



Article

Arabinoxylan and Pectin Metabolism in Crohn's Disease Microbiota: An *In Silico* Study

Carlos Sabater ^{1,2,*} , Inés Calvete-Torre ^{1,2} , Lorena Ruiz ^{1,2} and Abelardo Margolles ^{1,2}

¹ Department of Microbiology and Biochemistry of Dairy Products, Instituto de Productos Lácteos de Asturias-Consejo Superior de Investigaciones Científicas (IPLA-CSIC), Paseo Río Linares s/n, 33300 Villaviciosa, Asturias, Spain; ines.calvete@ipla.csic.es (I.C.-T.); lorena.ruiz@ipla.csic.es (L.R.); amargolles@ipla.csic.es (A.M.)

² Functionality and Ecology of Beneficial Microbes (MicroHealth) Group, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), 33011 Oviedo, Asturias, Spain

* Correspondence: carlos.sabater@csic.es; Tel.: +34-985-892131

Abstract: Inflammatory bowel disease is a chronic disorder including ulcerative colitis and Crohn's disease (CD). Gut dysbiosis is often associated with CD, and metagenomics allows a better understanding of the microbial communities involved. The objective of this study was to reconstruct *in silico* carbohydrate metabolic capabilities from metagenome-assembled genomes (MAGs) obtained from healthy and CD individuals. This computational method was developed as a mean to aid rationally designed prebiotic interventions to rebalance CD dysbiosis, with a focus on metabolism of emergent prebiotics derived from arabinoxylan and pectin. Up to 1196 and 1577 MAGs were recovered from CD and healthy people, respectively. MAGs of *Akkermansia muciniphila*, *Barnesiella viscericola* DSM 18177 and *Paraprevotella xyliniphila* YIT 11841 showed a wide range of unique and specific enzymes acting on arabinoxylan and pectin. These glycosidases were also found in MAGs recovered from CD patients. Interestingly, these arabinoxylan and pectin degraders are predicted to exhibit metabolic interactions with other gut microbes reduced in CD. Thus, administration of arabinoxylan and pectin may ameliorate dysbiosis in CD by promoting species with key metabolic functions, capable of cross-feeding other beneficial species. These computational methods may be of special interest for the rational design of prebiotic ingredients targeting at CD.

Keywords: arabinoxylan; pectin; metagenome-assembled genomes; carbohydrate metabolism; Crohn's disease; cross-feeding



Citation: Sabater, C.; Calvete-Torre, I.; Ruiz, L.; Margolles, A. Arabinoxylan and Pectin Metabolism in Crohn's Disease Microbiota: An *In Silico* Study. *Int. J. Mol. Sci.* **2022**, *23*, 7093. <https://doi.org/10.3390/ijms23137093>

Academic Editors:

Beatrice Cobucci-Ponzano,
Andrea Strazzulli and Marco Moracci

Received: 23 March 2022

Accepted: 22 June 2022

Published: 25 June 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

Inflammatory bowel disease (IBD) is a chronic and relapsing inflammatory disorder of the intestine with a multifactorial etiology. IBD mainly includes ulcerative colitis and Crohn's disease (CD) and its prevalence has been increasing worldwide [1,2]. It has been estimated that more than 1.5 million people in the United States and 2 million people in Europe suffer from IBD [3]. The progression of IBD involves complex interactions between the immune system, human microbiota and the environment [2,4]. In this regard, it is believed that dysbiosis of the human gut microbiota exacerbates IBD symptoms [4]. In this sense, metagenomic sequencing has been used to decipher alterations of the taxonomic composition and metabolic profiles of IBD microbiota as a mean to identify biomarkers of disease and targets for intervention [1]. However, metagenomics also allows recovering complete metagenome-assembled genomes (MAGs) from complex microbial communities and large datasets, which can provide valuable information on the metabolic potential of specific members of the community [5–7]. In addition, bioinformatics methods to analyse MAGs sequences in order to explore synergistic interactions between gut microbes have been developed [8]. In this regard, potential cross-feeding mechanisms between these gut bacteria can be elucidated *in silico*. Cross-feeding reduces competition between gut

microbes and enhances their growth. In this sense, essential symbionts and alternative symbionts involved in cross-feeding interactions can be determined. Essential symbionts comprise key microorganisms that occur in every minimal community of host MAGs needed to satisfy one specific metabolic function through metabolic cooperation of these bacteria. These functions involve the metabolisation of different compounds present in colon lumen. In contrast, alternative symbionts occur only in some of these minimal communities of cross-feeders. Therefore, any of the alternative symbionts can complete the missing metabolic functions of the minimal microbial community. These biologically relevant data cannot be obtained following assembly-free methods. However, the majority of studies have focused on the recovery of MAGs from healthy microbiota samples, and few have explored their potential to aid and reconstruct metabolic capacities in both health and dysbiosis states, with the final aim to assist the rational design of prebiotic interventions targeting gut microbiota modulation in specific population groups.

Genome annotation of MAGs can be used to study specific enzymatic activities such as those involved in carbohydrate metabolism [9]. Furthermore, MAGs could be used to build advanced genome-scale metabolic networks. These metabolic models reduce the complexity of large-scale microbiota into minimal communities with equivalent metabolic properties [8,10]. For example, genome-scale metabolic modelling revealed differential patterns in the bile acid metabolism of microbiomes of pediatric IBD patients [11]. To our knowledge, no previous attempts to elucidate symbiotic relationships in IBD microbiota and their influence on other metabolic activities such as its polysaccharide fermentation capability have been reported. In this regard, it has been described that the interaction of diet and the gut microbiota is perturbed in patients with IBD [12], while the intake of low fiber diets results in gut dysbiosis and promotes inflammation of the gut [12,13]. In contrast, pro- and prebiotic administration enhances the effectivity of IBD treatment [2]. Specifically, probiotic microorganisms belonging to *Bifidobacterium* and *Lactobacillus* genera may be beneficial for IBD remission. It should be considered that vegetable poly- and oligosaccharides comprise different families such as arabinoxylan and pectin-derived compounds that have been proposed as emerging prebiotics with enhanced bioactivity [14]. These prebiotic mixtures could be tailored to target specific diseases such as IBD [15,16]. In addition, previous studies suggest the beneficial effects of pectin, arabinoxylan and oligosaccharides derived from these substrates to ameliorate IBD symptoms in clinical research [17–19].

Arabinoxylan consist of a linear backbone of 1500 to 15,000 $\beta(1-4)$ D-xylopyranoside units, which can occur substituted with α -L-arabinofuranoside residues positioned on C-(O)-2 or C-(O)-3 [14,20]. In addition, xylose monomers may be substituted with glucuronic acid and its 4-O-methyl derivative. On the other hand, pectin is mainly composed of linear chains of α -1,4-D-galacturonic acid (GalA) called homogalacturonan. Moreover, ramified domains of pectin comprise alternate sequences of GalA and α -(1, 2) linked α -L-rhamnosyl residues, which may be substituted at O-4 with linear or branched oligosaccharides [14,21]. Therefore, microbial enzymes acting on these polysaccharide structures involved the following activities: arabinofuranosidases, xylosidases, xylanases, xylan glucuronidases and oxidases, acetyl xylan esterases and pectin esterases, polygalacturonases, rhamnosidases, and polygalacturonate and rhamnogalacturonan lyases. As an approach to aid design prebiotics based on arabinoxylan or pectin, targeted to ameliorate IBD gut dysbiosis, microbial enzymes involving these domains should be investigated.

Therefore, the aim of this study was to employ several computational methods to reconstruct *in silico* carbohydrate metabolic capabilities from MAGs recovered from the microbiota of both healthy and CD individuals. Glycosidase profiles and potential metabolic activity of gut microbes from healthy and CD microbiota have been compared. This bioinformatics approach was developed as a mean to aid rationally designed prebiotic interventions to rebalance the CD dysbiosis, with a focus on the metabolism of emergent prebiotics derived from arabinoxylan and pectin.

2. Results and Discussion

2.1. MAGs Recovery

Metabolic capabilities of gut microbes present in the microbiota of patients with CD and healthy individuals have been studied. Specifically, glycolytic activities of gut microbes and their synergistic interactions in the presence of pectin and arabinoxylan have been investigated *in silico*. Figure 1 illustrates the computational workflow used in this work. A total 1196 and 1577 MAGs were first recovered from CD and healthy fecal samples, respectively (Tables 1 and 2, Supplementary material Tables S1 and S2). As it can be seen, some of these MAGs could be identified at strain or species level while other sequences could be correctly identified only at genus or family level. Interestingly, several microbial clades (total number of different taxa $n = 41$) were also found in the microbiota of both patients with CD and healthy individuals (Tables 1 and 2). In this regard, these common taxa include health-promoting genera such as *Faecalibacterium*, *Akkermansia*, *Blautia* and *Paraprevotella* [14,22].

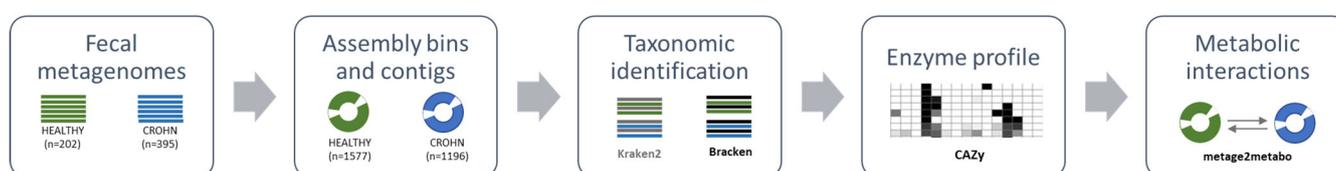


Figure 1. Computational workflow used in this work. Metagenome-assembled genomes (MAGs) were recovered from the microbiota of patients with Crohn’s disease and healthy individuals. Taxonomic identification of MAGs was performed using Kraken2 and Bracken software. MAGs sequences were mapped against the Carbohydrate-Active enZYMes Database (CAZy) to annotate glycosidase domains. Finally, metabolic interactions between MAGs were elucidated using metage2metabo software (v. 1.5.0) developed by Belcour et al. [8].

Table 1. Number of metagenome-assembled genomes (MAGs, $n = 998$) recovered from the microbiota of patients with Crohn’s disease and assigned to taxonomic clades that were also found in the microbiota of healthy individuals. MAGs were identified at family, genus, species or strain level. It should be noted that these common clades (total number of different taxa $n = 41$) were identical to those shown in Table 2.

MAGs Found in the Microbiota of Patients with Crohn’s Disease That were also Found in the Microbiota of Healthy Individuals	
Taxa	<i>n</i>
Lachnospiraceae	174
<i>Alistipes</i>	113
<i>Faecalibacterium prausnitzii</i>	100
<i>Ruminococcus</i>	96
<i>Sutterella</i>	67
<i>Prevotella</i>	61
<i>[Eubacterium] eligens</i> ATCC 27750	41
<i>Oscillibacter valericigenes</i> Sjm18–20	39
<i>Bacteroides</i>	36
<i>Akkermansia muciniphila</i>	35
<i>Escherichia coli</i>	33
<i>Barnesiella viscericola</i> DSM 18177	25
<i>Phascolarctobacterium</i>	23
<i>Acidaminococcus intestini</i> RyC-MR95	22
<i>Blautia</i>	17
Ruminococcaceae	16
Burkholderiales bacterium YL45	13
<i>Dialister</i>	13

Table 1. Cont.

MAGs Found in the Microbiota of Patients with Crohn's Disease That were also Found in the Microbiota of Healthy Individuals	
Taxa	n
<i>Flavonifractor plautii</i>	11
<i>Lachnoclostridium</i>	10
<i>Megasphaera</i>	6
<i>Paraprevotella xylaniphila</i> YIT 11841	6
<i>Anaerostipes hadrus</i>	5
<i>Odoribacter splanchnicus</i>	5
<i>Anaerotignum propionicum</i> DSM 1682	4
<i>Desulfovibrio vulgaris</i> str. Miyazaki F	4
Clostridiales bacterium CCNA10	3
<i>Bifidobacterium longum</i> subsp. <i>longum</i>	2
<i>Bifidobacterium pseudocatenulatum</i> DSM 20438 = JCM 1200 = LMG 10505	2
<i>Butyrivibrio</i>	2
<i>Haemophilus parainfluenzae</i> T3T1	2
<i>Megamonas funiformis</i>	2
<i>Ruthenibacterium lactatiformans</i>	2
<i>Collinsella aerofaciens</i>	1
<i>Enterococcus faecium</i> Com15	1
Erysipelotrichaceae	1
<i>Fusobacterium mortiferum</i> ATCC 9817	1
<i>Intestinibaculum porci</i>	1
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	1
<i>Klebsiella quasipneumoniae</i> subsp. <i>quasipneumoniae</i>	1
<i>Longibaculum</i> sp. KGMB06250	1
Total	998

Table 2. Number of metagenome-assembled genomes (MAGs, $n = 1361$) recovered from the microbiota of healthy individuals and assigned to taxonomic clades that were also found in the microbiota of patients with Crohn's disease. MAGs were identified at family, genus, species or strain level. It should be noted that these common clades (total number of different taxa $n = 41$) were identical to those shown in Table 1.

MAGs Found in the Microbiota of Healthy Individuals That Were also Found in the Microbiota of Patients with Crohn's Disease	
Taxa	n
Lachnospiraceae	198
<i>Ruminococcus</i>	196
<i>Faecalibacterium prausnitzii</i>	111
Ruminococcaceae	90
<i>Prevotella</i>	84
<i>[Eubacterium] eligens</i> ATCC 27750	79
<i>Phascolarctobacterium</i>	70
<i>Sutterella</i>	64
<i>Alistipes</i>	63
<i>Oscillibacter valericigenes</i> Sjm18–20	36
<i>Bacteroides</i>	34
<i>Barnesiella viscericola</i> DSM 18177	31
Burkholderiales bacterium YL45	28
<i>Escherichia coli</i>	27
<i>Akkermansia muciniphila</i>	26
<i>Collinsella aerofaciens</i>	20
<i>Blautia</i>	17
<i>Dialister</i>	16
<i>Bifidobacterium longum</i> subsp. <i>longum</i>	15

Table 2. Cont.

MAGs Found in the Microbiota of Healthy Individuals That Were also Found in the Microbiota of Patients with Crohn's Disease	
Taxa	n
<i>Megamonas funiformis</i>	15
<i>Megasphaera</i>	15
<i>Odoribacter splanchnicus</i>	15
<i>Bifidobacterium pseudocatenulatum</i> DSM 20438 = JCM 1200 = LMG 10505	13
<i>Lachnospiraceae</i>	12
<i>Paraprevotella xyliniphila</i> YIT 11841	12
<i>Anaerostipes hadrus</i>	11
<i>Desulfovibrio vulgaris</i> str. Miyazaki F	10
<i>Anaerostipes propionicum</i> DSM 1682	9
<i>Flavonifractor plautii</i>	7
<i>Klebsiella pneumoniae</i> subsp. <i>pneumoniae</i>	6
<i>Erysipelotrichaceae</i>	5
<i>Longibaculum</i> sp. KGMB06250	5
<i>Haemophilus parainfluenzae</i> T3T1	4
<i>Acidaminococcus intestini</i> RyC-MR95	3
<i>Butyrivibrio</i>	3
<i>Fusobacterium mortiferum</i> ATCC 9817	3
<i>Intestinibaculum porci</i>	3
<i>Klebsiella quasipneumoniae</i> subsp. <i>quasipneumoniae</i>	2
Clostridiales bacterium CCNA10	1
<i>Enterococcus faecium</i> Com15	1
<i>Ruthenibacterium lactatiformans</i>	1
Total	1361

MAGs from a few clades were exclusively recovered from one of the studied groups. MAGs from the clades *Parabacteroides distasonis*, *Roseburia intestinalis* L1-82, *Veillonella*, *Clostridium bolteae* and *Muribaculum* sp. TLL-A4 and other gut microbes were exclusively retrieved from IBD metagenomes (total number of different taxa $n = 25$, Supplementary material Table S1), whereas MAGs from unidentified species of *Parabacteroides*, *Clostridium*, *Roseburia*, *Muribaculum*, as well as *B. adolescentis* ATCC 15703 and other species were only retrieved from healthy metagenomes (total number of different taxa $n = 42$, Supplementary material Table S2).

As explained, several bacterial genera were found in both IBD and healthy metagenomes. These MAGs belong to common host species from the human colon. It has been reported that Lachnospiraceae were not significantly increased in patients with IBD [23]. In the present study, a similar number of MAGs of unidentified members of Lachnospiraceae and other genera belonging to this family such as *Blautia* were recovered from IBD patients and healthy participants. This microbial family belongs to the healthy core gut microbiota and encompasses the main producers of short-chain fatty acids (SCFAs). In addition, IBD typically involves a reduction in microbial diversity and the lower abundance of Firmicutes [1]. A higher number of MAGs of Firmicutes such as *Ruminococcus*, were recovered from healthy participants than from IBD patients (Tables 1 and 2). On the other hand, MAGs of gut microbes affected by CD dysbiosis, according to the literature, were obtained from both types of samples (healthy participants and CD patients) (Tables 1 and 2). In this regard, it has been described that IBD results in a decrease of bacteria with anti-inflammatory capacities such as *F. prausnitzii* when compared to healthy participants [1]. This species belonging to *Clostridium* cluster IV induces the proliferation of regulatory T cells associated with IBD. *F. prausnitzii* also exerts an anti-inflammatory effect by producing butyrate and microbial anti-inflammatory molecules [15,24]. A decrease in the concentration of SCFAs produced by *F. prausnitzii* and *Clostridium* clusters IV, XIVa and XVIII has been found in IBD patients [25]. Similarly, several species belonging to *Blautia*, *Clostridium*, *Roseburia* and *Ruminococcus* are decreased in patients with CD [1], including

children, adolescents and adults [1,23]. On the contrary, increases in Proteobacteria, mainly *E. coli*, and Bacteroidetes have been reported in patients with CD when compared to healthy subjects [1,26,27].

Some microbial clades identified in IBD and healthy microbiota such as *A. muciniphila* and *F. prausnitzii* comprise specialist primary degrading gut anaerobes involved in the breakdown of dietary fiber [28], leading to SCFAs production and health benefits [15]. Therefore, it is of great interest to study the carbohydrate metabolism of these microorganisms in order to design IBD targeting prebiotics.

2.2. Study of MAGs Glycosidases Acting on Arabinoxylan and Pectin

In order to deepen the carbohydrate metabolism potential of MAGs from gut microbes retrieved from CD and healthy individual's metagenomes, glycosidase analysis was performed. Specifically, characteristic profiles of glycosidases acting on arabinoxylan and pectin chains were investigated (Figures 2–4, Supplementary material Tables S3–S6, Supplementary material Figures S1 and S2). A schematic representation of enzyme domains acting on arabinoxylan and pectin is provided in Figure 2. As an approach to aid design prebiotics based on arabinoxylan or pectin, targeted to ameliorate IBD gut dysbiosis, first the enzyme machinery of the MAGs recovered from both IBD and healthy gut metagenomes to hydrolyse these carbohydrates was investigated.

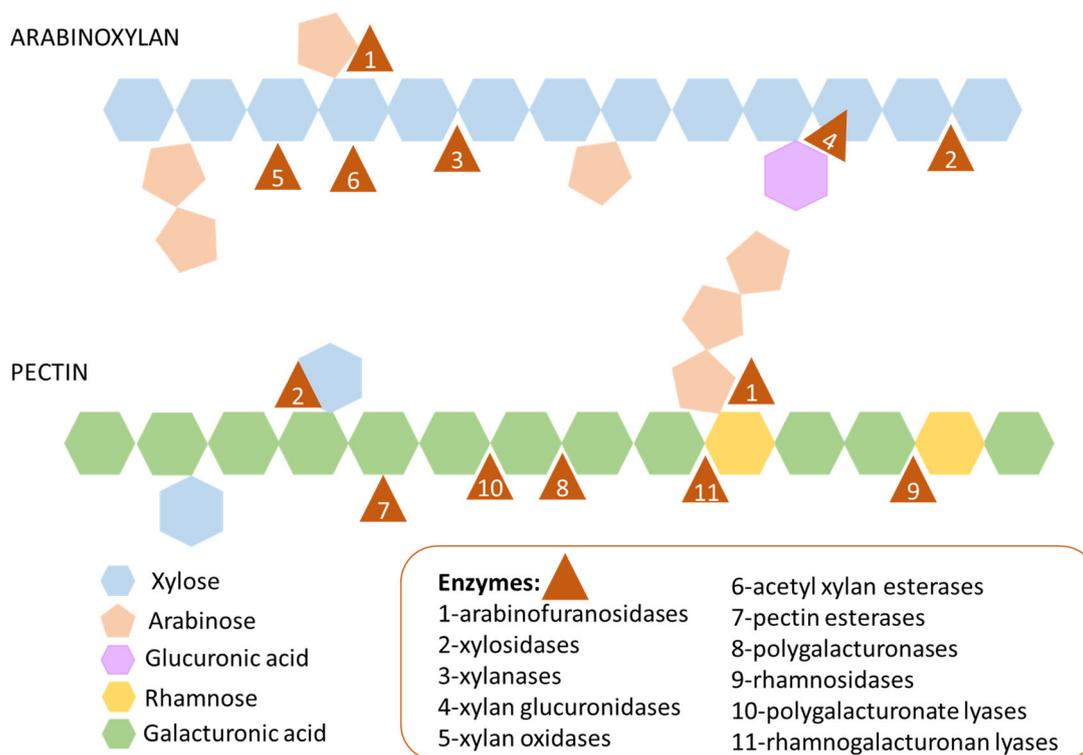


Figure 2. Schematic representation of enzyme domains acting on arabinoxylan and pectin that were annotated in metagenome-assembled genomes (MAGs) recovered from the microbiota of patients with Crohn's disease and healthy individuals.

Arabinoxylan degrading enzymes. Arabinoxylan degrading enzymes were annotated in several MAGs of a wide range of genera present in both IBD and healthy metagenomes including *Anaerostipes*, *Bacteroides*, *Bifidobacterium*, *Blautia*, *Butyrivibrio*, *Faecalibacterium*, unidentified members of Lachnospiraceae, *Prevotella*, *Paraprevotella* and *Ruminococcus* (Figures 3 and 4, Supplementary material Tables S3 and S4). The potential of several of these genera to metabolise emerging prebiotics was reported in a previous study [14]. With regard to CD samples, xylanase domains were identified in all MAGs of *Butyrivibrio* and *B. viscericola* DSM 18177, as well as in most MAGs of other gut microbes such as *P. xylaniphila* YIT 11841 (Figure 3, Supplementary material Table S3). Similar results were obtained for healthy microbiota samples (Figure 4, Supplementary material Table S4). Xylosidase domains were annotated in all MAGs of *B. pseudocatenulatum* DSM 20438 recovered from healthy microbiota samples, while xylan glucuronidases were identified in the majority of MAGs of *P. xylaniphila* YIT 11841 recovered from IBD and healthy metagenomes (Figures 3 and 4, Supplementary material Tables S3 and S4). Other functional domains of interest involve acetyl xylan esterases identified in most MAGs of *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841 recovered from IBD metagenomes (Figure 3, Supplementary material Table S3). These domains were also annotated in several MAGs recovered from healthy metagenomes (Figure 4, Supplementary material Table S4). In contrast, xylan oxidase domains were only found in MAGs of *E. faecium* Com15 recovered from healthy metagenomes (Figure 4, Supplementary material Table S4).

Arabinoxylan and pectin degrading enzymes. Enzyme domains involving both xylanases and arabinofuranosidases were present in the majority of MAGs of several clades including *Bacteroides*, Lachnospiraceae, *Ruminococcus*, *B. viscericola* DSM 18177, *O. splanchnicus* and *P. xylaniphila* YIT 11841 recovered from IBD and healthy metagenomes (Figures 3 and 4, Supplementary material Tables S3 and S4). It has been reported that arabinose residues are metabolised by a limited amount of gut microbes containing arabinofuranoside activities. In addition, few microbial species show full xylanolytic capacity so arabinoxylan are first hydrolysed by first degraders such as specific *Bacteroides* species. Hydrolysis products are then utilised by other bacteria that cannot grow on xylan, leading to cross-feeding interactions [14,29]. Unspecific domains comprising arabinofuranosidases, β -galactosidases and β -xylosidases were identified in most MAGs of *B. viscericola* DSM 18177 recovered from patients with CD (Figure 3, Supplementary material Table S3). Other domains involving multiple activities of interest such as acetyl xylan esterase and pectin acetylerase were mainly found in MAGs of *P. xylaniphila* YIT 11841 recovered from IBD and healthy metagenomes (Figures 3 and 4, Supplementary material Tables S3 and S4).

Pectin degrading enzymes. With regard to functional profiles of glycosidases acting on pectin, rhamnosidases were annotated in the majority of MAGs of *A. muciniphila*, *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841 from IBD and healthy participants (Figures 3 and 4, Supplementary material Tables S3 and S4). In addition, rhamnogalacturonan lyases were found in most MAGs of *P. xylaniphila* YIT 11841 recovered from IBD and healthy metagenomes (Figures 3 and 4, Supplementary material Tables S3 and S4).

Some glycosidases were identified only in MAGs recovered from metagenomes from patients with CD. It should be noted that these glycosidases were annotated in a limited number of MAGs belonging to the same clade, highlighting the role of metabolic variability between strains (Supplementary material Figure S1 and Table S5). On the other hand, several functional domains were found only in MAGs recovered from healthy individuals comprising starch-binding enzymes and cellulases (Supplementary material Figure S2 and Table S6). These activities were similar to those only identified in IBD metagenomes (Supplementary material Figure S1 and Table S5).

A wide number of characteristic glycosidase domains acting on arabinoxylan and pectin were annotated in MAGs of *A. muciniphila*, *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841 recovered from metagenomic samples of patients with CD and healthy individuals. These unique enzymatic activities involved xylan glucuronidases, acetyl xylan esterases, pectin acetylerases, rhamnosidases and rhamnogalacturonan lyases that were

not found in MAGs of other gut microbes. These results suggest that metabolic activities of these species might not be greatly affected by gut dysbiosis in CD. Therefore, the rational design of prebiotics derived from arabinoxylan and pectin could be of special interest to ameliorate CD symptoms through its metabolism by these core species. Previous studies suggest that administration of arabinoxylan oligosaccharides and high methoxylated pectin increased the abundance of *A. muciniphila* [30]. Similarly, prebiotic administration of a highly branched rhamnogalacturonan type I-enriched pectin increased the abundance of *Barnesiella* and other beneficial genera such as *Butyrivibrio*, *Roseburia*, *Flavonifractor*, *Acetivibrio* and *Clostridium* cluster IV [31]. Moreover, the ability of *Paraprevotella* genus to metabolise xylan and pectin leading to SCFAs production has been described [22].

2.3. Cross-Feeding between MAGs in the Presence of Arabinoxylan and Pectin

To gain a better understanding on the ecological impact of MAGs showing glycosidase domains acting on arabinoxylan and pectin to sustain gut microbiota ecosystem networks, potential cross-feeding mechanisms between these bacteria were elucidated *in silico* according to Belcour et al. [8]. The number of genes associated to different microbial metabolic activities annotated in MAGs ranged from 361 to 2074 on IBD samples, and from 315 to 2080 on samples from healthy individuals. This parameter was calculated according to Belcour et al. [8]. To illustrate potential cross-feeding mechanisms between different taxa of interest, arabinoxylan and pectin degraders (*A. muciniphila*, *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841) found in the microbiota of both healthy participants and patients with CD, were selected. These species and strains showed a wide range of characteristic enzyme activities acting on arabinoxylan and pectin (i.e., xylan glucuronidases and xylan esterases, pectin acetyl esterases, rhamnosidases and rhamnogalacturonan lyases) that were not found in other gut microbes (see Results and Discussion Section 2.2.). These three clades have been associated to several health benefits including SCFAs production [22,30,31]. On the other hand, MAGs of other microorganisms that were identified only in the microbiota of healthy individuals (i.e. these MAGs were not recovered from the microbiota of patients with CD, Supplementary material Table S2) were included in the metabolic interaction study (Figure 5). These clades might be greatly affected by gut dysbiosis. To compute the metabolic network, one MAG per clade showing the highest number of genes associated to metabolite production was chosen.

Several cross-feeding mechanisms were observed among pectin and arabinoxylan degraders (*A. muciniphila*, *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841) and those clades affected by gut dysbiosis (i.e., those that could only be recovered from healthy individuals) (Figure 5). Similar metabolic interactions were determined in the presence of pectin and arabinoxylan (Figure 5, Supplementary material Figure S3). *A. muciniphila* and *B. viscericola* DSM 18177 showed equivalent metabolic functions and potential synergistic relationships with *F. cylindroides* T2-87. On the other hand, cross-feeding mechanisms between *P. xylaniphila* YIT 11841 and *C. difficile* or *V. atypica* were elucidated. Interestingly, *F. cylindroides* T2-87, *C. difficile* and *V. atypica* showed potential cross-feeding mechanisms with the rest of gut microbes affected by gut dysbiosis (Figure 5).

Taking into account these limitations, computational models described in the present work allow predicting the number of microbial taxa involved in gas formation, although gas concentrations cannot be predicted by this pipeline. In this regard, no major differences were observed in the number of taxa involved in hydrogen production between microbial cross-feeding simulations in the presence of pectin, arabinoxylan and without prebiotics. It was found that up to 337 MAGs recovered from CD patients corresponding to 51 different taxa might contribute to hydrogen production. It should be considered that the amount of gas produced cannot be calculated by the models. Therefore, the suitability of pectin and arabinoxylan-derived compounds as prebiotics targeting at CD should be validated in future clinical studies.

3. Materials and Methods

3.1. Metagenome Selection

To investigate arabinoxylan and pectin metabolism by gut microbiota in patients with CD and healthy individuals, curated human metagenomes were retrieved using the standardised database HumanMetagenomeDB [34]. Unassembled reads produced by an Illumina platform showing a minimum sequencing depth of 10 million reads were selected. As a result, CD fecal metagenomes ($n = 395$) from three sequencing experiments [35–37] were retrieved: BioProject identity (ID) 46321 (sequence library IDs SRR495449–SRR497947), BioProject ID 321058 (sequence library IDs SRR3582131–SRR3582183) and BioProject ID 398089 (sequence library IDs SRR5935744–SRR5962905). These samples corresponded to fecal metagenomes from male and female participants from the United States of America belonging to the following age groups: children, teenagers, adults and elders. On the other hand, healthy metagenomes ($n = 202$) used as controls were retrieved from nine different studies and reference datasets [38–46]: BioProject ID 266076 (sequence library IDs ERR478964–ERR479598), BioProject ID 305507 (sequence library IDs SRR2992882–SRR2992961), BioProject ID 356544 (sequence library IDs SRR5088933–SRR5088943), BioProject ID 382085 (sequence library IDs ERR1912957–ERR1913126), BioProject ID 356102 (sequence library IDs SRR5091454–SRR5091619), BioProject ID 388263 (sequence library IDs SRR5813226–SRR5813545), BioProject ID 397664 (sequence library IDs SRR5925338–SRR5925343), BioProject ID 492716 (sequence library IDs SRR8113241–SRR8113264) and BioProject ID 520750 (sequence library IDs DRR127488–DRR162776). These samples corresponded to fecal metagenomes from male and female participants from several countries (Australia, Canada, China, France, Germany, United States of America and Japan) belonging to the adult and elder age groups.

3.2. Metagenome Assembly

MAGs were recovered according to the method described by Sabater et al. [9], using validated workflows for metagenome assembly and taxonomic classification [5]. These computational pipelines were used to process metagenomes from patients with CD and healthy individuals to further compare their metabolic activities (Figure 1).

Contaminant reads and low-quality sequences were separated *in silico* from microbial reads using Kneaddata (v0.7.4) and Trimmomatic (v0.39) software. For this purpose, minimum length of output reads was computed as 50 percent of the length of the input reads considering a sliding window of 4:20. Bowtie2 (v.2.4.2) was used to map metagenomic reads [47], against the reference databases Homo sapiens hg37 and human contamination Bowtie2 (v.2.3.5.1) in order to remove host contamination. Metagenome sequences were assembled using MEGAHIT v.1.2.9 software with default settings [48]. Maximum k-mer size was set at 127 in order to generate the following series of k-mers with lengths shorter than entire reads (k-21, k-31, k-41, k-51, k-61, k-71, k-81, k-91, k-101, k-111, k-121, k-127). Metagenome reads were mapped against the assembly using Bowtie2 [47]. Output bam files generated were sorted and indexed. Metagenome binning of contigs larger than 1.5 kilobases was performed using MetaBAT2 v.2.2.15 software with default settings [49]. Once MAGs were obtained, CheckM v.1.1.3 lineage-specific workflow was run to assess

their completeness and contamination [50]. Quality standards of MAGs were established according to Asnicar et al. [5]. MAGs showing completeness lower than 50% and contamination higher than 5% were discarded. Taxonomic classification of pruned MAGs at family, species genus and strain level, was performed following Kraken2 standard workflow [51]. Then, these taxonomic assignments were curated using Bracken software (Bayesian Reestimation of Abundance with Kraken) [52]. This software uses the taxonomic assignments made by Kraken2, along with information about the genomes themselves to accurately classify sequences even when a sample contains multiple near-identical species.

3.3. *In Silico Study of Arabinoxylan and Pectin Metabolism*

Metabolic capacity of MAGs obtained from patients with CD and healthy individuals to degrade arabinoxylan and pectin was assessed through comparative genomics. First, glycosidase functional domains of MAGs were annotated following “run_dbcan” pipeline developed by Zhang et al. [53]. To this aim, MAGs sequences were mapped against the Carbohydrate-Active enZYmes Database (CAZy, <http://www.cazy.org/last> accessed on 23 February 2022). To identify functional domains of interest (i.e., glycosidase domains acting on arabinoxylan and pectin chains), HMMER software for biosequence analysis [54], and Pfam database [55], were used. Coverage values higher than 0.95 were considered to provide high-quality annotated glycosidase sequences. Metabolic capabilities of MAGs were compared through hierarchical clustering calculated by the complete linkage method. In the complete linkage method, all pairwise dissimilarities between the elements in each cluster (i.e., glycosidase domains of MAGs acting on arabinoxylan and pectin) are calculated using the basic function “hclust” from R (v.3.6.2). Then, the largest dissimilarity value is chosen as the distance between clusters to yield more compact clusters.

Once complete glycosidase profiles of MAGs were elucidated, potential symbiotic relationships between MAGs from patients with CD and healthy individuals were studied *in silico*. For this purpose, MAGs were annotated using Prokka pipeline [56]. Then, standard Genbank files (.gbk) files containing sequences and annotations obtained by Prokka were selected to elucidate potential cross-feeding interaction mechanisms using metage2metabo software [8]. A set of nutrients (seeds) that may be present in human gut according to Belcour et al. [8], was provided to the algorithm in order to calculate potential metabolic networks in given nutritional conditions. Different simulations were performed in the presence of arabinoxylan and pectin.

4. Conclusions

A total of 1196 and 1577 MAGs have been recovered from CD and healthy fecal samples, respectively. Glycosidase profiles of MAGs have been annotated and functional domains acting on pectin and arabinoxylan have been identified. A broad range of enzymes acting on arabinoxylan and pectin were annotated in MAGs of *A. muciniphila*, *B. viscericola* DSM 18177 and *P. xylaniphila* YIT 11841 recovered from metagenome sequencing data of patients with CD and healthy individuals. In addition, potential cross-feeding mechanisms between these taxa and other gut microbes not found in CD in the presence of arabinoxylan and pectin have been elucidated. As a result, we propose that arabinoxylan and pectin administration may ameliorate gut dysbiosis in CD by promoting taxa through cross-feeding network interactions with key commensal species. Computational simulations herein presented provide a first insight into the prediction of complex microbial interactions within the gut ecosystem and how these can guide the rational design of prebiotic/synbiotic preparations targeting CD dysbiosis.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms23137093/s1>.

Author Contributions: Conceptualization, C.S., I.C.-T., L.R. and A.M.; methodology, C.S.; software, C.S., formal analysis; C.S., investigation, C.S., I.C.-T., L.R. and A.M.; writing—original draft preparation, C.S., writing—review and editing, C.S., I.C.-T., L.R. and A.M.; supervision, L.R. and A.M.;

funding acquisition; L.R. and A.M. All authors have read and agreed to the published version of the manuscript.

Funding: The work in our research groups was funded by the grants RTI2018-095021-J-I00 (funded by (MCIU/AEI/FEDER, UE) and AGL2016-78311-R (funded by (MINECO/AEI/FEDER, UE).

Acknowledgments: Carlos Sabater acknowledges his Postdoctoral research contract funded by the Instituto de Investigación Sanitaria del Principado de Asturias (ISPA) and Postdoctoral research contract Juan de la Cierva-Formación from Spanish Ministry of Science and Innovation (FJC2019-042125-I).

Conflicts of Interest: The authors declare no conflict of interest. Abelardo Margolles is co-founder and member of the scientific board of Microviable Therapeutics SL.

References

- Nishida, A.; Inoue, R.; Inatomi, O.; Bamba, S.; Naito, Y.; Andoh, A. Gut microbiota in the pathogenesis of inflammatory bowel disease. *Clin. J. Gastroenterol.* **2018**, *11*, 1–10. [\[CrossRef\]](#)
- Zhang, X.F.; Guan, X.X.; Tang, Y.J.; Sun, J.F.; Wang, X.K.; Wang, W.D.; Fan, J.M. Clinical effects and gut microbiota changes of using probiotics, prebiotics or synbiotics in inflammatory bowel disease: A systematic review and meta-analysis. *Eur. J. Nutr.* **2021**, *60*, 2855–2875. [\[CrossRef\]](#)
- Jairath, V.; Feagan, B.G. Global burden of inflammatory bowel disease. *Lancet Gastroenterol. Hepatol.* **2020**, *5*, 2–3. [\[CrossRef\]](#)
- Alshehri, D.; Saadah, O.; Mosli, M.; Edris, S.; Alhindi, R.; Bahieldin, A. Dysbiosis of gut microbiota in inflammatory bowel disease: Current therapies and potential for microbiota-modulating therapeutic approaches. *Bosn. J. Basic Med. Sci.* **2021**, *21*, 270. [\[CrossRef\]](#)
- Asnicar, F.; Berry, S.E.; Valdes, A.M.; Nguyen, L.H.; Piccinno, G.; Drew, D.A.; Leeming, E.; Gibson, R.; Le Roy, C.; Al Khatib, H.; et al. Microbiome connections with host metabolism and habitual diet from 1098 deeply phenotyped individuals. *Nat. Med.* **2021**, *27*, 321–332. [\[CrossRef\]](#)
- Bickhart, D.M.; Kolmogorov, M.; Tseng, E.; Portik, D.M.; Korobeynikov, A.; Tolstoganov, I.; Uritskiy, G.; Liachko, I.; Sullivan, S.T.; Shin, S.B.; et al. Generating lineage-resolved, complete metagenome-assembled genomes from complex microbial communities. *Nat. Biotechnol.* **2022**, *40*, 711–719. [\[CrossRef\]](#)
- Pasolli, E.; Asnicar, F.; Manara, S.; Zolfo, M.; Karcher, N.; Armanini, F.; Beghini, F.; Manghi, P.; Tett, A.; Ghensi, P.; et al. Extensive unexplored human microbiome diversity revealed by over 150,000 genomes from metagenomes spanning age, geography, and lifestyle. *Cell* **2019**, *176*, 649–662. [\[CrossRef\]](#)
- Belcour, A.; Frioux, C.; Aite, M.; Bretaudeau, A.; Hildebrand, F.; Siegel, A. Metage2Metabo, microbiota-scale metabolic complementarity for the identification of key species. *Elife* **2020**, *9*, e61968. [\[CrossRef\]](#)
- Sabater, C.; Ruiz, L.; Margolles, A. A Machine Learning Approach to Study Glycosidase Activities from Bifidobacterium. *Microorganisms* **2021**, *9*, 1034. [\[CrossRef\]](#)
- Zimmermann, J.; Kaleta, C.; Waschina, S. gapseq: Informed prediction of bacterial metabolic pathways and reconstruction of accurate metabolic models. *Genome Biol.* **2021**, *22*, 81. [\[CrossRef\]](#)
- Heinken, A.; Ravcheev, D.A.; Baldini, F.; Heirendt, L.; Fleming, R.M.; Thiele, I. Systematic assessment of secondary bile acid metabolism in gut microbes reveals distinct metabolic capabilities in inflammatory bowel disease. *Microbiome* **2019**, *7*, 75. [\[CrossRef\]](#)
- Sugihara, K.; Kamada, N. Diet–Microbiota Interactions in Inflammatory Bowel Disease. *Nutrients* **2021**, *13*, 1533. [\[CrossRef\]](#)
- Statovci, D.; Aguilera, M.; MacSharry, J.; Melgar, S. The impact of western diet and nutrients on the microbiota and immune response at mucosal interfaces. *Front. Immunol.* **2017**, *8*, 838. [\[CrossRef\]](#)
- Sabater, C.; Calvete-Torre, I.; Villamiel, M.; Moreno, F.J.; Margolles, A.; Ruiz, L. Vegetable waste and by-products to feed a healthy gut microbiota: Current evidence, machine learning and computational tools to design novel microbiome-targeted foods. *Trends Food Sci. Technol.* **2021**, *118*, 399–417. [\[CrossRef\]](#)
- Lin, T.L.; Shu, C.C.; Lai, W.F.; Tzeng, C.M.; Lai, H.C.; Lu, C.C. Investiture of next generation probiotics on amelioration of diseases—Strains do matter. *Med. Microecol.* **2019**, *1*, 100002. [\[CrossRef\]](#)
- Lordan, C.; Thapa, D.; Ross, R.P.; Cotter, P.D. Potential for enriching next-generation health-promoting gut bacteria through prebiotics and other dietary components. *Gut Microbes* **2020**, *11*, 1–20. [\[CrossRef\]](#)
- Hadji, H.; Bouchemal, K. Advances in the treatment of inflammatory bowel disease: Focus on polysaccharide nanoparticulate drug delivery systems. *Adv. Drug Deliv. Rev.* **2022**, *181*, 114101. [\[CrossRef\]](#)
- Sabater, C.; Molina-Tijeras, J.A.; Vezza, T.; Corzo, N.; Montilla, A.; Utrilla, P. Intestinal anti-inflammatory effects of artichoke pectin and modified pectin fractions in the dextran sulfate sodium model of mice colitis. Artificial neural network modelling of inflammatory markers. *Food Funct.* **2019**, *10*, 7793–7805. [\[CrossRef\]](#)
- Shah, A.; Masoodi, F.A.; Gani, A.; Gani, A.; Noor, N.; Fazli, A. Arabinoxylans. In *Food Biopolymers: Structural, Functional and Nutraceutical Properties*; Gani, A., Ashwar, B.A., Eds.; Springer: Cham, Switzerland, 2021; pp. 173–186. [\[CrossRef\]](#)

20. Mathew, S.; Aronsson, A.; Karlsson, E.N.; Adlercreutz, P. Xylo- and arabinoxylooligosaccharides from wheat bran by endoxylanases, utilisation by probiotic bacteria, and structural studies of the enzymes. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 3105–3120. [[CrossRef](#)]
21. Tan, H.; Nie, S. Deciphering diet-gut microbiota-host interplay: Investigations of pectin. *Trends Food Sci. Technol.* **2020**, *106*, 171–181. [[CrossRef](#)]
22. Li, Y.; Yang, H.; Xu, L.; Wang, Z.; Zhao, Y.; Chen, X. Effects of dietary fiber levels on cecal microbiota composition in geese. *Asian Australas. J. Anim. Sci.* **2018**, *31*, 1285. [[CrossRef](#)]
23. Vacca, M.; Celano, G.; Calabrese, F.M.; Portincasa, P.; Gobetti, M.; De Angelis, M. The controversial role of human gut lachnospiraceae. *Microorganisms* **2020**, *8*, 573. [[CrossRef](#)]
24. Breyner, N.M.; Michon, C.; de Sousa, C.S.; Vilas Boas, P.B.; Chain, F.; Azevedo, V.A.; Langella, P.; Chatel, J.M. Microbial anti-inflammatory molecule (MAM) from *Faecalibacterium prausnitzii* shows a protective effect on DNBS and DSS-induced colitis model in mice through inhibition of NF- κ B pathway. *Front. Microbiol.* **2017**, *8*, 114. [[CrossRef](#)]
25. Takahashi, K.; Nishida, A.; Fujimoto, T.; Fujii, M.; Shioya, M.; Imaeda, H.; Inatomi, O.; Bamba, S.; Andoh, A.; Sugimoto, M. Reduced abundance of butyrate-producing bacteria species in the fecal microbial community in CD. *Digestion* **2016**, *93*, 59–65. [[CrossRef](#)]
26. Frank, D.N.; Amand, A.L.S.; Feldman, R.A.; Boedeker, E.C.; Harpaz, N.; Pace, N.R. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 13780–13785. [[CrossRef](#)]
27. Manichanh, C.; Rigottier-Gois, L.; Bonnaud, E.; Gloux, K.; Pelletier, E.; Frangeul, L.; Nalin, R.; Jarrin, C.; Chardon, P.; Marteau, P.; et al. Reduced diversity of faecal microbiota in CD revealed by a metagenomic approach. *Gut* **2006**, *55*, 205–211. [[CrossRef](#)]
28. Duncan, S.H.; Russell, W.R.; Quartieri, A.; Rossi, M.; Parkhill, J.; Walker, A.W.; Flint, H.J. Wheat bran promotes enrichment within the human colonic microbiota of butyrate-producing bacteria that release ferulic acid. *Environ. Microbiol.* **2016**, *18*, 2214–2225. [[CrossRef](#)]
29. Zeybek, N.; Rastall, R.A.; Buyukkileci, A.O. Utilization of xylan-type polysaccharides in co-culture fermentations of *Bifidobacterium* and *Bacteroides* species. *Carbohydr. Polym.* **2020**, *236*, 116076. [[CrossRef](#)]
30. Zhai, Q.; Feng, S.; Arjan, N.; Chen, W. A next generation probiotic, *Akkermansia muciniphila*. *Crit. Rev. Food Sci. Nutr.* **2019**, *59*, 3227–3236. [[CrossRef](#)]
31. Zhu, K.; Mao, G.; Wu, D.; Yu, C.; Cheng, H.; Xiao, H.; Ye, X.; Linhardt, R.J.; Orfila, C.; Chen, S. Highly branched RG-I domain enrichment is indispensable for pectin mitigating against high-fat diet-induced obesity. *J. Agric. Food Chem.* **2020**, *68*, 8688–8701. [[CrossRef](#)]
32. Lindsay, J.O.; Whelan, K.; Stagg, A.J.; Gobin, P.; Al-Hassi, H.O.; Rayment, N.; Kamm, M.A.; Knight, S.C.; Forbes, A. Clinical, microbiological, and immunological effects of fructo-oligosaccharide in patients with Crohn's disease. *Gut* **2006**, *55*, 348–355. [[CrossRef](#)]
33. Wilson, B.; Whelan, K. Prebiotic inulin-type fructans and galacto-oligosaccharides: Definition, specificity, function, and application in gastrointestinal disorders. *J. Gastroenterol. Hepatol.* **2017**, *32*, 64–68. [[CrossRef](#)]
34. Kasmanas, J.C.; Bartholomäus, A.; Corrêa, F.B.; Tal, T.; Jehmlich, N.; Herberth, G.; von Bergen, M.; Stadler, P.F.; de Leon Ferreira de Carvalho, A.C.P.; da Rocha, U.N. HumanMetagenomeDB: A public repository of curated and standardized metadata for human metagenomes. *Nucleic Acids Res.* **2021**, *49*, D743–D750. [[CrossRef](#)]
35. Fraser-Liggett, C. Metagenomic Analysis of the Structure and Function of the Human Gut Microbiota in CD. *Nat. Preced.* **2010**, *1*. [[CrossRef](#)]
36. Lloyd-Price, J.; Arze, C.; Ananthakrishnan, A.N.; Schirmer, M.; Avila-Pacheco, J.; Poon, T.W.; Andrews, E.; Ajami, N.J.; Bonham, K.S.; Brislawn, C.J.; et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **2019**, *569*, 655–662. [[CrossRef](#)]
37. Vaughn, B.P.; Vatanen, T.; Allegretti, J.R.; Bai, A.; Xavier, R.J.; Korzenik, J.; Gevers, D.; Ting, A.; Robson, S.C.; Moss, A.C. Increased intestinal microbial diversity following fecal microbiota transplant for active CD. *Inflamm. Bowel Dis.* **2016**, *22*, 2182–2190. [[CrossRef](#)]
38. Bedarf, J.R.; Hildebrand, F.; Coelho, L.P.; Sunagawa, S.; Bahram, M.; Goeser, F.; Bork, P.; Wüllner, U. Functional implications of microbial and viral gut metagenome changes in early stage L-DOPA-naïve Parkinson's disease patients. *Genome Med.* **2017**, *9*, 38. [[CrossRef](#)]
39. Džunková, M.; Low, S.J.; Daly, J.N.; Deng, L.; Rinke, C.; Hugenholtz, P. Defining the human gut host-phage network through single-cell viral tagging. *Nat. Microbiol.* **2019**, *4*, 2192–2203. [[CrossRef](#)]
40. Fukuyama, J.; Rumker, L.; Sankaran, K.; Jeganathan, P.; Dethlefsen, L.; Relman, D.A.; Holmes, S.P. Multidomain analyses of a longitudinal human microbiome intestinal cleanout perturbation experiment. *PLoS Comput. Biol.* **2017**, *13*, e1005706. [[CrossRef](#)]
41. Hiseni, P.; Rudi, K.; Wilson, R.C.; Hegge, F.T.; Snipen, L. HumGut: A comprehensive human gut prokaryotic genomes collection filtered by metagenome data. *Microbiome* **2021**, *9*, 165. [[CrossRef](#)]
42. Minot, S.; Grunberg, S.; Wu, G.D.; Lewis, J.D.; Bushman, F.D. Hypervariable loci in the human gut virome. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3962–3966. [[CrossRef](#)]

43. Nanjundappa, R.H.; Ronchi, F.; Wang, J.; Clemente-Casares, X.; Yamanouchi, J.; Umeshappa, C.S.; Yang, Y.; Blanco, J.; Bassolas-Molina, H.; Salas, A.; et al. A gut microbial mimic that hijacks diabetogenic autoreactivity to suppress colitis. *Cell* **2017**, *171*, 655–667. [[CrossRef](#)]
44. Petersen, L.M.; Bautista, E.J.; Nguyen, H.; Hanson, B.M.; Chen, L.; Lek, S.H.; Sodergren, E.; Weinstock, G.M. Community characteristics of the gut microbiomes of competitive cyclists. *Microbiome* **2017**, *5*, 98. [[CrossRef](#)]
45. 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](#)]
46. 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](#)]
47. Langmead, B.; Wilks, C.; Antonescu, V.; Charles, R. Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* **2019**, *35*, 421–432. [[CrossRef](#)]
48. Li, D.; Luo, R.; Liu, C.M.; Leung, C.M.; Ting, H.F.; Sadakane, K.; Yamashita, H.; Lam, T.W. MEGAHIT v1.0: A fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* **2016**, *102*, 3–11. [[CrossRef](#)]
49. Kang, D.D.; Li, F.; Kirton, E.; Thomas, A.; Egan, R.; An, H.; Wang, Z. MetaBAT 2: An adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* **2019**, *7*, e7359. [[CrossRef](#)]
50. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)]
51. Wood, D.E.; Lu, J.; Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **2019**, *20*, 257. [[CrossRef](#)]
52. Lu, J.; Breitwieser, F.P.; Thielen, P.; Salzberg, S.L. Bracken: Estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **2017**, *3*, e104. [[CrossRef](#)]
53. Zhang, H.; Yohe, T.; Huang, L.; Entwistle, S.; Wu, P.; Yang, Z.; Busk, P.K.; Yin, Y. dbCAN2: A meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **2018**, *46*, W95–W101. [[CrossRef](#)]
54. Mistry, J.; Chuguransky, S.; Williams, L.; Qureshi, M.; Salazar, G.A.; Sonnhammer, E.L.; Tosatto, S.C.E.; Paladin, L.; Rai, S.; Richardson, L.J.; et al. Pfam: The protein families database in 2021. *Nucleic Acids Res.* **2021**, *49*, D412–D419. [[CrossRef](#)]
55. Mistry, J.; Finn, R.D.; Eddy, S.R.; Bateman, A.; Punta, M. Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* **2013**, *41*, e121. [[CrossRef](#)]
56. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **2014**, *30*, 2068–2069. [[CrossRef](#)]