We mapped the identified archaeal genes against GO annotations to discern their biological roles. Significantly, we observed a negative correlation between CRC and the archaeal protein, beta-galactosidase BgaH ($p = 0.002$), which plays a role in galactose metabolic process. Similarly, there was a negative association with the methanogen homoacetylase large subunit ($p = 0.019$), an archaeal protein involved in the biosynthesis of leucine. On the other hand, the polyferredoxin protein MvhB, associated with energy metabolism, showed a positive correlation with CRC ($p = 0.045$) (Table 5).

Table 5. Association of proteins mostly or exclusively detected in archaea with CRC.

Proteins	Spearman's Correlation Coefficient (r)	p -Value	% of UniprotKB Search Results with Taxonomy: Archaea and Query: Protein	GO Annotation
Beta-galactosidase BgaH	−0.155543828	0.002	100%	Galactose Metabolic Process
Methanogen homoacetylase large subunit	−0.11878052	0.019	100%	Leucine biosynthetic process
Polyferredoxin protein MvhB	0.101673216	0.045	96%	Energy Metabolism

4. Discussion

CRC presents a public health problem; its incidence rate is estimated to increase from 1,931,590 new cases in 2020 to 3,154,674 by 2040 [1]. To prevent this huge increase, several scientific researchers have been trying to understand the mechanism of colorectal carcinogenesis, its initiation, and its development. Despite the multifactorial nature of colorectal cancer (CRC), the gut microbiota were found to be correlated with this pathology [40]. Unlike the bacteriome, which is the most analyzed component of the gut microbiota, it was only in 2020 that the fecal archaeal DNA of CRC patients was profiled and compared with that of healthy subjects from China [13].

Here, we employed several user-friendly methods to analyze the gut archaeomes of CRC patients and healthy individuals in a large cohort from different populations, including Galaxy, MicrobiomeAnalysit, and SPSS, for bioinformatics and statistical analyses. Galaxy offers a diverse set of bioinformatic tools specifically designed for metagenomic analysis that allow reproducible analysis workflows, collaboration, and complexity without

programming experience [41]. Meanwhile, MicrobiomeAnalysit enables complex analysis without bioinformatic expertise or programming knowledge. Together, Galaxy and MicrobiomeAnalysit provide a comprehensive and in-depth analysis of metagenomic data, accommodating users of varying experience levels [42]. We used a Python script for the functional analysis to investigate any potential relationships between the discovered genes and CRC. This method was required due to the large number of identified genes which were too numerous for SPSS to analyze.

Taxonomically, after metagenomic data filtering using MicrobiomeAnalysit, a total of 217 archaeal species belonging to different phyla (Euryarchaeota, Thermoproteota, Nitrososphaerota, Candidatus Geothermarchaeota, Candidatus Micrarchaeota, and Candidatus Lokiarchaeota) were found in healthy and CRC patients. Interestingly, we detected the presence of Candidatus Micrarchaeota belonging to the DPANN superphylum, and we revealed its strong positive correlation with CRC gut microbiomes ($p < 0.001$).

Statistical tests performed to compare the abundance of archaeal species between CRC patients and healthy subjects showed the enrichment of methanogens in healthy gut microbiomes, except for *M. smithii*, which was found to be strongly correlated with the CRC phenotype. Previous studies identified the human gut archaeome and found that it consisted mostly of methane-producing archaea (methanogens), especially Methanobacteriales [43]. Regarding the abundance of *M. smithii* in the CRC gut microbiomes, previous research showed conflicting results, and it was reported that this archaeon was associated with Indian CRC patients [44], unlike the Chinese [13] and American [45] populations, which showed its association with the healthy phenotype. We also highlighted the abundance of *Halobacteria* sp. in the CRC gut microbiomes, confirming the previous study of Coker et al. (2020), in which they demonstrated that *Natrinema* sp. J7-2 exhibited a gradual increase in abundance from the control group to patients with adenoma and then to those with CRC; furthermore, they proposed an association between the enrichment of some Haloarchaea species (*Halorubrum* sp., *Natrinema* sp., and *Halococcus* sp.) in CRC and the intake of salty food [13]. Several studies have reported a positive correlation between the consumption of salty foods and an increased risk of developing CRC [46–48]. This salty food habit, by promoting the proliferation of Haloarchaea, can disrupt the intestinal microbiome not only by altering the availability of nutrients on which intestinal bacteria depend for their growth and survival but also by producing “haloarcheocines”, which are antibiotics produced by the halophile archaea [49]. Another study in which a metagenomic analysis was performed on formalin-fixed paraffin-embedded tissues from Tunisian CRC patients showed a significant positive correlation between *Halobacteria* in tumor tissues compared to healthy mucosa [50].

Archaeoglobi, an archaeal class, was found to be correlated with CRC gut microbiomes according to our statistical analysis. *F. placidus* belonging to the Archaeoglobi was highly correlated with CRC microbiomes ($p < 0.001$). This archaeon is the first hyperthermophile known to grow anaerobically by oxidizing acetate and reducing ferric iron (Fe^{3+}) to ferrous iron (Fe^{2+}) [51]. Acetate has been shown to inhibit CRC cell proliferation and induce apoptosis [52], so its deficiency could be considered a risk factor for CRC. The reduction of ferric iron (Fe^{3+}) to ferrous iron (Fe^{2+}) can generate reactive oxygen species (ROS) [53]. In fact, the role of ROS in carcinogenesis remains controversial, as ROS can have both pro-cancerous and anti-cancerous effects depending on their concentration. At moderate concentrations, ROS can contribute to the initiation, development, and progression of malignant neoplasms through the promotion of cell proliferation, invasion, and angiogenesis. However, at high concentrations, ROS can cause cancer cell apoptosis, leading to tumor suppression [54]. Otherwise, other species of archaea (*Ferroplasma* sp. [55]) and bacteria (*Thiobacillus ferrooxidans* or *Gallionella ferruginea* [56] *Acidithiobacillus ferrooxidans* [57]) could oxidize Fe^{2+} , which can keep a limited concentration of the ferrous iron, thus inducing low concentrations of ROS. Regarding ferric iron and its relationship to CRC and the gut microbiome, a previous study found a significant correlation between microbial siderophores and advanced stages of CRC; taken together [58], we suggest the involvement of gut microbes, presented essentially

by siderophores-producing bacteria and *F. placidus* in the iron deficiency in CRC patients. Due to the essential role of iron in supporting immune functions, a lack of iron may lead to an elevated risk of developing cancer by affecting the process of tumor formation and immune surveillance, as well as potentially modifying the immune environment within the tumor [59].

A DPANN archaeon, *Candidatus Mancarchaeum acidiphilum*, was detected within the studied human gut metagenomic data. Previously, this archaeon was identified in acidic ecosystems [60]. Interactions between this archaeon and protuberances from another archaeal order, Thermoplasmatales, suggested potential metabolite, energy, and macromolecule (including DNA) exchange [61]. Although we did not find a significant correlation with Thermoplasmatales, we observed its positive association with *F. placidus* and *Natrinema* sp. J7-2 in the CRC phenotype context. Evidence from other studies, including findings on *Thermotoga maritima* [62] and *Methylobacterium extorquens* [16,63], indicates that archaeal DNA can be shared among both archaeal and non-archaeal species. When considering the association of DPANN archaea with pathogens, there are risks related to the spread of genes contributing to carcinogenic properties, virulence, resistance, and metabolism [64]. Such transfers could influence gut microorganisms, potentially exacerbating CRC development. The prevalence of archaeal microorganisms within gut microbiota is limited. Still, their significance should not be overlooked. To illustrate this point, consider viruses: although they constitute only about 0.2% of gut microbiota, they can incorporate their genes into the human genome, potentially triggering oncogenes [65,66]. Remarkably, a significant portion of genes in archaea, especially in the DPANN group, are classified as “hypothetical” proteins or genomic “dark matter”, accounting for 30% to possibly 80% of their total content [67]. While these genes are still uncharacterized, it should be noted that they may be able to encode pathogenic phenotypes, underscoring the necessity of thorough functional genomics investigations. Characteristics of DPANN archaea, such as reduced genome sizes [22], may indicate fewer obstacles to gene transfer or integration. There is a conceivable, though distant, route for gene transfer if DPANN archaea interact with viruses also infecting eukaryotic cells [68]. Such viruses could assimilate genes from DPANN archaea and later infuse them into a eukaryotic host.

At a functional level, we aimed to understand the unique pathogenic potential of archaea by focusing on genes specific to them. This emphasis arises because many genes can be found across various cellular organisms, including bacteria, eukaryotes, and archaea. Hence, simply mapping metagenomic data against the functional database of archaea might not delineate the distinct role of archaea in colorectal carcinogenesis. When performing Spearman correlation then mapping against UniProtKB, we found a significant correlation with three proteins exclusive to archaeal species. Notably, one of these is the MvhB-type polyferredoxin. Although all organisms include a variety of ferredoxin proteins, the “MvhB-type polyferredoxin” variant is only found in methanogenic archaea and not in other archaeal groups or bacteria [69]. This protein acts as an electron transfer chain in potential redox–enzyme complexes and plays a pivotal role in electron transfer processes. It is composed of three to seven 2[4Fe-4S]-ferredoxin units [69]. Our findings highlight a significant positive association between this protein and CRC. Given that polyferredoxins have an iron–sulfur composition, their production is closely tied to iron availability. Thus, an increase in MvhB-type polyferredoxin production might signal enhanced iron uptake by archaea. Consequently, we propose a potential risk associated with increased archaeal iron intake in CRC, both taxonomically and functionally. We also revealed that BgaH was inversely correlated with CRC. Delving into BgaH’s enzyme function, we found it primarily targets β -d-galactosides lactulose as key substrates [70], which suppresses harmful bacteria and enhances the production of beneficial metabolites like short-chain fatty acids (SCFAs) [71]. We also identified a negative correlation between the methanogen homoaconitase large subunit and the prevalence of CRC. This enzyme is critical for the biosynthesis of leucine, an amino acid recently found to be inversely associated with CRC as well [72]. Our findings are consistent with a prior study in which a decrease in microbial

genes associated with the production of branched-chain amino acids, specifically leucine, valine, and isoleucine, was highlighted during carcinogenesis [73].

Our study presents the first metagenomic analysis of the gut microbiome in Tunisian CRC patients. We focused on sequencing DNA taken from tumor-adjacent FFPE tissues from these patients. Significantly, the discovery of archaeal sequences in our samples underscores the viability of conducting an archaeome analysis from FFPE tissues using an affordable DNA extraction protocol and easy-to-use bioinformatic tools. Moreover, our findings highlight a notable enrichment of Halobacteria in the tumor mucosa when compared to the normal mucosa in these CRC patients. Based on these findings, it is important to highlight that the gut microbiome has been shown to have the ability to influence treatment response in several ways, including drug metabolism, immunomodulation, inflammation, and the tumor microenvironment [74,75]. Our findings might open the door for the creation of therapeutic and preventative strategies for colorectal cancer given this wider context.

In terms of limitations, our study's inability to conduct a comprehensive comparison of archaeomes across the various stages of colorectal carcinogenesis has limited our insight into their evolutionary patterns and influence on disease progression. Additionally, a significant number of genes in archaea are classified as "hypothetical" proteins or termed genomic "dark matter." As the exact functional roles of these genes remain undetermined, further research is urgently needed to discern their impact on human health and disease.

5. Conclusions

In conclusion, we explored the role of gut archaea in colorectal carcinogenesis and identified significant associations between CRC and specific archaeal species, notably *Ferroglobus placidus* and *Natrinema* sp. J7-2. Additionally, we detected the presence of the DPANN superphylum in the gut, indicating potential gene transfers that might impact CRC progression. We further emphasized the significance of certain proteins unique to archaeal species, such as the polyferredoxin protein MvhB, Beta-galactosidase BgaH, and Methanogen homoaconitase large subunit, underscoring the importance of comprehending the distinct pathogenic potential of archaea in CRC.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microorganisms11112742/s1>, Supplementary Data S1: Participant characterization; Supplementary Data S2: Kraken abundance table; Supplementary Data S3: Processed Kraken abundance; Supplementary Data S4: Biostatistical analysis (*): statistically significant correlation at $p < 0.05$. (**): highly significant correlation at $p < 0.01$.); Supplementary Data S5: Prokka annotation; Supplementary Data S6: Biostatistical analysis for the functional analysis; Supplementary Data S7: Uniprot annotation.

Author Contributions: Conceptualization, N.E.H.M., M.Y., H.O.H., L.A.K., G.G. and R.G.; Methodology, I.B., H.O.H. and G.G.; Software, N.E.H.M., I.B. and M.Y.; Validation, J.-C.L., L.A.K. and R.G.; Investigation, N.E.H.M. and H.O.H.; Resources, N.E.H.M.; Data curation, R.G.; Writing—original draft, N.E.H.M., I.B., M.Y. and H.O.H.; Writing—review & editing, J.-C.L., L.A.K., G.G. and R.G.; Visualization, R.G.; Supervision, L.A.K., G.G. and 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.

Data Availability Statement: All the data used for this study (metagenomic data) are already presented in the Supplementary Data and available online.

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

References

1. Xi, Y.; Xu, P. Global colorectal cancer burden in 2020 and projections to 2040. *Transl. Oncol.* **2021**, *14*, 101174. [[CrossRef](#)]
2. Tilg, H.; Adolph, T.E.; Gerner, R.R.; Moschen, A.R. The intestinal microbiota in colorectal cancer. *Cancer Cell* **2018**, *33*, 954–964. [[CrossRef](#)]
3. Wang, X.; Yang, Y.; Huycke, M.M. Commensal bacteria drive endogenous transformation and tumour stem cell marker expression through a bystander effect. *Gut* **2015**, *64*, 459–468. [[CrossRef](#)]
4. Lucas, C.; Barnich, N.; Nguyen, H.T.T. Microbiota, inflammation and colorectal cancer. *Int. J. Mol. Sci.* **2017**, *18*, 1310. [[CrossRef](#)]
5. Tjalsma, H.; Boleij, A.; Marchesi, J.R.; Dutilh, B.E. A bacterial driver–passenger model for colorectal cancer: Beyond the usual suspects. *Nat. Rev. Microbiol.* **2012**, *10*, 575–582. [[CrossRef](#)]
6. Massimino, L.; Lovisa, S.; Lamparelli, L.A.; Danese, S.; Ungaro, F. Gut eukaryotic virome in colorectal carcinogenesis: Is that a trigger? *Comput. Struct. Biotechnol. J.* **2021**, *19*, 16–28. [[CrossRef](#)]
7. Emler, C.; Ruffin, M.; Lamendella, R. Enteric virome and carcinogenesis in the gut. *Dig. Dis. Sci.* **2020**, *65*, 852–864. [[CrossRef](#)]
8. Coker, O.O. Non-bacteria microbiome (virus, fungi, and archaea) in gastrointestinal cancer. In *Microbiome in Gastrointestinal Cancer*; Springer: Singapore, 2023; pp. 91–106.
9. Kaźmierczak-Siedlecka, K.; Dvořák, A.; Folwarski, M.; Daca, A.; Przewłócka, K.; Makarewicz, W. Fungal gut microbiota dysbiosis and its role in colorectal, oral, and pancreatic carcinogenesis. *Cancers* **2020**, *12*, 1326. [[CrossRef](#)]
10. Coker, O.O.; Nakatsu, G.; Dai, R.Z.; Wu, W.K.K.; Wong, S.H.; Ng, S.C.; Chan, F.K.L.; Sung, J.J.Y.; Yu, J. Enteric fungal microbiota dysbiosis and ecological alterations in colorectal cancer. *Gut* **2019**, *68*, 654–662. [[CrossRef](#)]
11. Bang, C.; Schmitz, R.A. Archaea: Forgotten players in the microbiome. *Emerg. Top. Life Sci.* **2018**, *2*, 459–468.
12. Guindo, C.O.; Drancourt, M.; Grine, G. Digestive tract methanodrome: Physiological roles of human micro-biota-associated methanogens. *Microb. Pathog.* **2020**, *149*, 104425. [[CrossRef](#)]
13. Coker, O.O.; Wu, W.K.K.; Wong, S.H.; Sung, J.J.; Yu, J. Altered gut archaea composition and interaction with bacteria are associated with colorectal cancer. *Gastroenterology* **2020**, *159*, 1459–1470. [[CrossRef](#)]
14. Moissl-Eichinger, C.; Pausan, M.; Taffner, J.; Berg, G.; Bang, C.; Schmitz, R.A. Archaea are interactive components of complex microbiomes. *Trends Microbiol.* **2018**, *26*, 70–85. [[CrossRef](#)]
15. Nkamga, V.D.; Henrissat, B.; Drancourt, M. Archaea: Essential inhabitants of the human digestive micro-biota. *Hum. Microbiome J.* **2017**, *3*, 1–8. [[CrossRef](#)]
16. Eckburg, P.B.; Lepp, P.W.; Relman, D.A. Archaea and their potential role in human disease. *Infect. Immun.* **2003**, *71*, 591–596. [[CrossRef](#)]
17. Ochman, H.; Moran, N.A. Genes lost and genes found: Evolution of bacterial pathogenesis and symbiosis. *Science* **2001**, *292*, 1096–1099. [[CrossRef](#)]
18. Million, M.; Tidjani Alou, M.; Khelafia, S.; Bachar, D.; Lagier, J.C.; Dione, N.; Brah, S.; Hugon, P.; Lombard, V.; Armougom, F.; et al. Increased gut redox and depletion of anaerobic and methanogenic prokaryotes in severe acute malnutrition. *Sci. Rep.* **2016**, *6*, 26051. [[CrossRef](#)]
19. Scanlan, P.D.; Shanahan, F.; Marchesi, J.R. Human methanogen diversity and incidence in healthy and diseased colonic groups using *mcrA* gene analysis. *BMC Microbiol.* **2008**, *8*, 79. [[CrossRef](#)]
20. Grine, G.; Lotte, R.; Chirio, D.; Chevalier, A.; Raoult, D.; Drancourt, M.; Ruimy, R. Co-culture of *Methanobrevibacter smithii* with enterobacteria during urinary infection. *eBioMedicine* **2019**, *43*, 333–337. [[CrossRef](#)]
21. Cai, M.; Kandalai, S.; Tang, X.; Zheng, Q. Contributions of human-associated archaeal metabolites to tumor microenvironment and carcinogenesis. *Microbiol. Spectr.* **2022**, *10*, e02367-21. [[CrossRef](#)]
22. Rinke, C.; Schwientek, P.; Sczyrba, A.; Ivanova, N.N.; Anderson, I.J.; Cheng, J.F.; Darling, A.; Malfatti, S.; Swan, B.K.; Gies, E.A.; et al. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **2013**, *499*, 431–437. [[CrossRef](#)] [[PubMed](#)]
23. Zhou, Z.; Liu, Y.; Anantharaman, K.; Li, M. The expanding Asgard archaea invoke novel insights into Tree of Life and eukaryogenesis. *mLife* **2022**, *1*, 374–381. [[CrossRef](#)]
24. Baker, B.J.; De Anda, V.; Seitz, K.W.; Dombrowski, N.; Santoro, A.E.; Lloyd, K.G. Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* **2020**, *5*, 887–900. [[CrossRef](#)] [[PubMed](#)]
25. Dombrowski, N.; Lee, J.H.; Williams, T.A.; Offre, P.; Spang, A. Genomic diversity, lifestyles and evolutionary origins of DPANN archaea. *FEMS Microbiol. Lett.* **2019**, *366*, fnz008. [[CrossRef](#)] [[PubMed](#)]
26. Burns, M.B.; Lynch, J.; Starr, T.K.; Knights, D.; Blekhan, R. Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment. *Genome Med.* **2015**, *7*, 55. [[CrossRef](#)] [[PubMed](#)]
27. Liu, Y.; Dai, M. Trimethylamine N-oxide generated by the gut microbiota is associated with vascular inflammation: New insights into atherosclerosis. *Mediat. Inflamm.* **2020**, *2020*, 4634172. [[CrossRef](#)] [[PubMed](#)]
28. Murros, K.E. Hydrogen sulfide produced by gut bacteria may induce Parkinson’s disease. *Cells* **2022**, *11*, 978. [[CrossRef](#)]
29. Feng, Q.; Liang, S.; Jia, H.; Stadlmayr, A.; Tang, L.; Lan, Z.; Zhang, D.; Xia, H.; Xu, X.; Jie, Z.; et al. Gut microbiome development along the colorectal adenoma-carcinoma sequence. *Nat. Commun.* **2015**, *6*, 6528. [[CrossRef](#)]
30. Wirbel, J.; Pyl, P.T.; Kartal, E.; Zych, K.; Kashani, A.; Milanese, A.; Fleck, J.S.; Voigt, A.Y.; Palleja, A.; Ponnudurai, R.; et al. Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* **2019**, *25*, 679–689. [[CrossRef](#)]

31. Thomas, A.M.; Manghi, P.; Asnicar, F.; Pasolli, E.; Armanini, F.; Zolfo, M.; Beghini, F.; Manara, S.; Karcher, N.; Pozzi, C.; et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat. Med.* **2019**, *25*, 667–678. [[CrossRef](#)] [[PubMed](#)]
32. Yachida, S.; Mizutani, S.; Shiroma, H.; Shiba, S.; Nakajima, T.; Sakamoto, T.; Watanabe, H.; Masuda, K.; Nishimoto, Y.; Kubo, M.; et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* **2019**, *25*, 968–976. [[CrossRef](#)] [[PubMed](#)]
33. Yu, J.; Feng, Q.; Wong, S.H.; Zhang, D.; Yi Liang, Q.; Qin, Y.; Tang, L.; Zhao, H.; Stenvang, J.; Li, Y.; et al. Metagenomic analysis of fecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut* **2017**, *66*, 70–78. [[CrossRef](#)] [[PubMed](#)]
34. Dhakan, D.B.; Maji, A.; Sharma, A.K.; Saxena, R.; Pulikkan, J.; Grace, T.; Gomez, A.; Scaria, J.; Amato, K.R.; Sharma, V.K. The unique composition of Indian gut microbiome, gene catalogue, and associated fecal metabolome deciphered using multi-omics approaches. *Gigascience* **2019**, *8*, giz004. [[CrossRef](#)] [[PubMed](#)]
35. Wood, D.E.; Salzberg, S.L. Kraken: Ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **2014**, *15*, R46. [[CrossRef](#)]
36. Li, D.; Liu, C.M.; Luo, R.; Sadakane, K.; Lam, T.W. MEGAHIT: An ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **2015**, *31*, 1674–1676. [[CrossRef](#)]
37. Polunina, P.; Batut, B. Assembly of Metagenomic Sequencing Data (Galaxy Training Materials). Available online: <https://training.galaxyproject.org/training-material/topics/assembly/tutorials/metagenomics-assembly/tutorial.html> (accessed on 1 June 2023).
38. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)] [[PubMed](#)]
39. Stojiljković, M. NumPy, SciPy, and Pandas: Correlation with Python. Available online: <https://realpython.com/numpy-scipy-pandas-correlation-python/> (accessed on 9 June 2023).
40. Kostic, A.D.; Chun, E.; Robertson, L.; Glickman, J.N.; Gallini, C.A.; Michaud, M.; Clancy, T.E.; Chung, D.C.; Lochhead, P.; Hold, G.L.; et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe* **2013**, *14*, 207–215. [[CrossRef](#)]
41. Afgan, E.; Baker, D.; Batut, B.; Van Den Beek, M.; Bouvier, D.; Čech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B.A.; et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W537–W544. [[CrossRef](#)]
42. Chong, J.; Liu, P.; Zhou, G.; Xia, J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. *Nat. Protoc.* **2020**, *15*, 799–821. [[CrossRef](#)]
43. Gaci, N.; Borrel, G.; Tottey, W.; O’Toole, P.W.; Brugère, J.F. Archaea and the human gut: New beginning of an old story. *World J Gastroenterol.* **2014**, *20*, 16062. [[CrossRef](#)]
44. Gupta, A.; Dhakan, D.B.; Maji, A.; Saxena, R.; PK, V.P.; Mahajan, S.; Pulikkan, J.; Kurian, J.; Gomez, A.M.; Scaria, J.; et al. Association of *Flavonifractor plautii*, a flavonoid-degrading bacterium, with the gut microbiome of colorectal cancer patients in India. *MSystems* **2019**, *4*, e00438-19. [[CrossRef](#)]
45. Hansen, E.E.; Lozupone, C.A.; Rey, F.E.; Wu, M.; Guruge, J.L.; Narra, A.; Goodfellow, J.; Zaneveld, J.R.; McDonald, D.T.; Goodrich, J.A.; et al. Pan-genome of the dominant human gut-associated archaeon, *Methanobrevibacter smithii*, studied in twins. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (Suppl. S1), 4599–4606. [[CrossRef](#)] [[PubMed](#)]
46. Hoshiyama, Y.; Sekine, T.; Sasaba, T. A case-control study of colorectal cancer and its relation to diet, cigarettes, and alcohol consumption in Saitama Prefecture, Japan. *Tohoku J. Exp. Med.* **1993**, *171*, 153–165. [[CrossRef](#)] [[PubMed](#)]
47. Chottanapund, S.; Chamroonsawasdi, K.; Tunyasitthisundhorn, P.; Aekplakorn, W.; Silpasuwan, P.; Anantachoti, P.; Rojroongwasinkul, N.; Damapong, S.; Sornpaisarn, B.; Rojanapithayakorn, W.; et al. Modifiable Factors and Colon Cancer Risk in Thai Population. *Asian Pac. J. Cancer Prev.* **2021**, *22*, 37–43. [[CrossRef](#)]
48. Yusof, A.S.; Isa, Z.M.; Shah, S.A. Perceptions of Malaysian colorectal cancer patients regarding dietary in-take: A qualitative exploration. *Asian Pac. J. Cancer Prev.* **2013**, *14*, 1151–1154. [[CrossRef](#)]
49. de Castro, I.; Mendo, S.; Caetano, T. Antibiotics from haloarchaea: What can we learn from comparative genomics? *Mar. Biotechnol.* **2020**, *22*, 308–316. [[CrossRef](#)]
50. Mathlouthi, N.E.H.; Ouamrou Hama, H.; Belguith, I.; Charfi, S.; Boudawara, T.; Lagier, J.C.; Ammar Keskes, L.; Grine, G.; Gdoura, R. Colorectal Cancer Archaeome: A Metagenomic Exploration, Tunisia. *Curr. Issues Mol. Biol.* **2023**, *45*, 7572–7581. [[CrossRef](#)] [[PubMed](#)]
51. Tor, J.M.; Kashefi, K.; Lovley, D.R. Acetate oxidation coupled to Fe (III) reduction in hyperthermophilic microorganisms. *Appl. Environ. Microbiol.* **2001**, *67*, 1363–1365. [[CrossRef](#)] [[PubMed](#)]
52. Ferro, S.; Azevedo-Silva, J.; Casal, M.; Côte-Real, M.; Baltazar, F.; Preto, A. Characterization of acetate transport in colorectal cancer cells and potential therapeutic implications. *Oncotarget* **2016**, *7*, 70639–70653. [[CrossRef](#)]
53. Bennett, B.D.; Gralnick, J.A. Mechanisms of toxicity by and resistance to ferrous iron in anaerobic systems. *Free Radic. Biol. Med.* **2019**, *140*, 167–171. [[CrossRef](#)] [[PubMed](#)]
54. Aggarwal, V.; Tuli, H.S.; Varol, A.; Thakral, F.; Yerer, M.B.; Sak, K.; Varol, M.; Jain, A.; Khan, M.A.; Sethi, G. Role of Reactive Oxygen Species in Cancer Progression: Molecular Mechanisms and Recent Advancements. *Biomolecules* **2019**, *9*, 735. [[CrossRef](#)] [[PubMed](#)]

55. Golyshina, O.V.; Timmis, K.N. Ferroplasma and relatives, recently discovered cell wall-lacking archaea making a living in extremely acid, heavy metal-rich environments. *Environ. Microbiol.* **2005**, *7*, 1277–1288. [[CrossRef](#)] [[PubMed](#)]
56. Widdel, F.; Schnell, S.; Heising, S.; Ehrenreich, A.; Assmus, B.; Schink, B. Ferrous iron oxidation by anoxygenic phototrophic bacteria. *Nature* **1993**, *362*, 834–836. [[CrossRef](#)]
57. Meruane, G.; Vargas, T. Bacterial oxidation of ferrous iron by *Acidithiobacillus ferrooxidans* in the pH range 2.5–7.0. *Hydrometallurgy* **2003**, *71*, 149–158. [[CrossRef](#)]
58. 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. [[CrossRef](#)] [[PubMed](#)]
59. 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](#)]
60. Baker, B.J.; Comolli, L.R.; Dick, G.J.; Hauser, L.J.; Hyatt, D.; Dill, B.D.; Land, M.L.; VerBerkmoes, N.C.; Hettich, R.L.; Banfield, J.F. Enigmatic, ultrasmall, uncultivated Archaea. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 8806–8811. [[CrossRef](#)]
61. Golyshina, O.V.; Toshchakov, S.V.; Makarova, K.S.; Gavrillov, S.N.; Korzhenkov, A.A.; La Cono, V.; Arcadi, E.; Nechitaylo, T.Y.; Ferrer, M.; Kublanov, I.V.; et al. 'ARMAN' archaea depend on association with eu-ryarchaeal host in culture and in situ. *Nat. Commun.* **2017**, *8*, 60. [[CrossRef](#)]
62. Nesbo, C.L.; L'Haridon, S.; Stetter, K.O.; Doolittle, W.F. Phylogenetic analyses of two "archaeal" genes in *Thermotoga maritima* reveal multiple transfers between archaea and bacteria. *Mol. Biol. Evol.* **2001**, *18*, 362–375. [[CrossRef](#)]
63. Chistoserdova, L.; Vorholt, J.A.; Thauer, R.K.; Lidstrom, M.E. C1 transfer enzymes and coenzymes linking methylotrophic bacteria and methanogenic Archaea. *Science* **1998**, *281*, 99–102. [[CrossRef](#)]
64. Deng, Y.; Xu, H.; Su, Y.; Liu, S.; Xu, L.; Guo, Z.; Wu, J.; Cheng, C.; Feng, J. Horizontal gene transfer contributes to virulence and antibiotic resistance of *Vibrio harveyi* 345 based on complete genome sequence analysis. *BMC Genom.* **2019**, *20*, 761. [[CrossRef](#)] [[PubMed](#)]
65. Péneau, C.; Imbeaud, S.; La Bella, T.; Hirsch, T.Z.; Caruso, S.; Calderaro, J.; Paradis, V.; Blanc, J.F.; Letouzé, E.; Nault, J.C.; et al. Hepatitis B virus integrations promote local and distant oncogenic driver alterations in hepatocellular carcinoma. *Gut* **2022**, *71*, 616–626. [[CrossRef](#)] [[PubMed](#)]
66. Bousali, M.; Papatheodoridis, G.; Paraskevis, D.; Karamitros, T. Hepatitis B virus DNA integration, chronic infections and hepatocellular carcinoma. *Microorganisms* **2021**, *9*, 1787. [[CrossRef](#)] [[PubMed](#)]
67. Makarova, K.S.; Wolf, Y.I.; Koonin, E.V. Towards functional characterization of archaeal genomic dark matter. *Biochem. Soc. Trans.* **2019**, *47*, 389–398. [[CrossRef](#)] [[PubMed](#)]
68. Krupovic, M.; Cvirkaite-Krupovic, V.; Iranzo, J.; Prangishvili, D.; Koonin, E.V. Viruses of archaea: Structural, functional, environmental and evolutionary genomics. *Virus Res.* **2018**, *244*, 181–193. [[CrossRef](#)]
69. Watanabe, T.; Shima, S. MvhB-type polyferredoxin as an electron-transfer chain in putative redox-enzyme complexes. *Chem. Lett.* **2021**, *50*, 353–360. [[CrossRef](#)]
70. Holmes, M.L.; Dyall-Smith, M.L. Sequence and expression of a halobacterial β -galactosidase gene. *Mol. Microbiol.* **2000**, *36*, 114–122. [[CrossRef](#)]
71. Karakan, T.; Tuohy, K.M.; Janssen-van Solingen, G. Low-dose lactulose as a prebiotic for improved gut health and enhanced mineral absorption. *Front. Nutr.* **2021**, *8*, 672925. [[CrossRef](#)]
72. Rothwell, J.A.; Bešević, J.; Dimou, N.; Breeur, M.; Murphy, N.; Jenab, M.; Wedekind, R.; Viallon, V.; Ferrari, P.; Achaintre, D.; et al. Circulating amino acid levels and colorectal cancer risk in the European Prospective Investigation into Cancer and Nutrition and UK Biobank cohorts. *BMC Med.* **2023**, *21*, 80. [[CrossRef](#)]
73. Gao, R.; Zhu, Y.; Kong, C.; Xia, K.; Li, H.; Zhu, Y.; Zhang, X.; Liu, Y.; Zhong, H.; Yang, R.; et al. Alterations, interactions, and diagnostic potential of gut bacteria and viruses in colorectal cancer. *Front. Cell. Infect. Microbiol.* **2021**, *11*, 657867. [[CrossRef](#)]
74. Routy, B.; Gopalakrishnan, V.; Daillère, R.; Zitvogel, L.; Wargo, J.A.; Kroemer, G. The gut microbiota influences anticancer immunosurveillance and general health. *Nat. Rev. Clin. Oncol.* **2018**, *15*, 382–396. [[CrossRef](#)] [[PubMed](#)]
75. Fessler, J.; Matson, V.; Gajewski, T.F. Exploring the emerging role of the microbiome in cancer immunotherapy. *J. Immunother. Cancer* **2019**, *7*, 108. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.