fermentation

MDPI

Article

Exploring the Intestinal Microbial Community of Lantang Pigs through Metagenome-Assembled Genomes and Carbohydrate Degradation Genes

Jianbo Yang [1,2,†], Ying Fan [1,2,†], Rui Jin [1,2], Yunjuan Peng [1,2], Jianmin Chai [1,2], Xiaoyuan Wei [3], Yunxiang Zhao [4], Feilong Deng [1,2], Jiangchao Zhao [3,*] and Ying Li [1,2,*]

[1] Guangdong Provincial Key Laboratory of Animal Molecular Design and Precise Breeding, College of Life Science and Engineering, Foshan University, Foshan 528225, China; jianbo952@gmail.com (J.Y.); baicai727@outlook.com (Y.F.); 19514403220@163.com (R.J.); yunjuanpeng@163.com (Y.P.); jchai@uark.edu (J.C.); fdeng@fosu.edu.cn (F.D.)
[2] School of Life Science and Engineering, Foshan University, Foshan 528225, China
[3] Department of Animal Science, Division of Agriculture, University of Arkansas, Fayetteville, AR 72701, USA; xw010@uark.edu
[4] College of Animal Science & Technology, Guangxi University, Nanning 530004, China; yunxiangzhao@126.com
* Correspondence: jzhao77@uark.edu (J.Z.); yingli@fosu.edu.cn (Y.L.)
† These authors contributed equally to this work.

Abstract: High-fiber, low-cost agricultural byproducts offer a sustainable alternative for mitigating the competition for crops between humans and livestock. Pigs predominantly utilize dietary fibers through the process of microbial fermentation within the gut. This study explored the gut microbiota and the capacity for carbohydrate degradation in 30 individual Lantang pigs, a breed indigenous to China. Through metagenomic analysis, a total of 671 metagenome-assembled genomes (MAGs) were assembled and assigned into 14 bacterial and 1 archaeal phylum, including 97 species from uncultured microbes. The phylum with the highest abundance were identified as Bacillota_A, Bacteroidota, and Bacillota. Remarkably, the investigation revealed nearly 10,000 genes implicated in the degradation of carbohydrates, with a pronounced prevalence within five principal bacterial genera: *Prevotella*, *Cryptobacteroides*, *Gemmiger*, *Vescimonas*, and *Faecousia*. Additionally, 87 distinct types of carbohydrate-degrading enzymes were exclusively identified within the gut microbiota of the Lantang pig. These insights not only enhance our understanding of the microbial diversity specific to native Chinese pig breeds but also augment the body of research regarding porcine fiber degradation capabilities. The implications of this study are twofold: it provides strategic directions for optimizing feed efficiency and reducing breeding costs, and it furnishes an expanded gene pool for the microbial synthesis of industrial enzymes in the future.

Keywords: dietary fiber; enzymes; archaea; *Prevotella*; Chinese native pig

Citation: Yang, J.; Fan, Y.; Jin, R.; Peng, Y.; Chai, J.; Wei, X.; Zhao, Y.; Deng, F.; Zhao, J.; Li, Y. Exploring the Intestinal Microbial Community of Lantang Pigs through Metagenome-Assembled Genomes and Carbohydrate Degradation Genes. *Fermentation* **2024**, *10*, 207. https://doi.org/10.3390/fermentation10040207

Academic Editor: Alessandra Pino

Received: 24 February 2024
Revised: 3 April 2024
Accepted: 8 April 2024
Published: 11 April 2024

1. Introduction

China, being the foremost consumer of pork, harbors a diverse array of native pig breeds [1,2]. The Lantang pig, a breed indigenous to South China, is known by its excellent traits, such as strong adaptability to hot and humid environments, as well as a preference for a high-fiber diet [3]. Nonetheless, the majority of mammals, including Lantang pigs, are lacking in the enzymes necessary for digesting complex carbohydrates [4]. The degradation of dietary carbohydrates and host-derived polysaccharides primarily relies on the gut microbiota and the numerous carbohydrate-active enzymes (CAZymes) produced by the primary degraders [5]. CAZymes are naturally fermented by microbes—in short, fiber-degrading microbes that produce enzymes such as cellulases—which break down cellulose and other complex polysaccharides and indigestible plant fibers in the gut into smaller

sugar molecules that can be fermented by the colonic microbiota into short-chain fatty acids (SCFAs), primarily acetate, ropionate, and butyrate [6–8]. A recent study has discovered that the levels of SCFAs are notably higher in the colons of Jinhua pigs, which have a lower fat content [9]. These compounds play crucial roles in mammalian energy metabolism, intestinal physiology, and immune function [10,11].

Plant cell walls are composed of multiple complex biopolymers, including cellulose, hemicellulose, and lignin [12,13]. These substances impart toughness and resistance to plant cell walls. Breaking the chemical bonds and structures within plant cell walls requires the involvement of multiple glycoside hydrolase (GH) families [14]. Currently, these enzyme activities are classified into 187 GH families based on their sequence and their functional and structural characteristics. In addition to GH activity enzymes, the degradation of complex polysaccharides also involves carbohydrate esterases (CE) and polysaccharide lyases (PL) [15,16]. The degradative capabilities of enzymes are of significant importance in biotechnology and industrial applications [17,18].

The rapid development of high-throughput sequencing technologies and bioinformatics tools have improved in-depth studies of the gastrointestinal microbiota of animals. The mammalian gut microbiome represents a complex and dynamic microbial ecosystem, influenced by genetic and environmental factors, with diet being a decisive factor in the variation of gut microbial communities [19–21]. Native pig breeds exhibit superior fiber digestion capabilities [22]. A comparison of the gut microbiota between Ningxiang and Large White pigs revealed that the *Ruminococcaceae*_NK4A214_group, *Parabacteroides*, and *Ruminantium*_group genera are more abundant in Ningxiang pigs, all of which have the capacity to degrade plant polysaccharides [23]. The gut microbiome of Lantang pigs is primarily composed of Bacillota_A and Bacteroidota, formerly recognized as Firmicutes and Bacteroidetes, respectively, which are also prevalent in the human gut and the rumen [24,25].

Despite extensive research on the meat quality and growth performance of Lantang pigs [26], studies on the specific mechanisms of fiber degradation, influencing factors, and comparisons with other native pig breeds are relatively limited. Within this context, we hypothesize that, compared to other pig breeds such as Duroc pigs, the gastrointestinal microbiota of Lantang pigs possesses a unique composition and demonstrates specialized functionality, particularly in the diversity and activity of CAZymes, leading to a superior capacity for fiber degradation. We posit that this fibrolytic advantage primarily stems from a higher abundance and broader diversity of specific microbial taxa and CAZymes within the gut of Lantang pigs. The synergistic effect of these components significantly enhances the efficiency of plant fiber breakdown and utilization, thereby positively impacting the energy metabolism, intestinal health, and overall growth performance of Lantang pigs.

To reveal unknown microbial species in the gut capable of degrading complex carbohydrates, we performed shotgun metagenomic sequencing on the fecal samples of Lantang pigs. By assembling MAGs and analyzing the abundance features of these MAGs, we identified carbohydrate families unique to the Lantang pig's gut. Moreover, we discovered an abundance of carbohydrate-degrading enzymes in the MAGs of uncultured species. This article aims to provide comprehensive profiles of the CAZymes in Lantang pigs, which would provide guidance for the microbial resource utilization of Lantang pigs.

2. Materials and Methods

2.1. Experimental Design

This study aimed to analyze and compare the gut microbiota composition of Lantang pigs and Duroc pigs. Conducted at a commercial breeding farm in Shaoguan City, Guangdong Province, China, a total of 60 pigs, comprising 30 Lantang pigs and 30 Duroc pigs, were selected as the sample sources. The Duroc pigs were aged 50 $\pm$ 5 days, with an average weight of 11.8 kg, while the Lantang pigs were aged 95 $\pm$ 5 days, with an average weight of 17.3 kg. After weaning at 21 days, the pigs were transferred to individual pens, which were spacious enough to allow them free movement and rotation, ensuring basic

standards of animal welfare were met. Throughout the experimental period, we fed the pigs with commercial feed. Feeding was ceased one day prior to sample collection, though water was freely available.

Samples were collected using cotton swabs to obtain rectal feces. These samples were immediately placed into pre-labeled plastic cryogenic tubes, then immersed in liquid nitrogen to preserve freshness and prevent alterations in microbial composition. Within 12 h of collection, the samples were transferred to a −80 °C freezer in the laboratory for subsequent analysis.

It is noteworthy that all pigs remained healthy throughout the study period and were not administered any antibiotics or probiotics to avoid affecting the composition of the gut microbiota. This research strictly followed the "Regulations on the Management of Laboratory Animals" established by the Ministry of Science and Technology of the People's Republic of China, ensuring the scientific integrity and ethical compliance of the experiment. Furthermore, the study received approval from the Animal Ethics Committee of Foshan University (Approval no. 2019029902).

2.2. Genomic DNA Extraction and Metagenomic Sequencing

Bacterial DNA was extracted utilizing TIANGEN reagents by Novogene Bioinformatics Technology Co., Ltd. (Beijing, China), following the manufacturer's protocols. To ensure sample freshness and experimental accuracy, samples were preserved in liquid nitrogen and promptly transported to the laboratory for storage at −80 °C. During DNA extraction, fecal material was placed in 2 mL centrifuge tubes containing 1.5 mL of PBS buffer, followed by centrifugation at 6000 rpm for 5 min. Gel electrophoresis using agarose assessed DNA purity and integrity, while the Nanodrop system determined DNA purity through the OD 260/280 ratio. Accurate DNA quantification was performed using Qubit 2.0. Upon passing quality checks, DNA underwent random fragmentation using a Covaris ultrasonicator (Covaris, Woburn, MA, USA), followed by end repair, A-tailing, adapter ligation, purification, and PCR amplification to complete the entire library preparation process. After the library construction, initial quantification using Qubit 2.0 was followed by library dilution, fragment size analysis using an Agilent 2100 (Agilent, Santa Clara, CA, USA), and precise quantification of effective library concentration using Q-PCR to ensure library quality. Based on the effective concentration and the desired volume of target data for sequencing, different libraries were pooled for clustering on the cBOT system, followed by high-throughput sequencing on the Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA).

2.3. Bioinformatics

Quality control of raw reads was conducted using Trimmomatic v0.39 [27] with the following parameters: -threads 30 LEADING:3 TRAILING:3 SLIDINGWINDOW:5:20 MINLEN:60. Subsequently, host contamination was removed by mapping all trimmed reads to the pig reference genome (Scrofa 11.1) using bowtie2 v2.5.1 [28,29]. Quality reports for the metagenomic reads of each sample were generated using seqkit v2.6.1 [30].

To reconstruct microbial genomes from metagenomic data, we employed the spades v3.15.5 [31] tool for metagenomic assembly and contig construction. Contigs from each sample were binned using the MetaBAT2 v2.12.1 [32]. A total of 2758 MAGs were obtained, which were then assessed for completeness and contamination using the CheckM v1.2.2 [33]; MAGs with completeness greater than 50% and contamination less than 10% were retained [34,35], resulting in 1302 MAGs. Finally, the MAGs were deduplicated using dRep v3.4.5 [36] with the following parameters: -comp 50 -con 10 -sa 0.99 -p 30. This resulted in 671 non-redundant MAGs for further analysis. Additionally, after merging the contigs assembled from all samples, a complete gene set was constructed using Prodigal v2.6.3 [37], followed by deduplication with the CD-hit v4.8.1 [38].

The GTDB-tk v1.0.2 [39] tool and the latest GTDB database release v214—containing 394,932 genomes representing 80,789 bacterial species and 7777 genomes representing

4416 archaeal species—were used to infer the taxonomic lineage and construct phylogenetic trees of MAGs, visualized using iTol v6 [40]. To explore the carbohydrase profile of the Lantang pig, the latest database from CAZy [41] was downloaded. All genomes were then annotated for CAZymes using Diamond v2.1.8 [42].

3. Results

3.1. Metagenome-Assembled Genomes

In this study, approximately 217.85 billion reads were obtained through metagenomic sequencing. Following quality filtering, clean reads were assembled using genome assembly tools. Ensuring completeness of more than 50% and contamination of less than 10%, 1302 MAGs were obtained. Of these, 491 MAGs had completeness greater than 90% and contamination less than 5%, including 11 MAGs with 100% completeness and 0% contamination. Subsequently, duplicates were removed based on 99% average nucleotide identity (ANI) to eliminate redundant MAGs. Eventually, a non-redundant MAG set was composed of 671 MAGs.

The assembly quality statistics of the 671 MAGs are summarized in Supplementary Table S1. The genome size of these MAG ranges from 0.55 Mb to 4.86 Mb, with an average size of 1.86 Mb. The length of N50 for MAGs ranges from 4 kb to 325 kb. Among the 671 MAGs, 241 MAGs (35.92%) exhibit assembly quality nearing completeness (with completeness greater than 90% and contamination less than 5%). A total of 121 MAGs showed an extremely high quality with a completeness of more than 95% and contamination of less than 5%, including three MAGs with 100% completeness and 0% contamination, belonging to the genera *Megasphaera*, *Ellagibacter*, and *Methanobrevibacter_A*, with *Methanobrevibacter_A* being of archaeal origin. As previously mentioned in our article, archaea play a crucial role in the porcine gut [43].

3.2. Classification of Genomes Based on Taxonomy

The attained MAGs were compared against the Genome Taxonomy Database (GTDB), facilitating an objective classification analysis across both bacterial and archaeal categories. Figure 1 presents a phylogenetic tree of 669 non-redundant genomes. Since only two archaea were classified, both within the Euryarchaeota phylum, under the genus *Methanobrevibacter_A*—with one being *Methanobrevibacter_A smithii* and the other not identified in the database—we only display the classification of 669 bacterial genomes here. This phylogenetic tree visually illustrates the proportional representation of each bacterial phylum. The most abundant phylum were Bacillota_A (375 MAGs, 56%), Bacteroidota (114 MAGs, 17%), and Bacillota (105 MAGs, 16%), collectively comprising 89% of the total phylum identified. The remaining phylum included the following: Actinomycetota (29 MAGs), Spirochaetota (21 MAGs), Bacillota_C (11 MAGs), Cyanobacteriota (3 MAGs), Chlamydiota (3 MAGs), Pseudomonadota (2 MAGs), Fibrobacterota (2 MAGs), Patescibacteria (1 MAG), Fusobacteriota (1 MAG), Desulfobacterota (1 MAG), and Campylobacterota (1 MAG). The 14 phylum were further classified into 45 families, with the three most abundant families all stemming from the Bacillota_A phylum: Lachnospiraceae (109), Oscillospiraceae (76), and Ruminococcaceae (62). At the genus level, a total of 199 genus were classified, with the most abundant being *Prevotella* (43), *Cryptobacteroides* (31), and *Gemmiger* (26). Additionally, a genus within the Ruminococcaceae family was not identified in the database and is represented by the placeholder "g__". At the species level, annotations were made for 303 species, with the top three species by abundance being *Cryptobacteroides sp004552965* (13), *UBA636 sp002299675* (12), and *Collinsella sp002391315* (11), respectively. Notably, the number of species represented by the placeholder "s__" amounts to 97. This indicates that there are 97 species yet to be cultured, underscoring the potential for further enriching the GTDB in the future. The bacterial abundance distribution at the phylum and genus levels in Lantang and Duroc pigs is illustrated in Figure 2. In comparison with the microbial communities of Duroc pigs, overall, the total bacterial quantity and diversity within the intestinal tract of Lantang pigs are significantly higher than those observed in Duroc pigs.

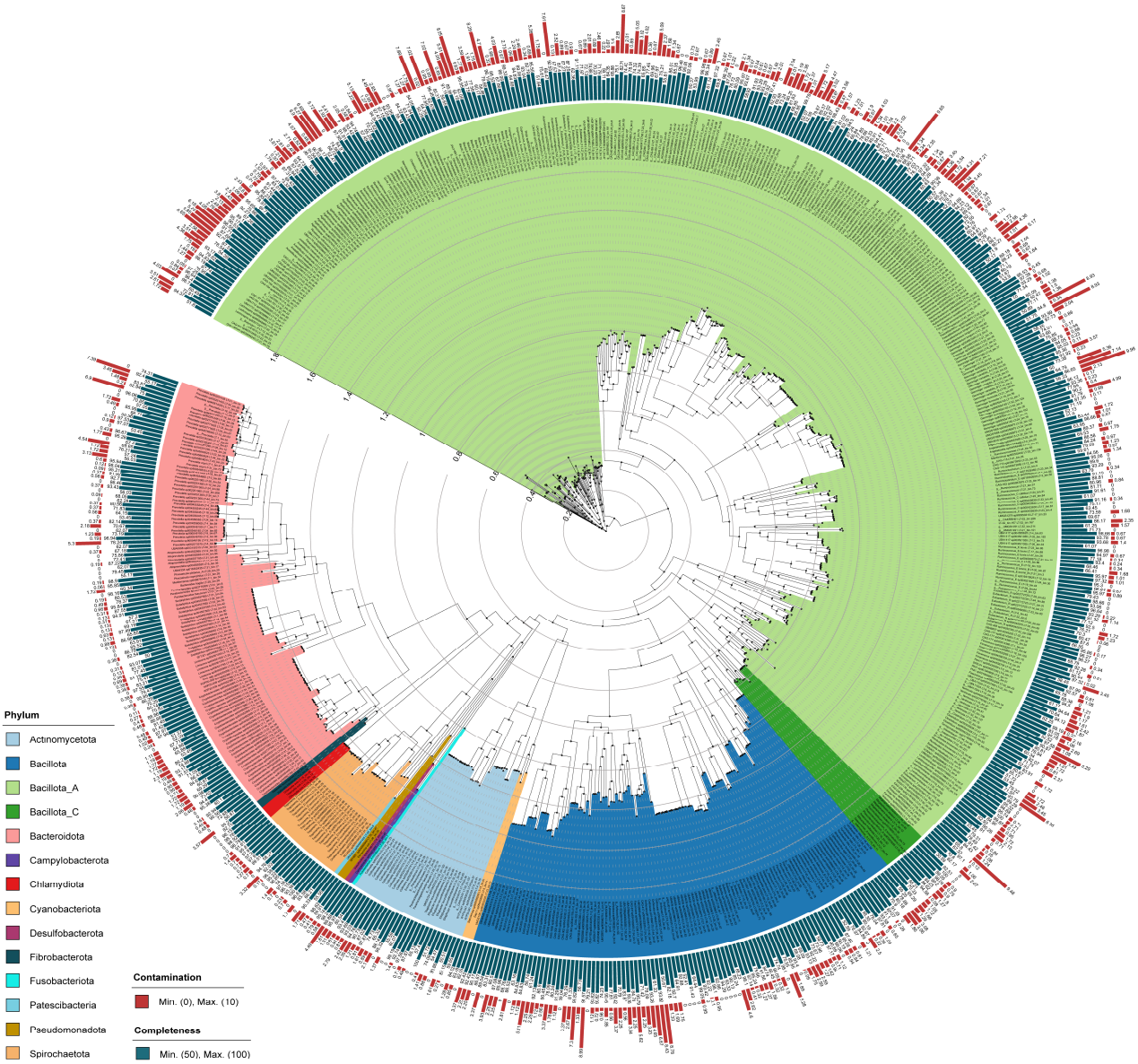

Figure 1. Evolutionary tree of intestinal microorganisms of Langtang pigs. The first circle is the species distribution of the assembled genome, different colors represent different categories, the second circle represents the completeness, and the third circle represents the degree of contamination.

Overall, all 671 genomes could be classified to the family level at least, using the GTDB database (see Supplementary Table S1). Notably, certain bacteria have been confirmed to possess fiber-degrading capabilities, for instance, 43 MAGs represent the genus *Prevotella*, which exhibits significant fibrolytic activity [44]. It can produce xylanase to break down plant fibers [45]. *Prevotella* also contains numerous enzymes capable of degrading polysaccharides [46]. Additionally, two MAGs represent the genus *Ruminococcus*, which includes proficient cellulose-degrading species such as *Ruminococcus albus* and *Ruminococcus flavefaciens*. These species can produce copious amounts of cellulase and hemicellulase to decompose fibers. *Ruminococcus flavefaciens* is capable of degrading cellulose and hemicellulose via the cellulolytic pathway [47]. Cellulosomes represent sophisticated multienzymatic assemblies synthesized by anaerobic cellulolytic bacteria to facilitate the breakdown of lignocellulosic biomass. These complexes consist of a scaffoldin core, serving as the structural foundation, to which diverse enzymatic units are attached for biomass degradation.

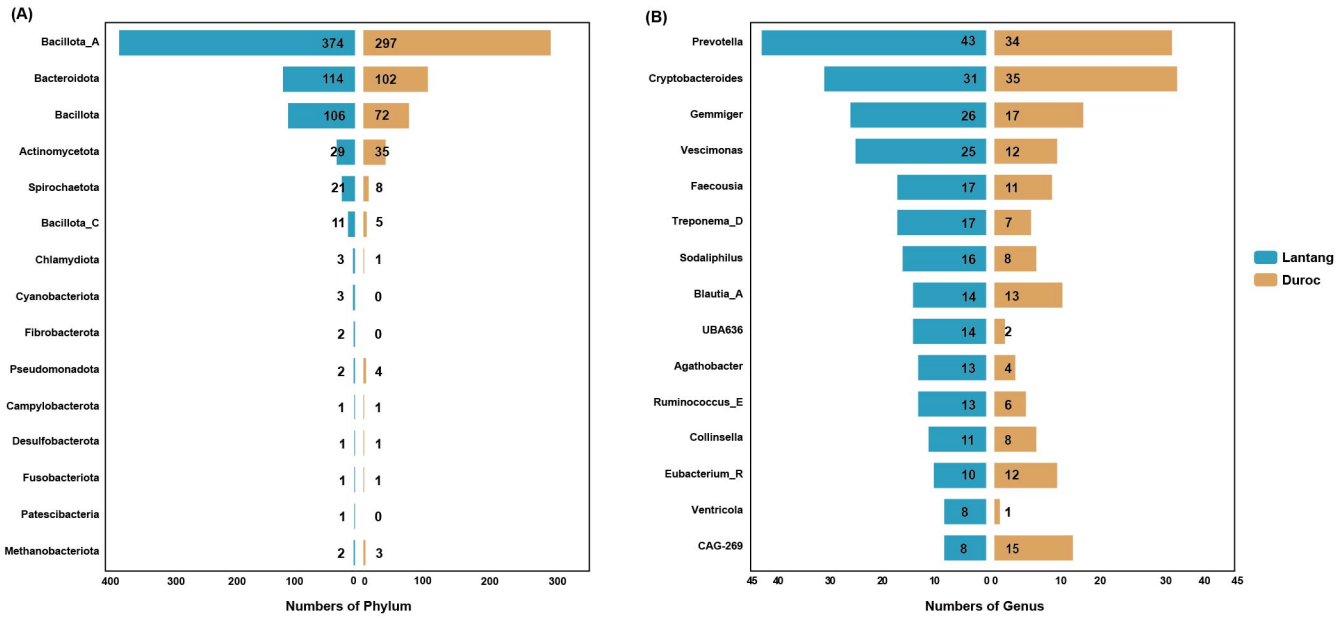

Figure 2. Quantities of bacteria in Lantang pigs and Duroc pigs at different taxonomic levels. (**A**) The graph represents the bacterial quantities at different phylum levels. (**B**) The graph represents the bacterial quantities at different genus levels.

3.3. Overview of Carbohydrate Genes in the Microbiome

A total of 671 MAGs predicted 911,124 protein-coding genes. We identified 19,264 genes encoding putative CAZy enzymes (Supplementary Table S2). According to the CAZy enzyme classification, the gene library with carbohydrate-active domains was divided into six categories [48], including 8621 glycoside hydrolases (GHs), 7081 glycosyltransferases (GTs), 2269 carbohydrate-binding module enzymes (CBMs), 904 carbohydrate esterases (CEs), 210 auxiliary activities enzymes (AAs), and 178 polysaccharide lyases (PLs). GHs and CEs are the primary enzymes for fiber degradation, accounting for approximately 50% of all CAZyme-encoding genes. At various taxonomic levels of bacterial phylum, the relative abundance of carbohydrate enzymes of different categories is depicted in Figure 3A. Numerous studies have demonstrated that among carbohydrate enzymes, the GH, CE, and PL families constitute degradative enzymes [49]. Gene counts pertaining to distinct phylum have been summarized and are presented in Table 1. On average, members of the Fusobacteriota phylum exhibit the highest encoding capacity for carbohydrate-degrading enzymes (73), exceeding those in Cyanobacteriota (37), Campylobacterota (27), and Spirochaetota (25). Notably, Bacillota_A, with the largest genome representation (356 out of 671 genomes), demonstrates an average per-genome count of 18 for GHs, PLs, and CEs. The comprehensive distribution of genes associated with carbohydrate degradation at the family level is delineated in Supplementary Table S3. Bacterial taxa within the Lachnospiraceae, Bacteroidaceae, Oscillospiraceae, Ruminococcaceae, and Muribaculaceae families are recognized for their distinguished fibrolytic attributes [50,51], with the number of genes encoding GHs being 2, 141, 840, 819, 690, and 112, respectively. To specifically explore the microbial potential for dietary fiber degradation in the gut of Lantang pigs, we counted the number of carbohydrate-degrading enzyme genes and their corresponding degradation family numbers based on the encoded carbohydrate enzyme classification information, as shown in Figure 3B. In the course of this study, we identified a comprehensive array of 6232 carbohydrate-degrading enzyme families and 9703 corresponding degradative genes. Among these, 90 bacterial species, which lacked annotations in the GTDB database, were represented in the analysis by their generic classifications. Notably, the phylum Bacillota_A, Spirochaetota, and Bacteroidota were characterized by their substantial contribution to the repertoire of carbohydrate-degrading enzymes. Within the Bacillota_A phylum specifically, an exceptionally high prevalence of degradative enzymes was documented, underscoring

its significant role in carbohydrate degradation, with the top three being MAG026 (129 GHs, 14 CEs, 1 PL), MAG468 (122 GHs, 13 CEs, 3 PLs), and MAG123 (GHs 119, CEs 11, PLs 4). Notably, large numbers of GHs were also found in some unnamed bacterial families, such as UBA9506 (averaging 109 GH families per genome), HGM11514 (69), and CAG-74 (58); interestingly, these also belong to the Bacillota_A phylum. From this, we can infer that the Bacillota_A phylum plays a dominant role in the process of fiber degradation.

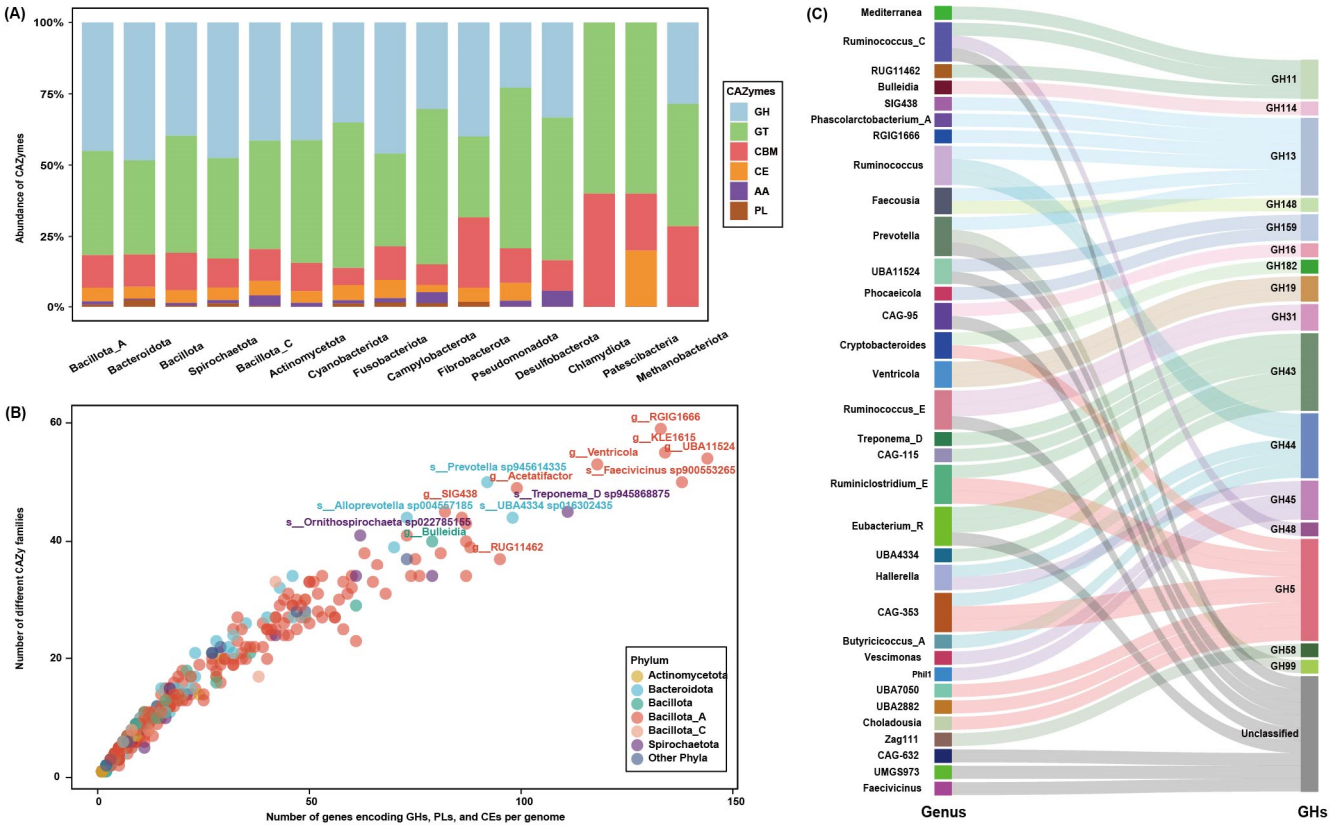

Figure 3. Comprehensive assessment of carbohydrate-degrading enzyme profiles: phylum-level distribution, family and gene abundance, and unique GH enzyme characteristics in the Lantang pig intestinal microbiota. (**A**) The proportion of bacterial phylum contained in different carbohydrate enzymes; (**B**) distribution of the number of carbohydrate-degrading enzyme families and the number of carbohydrate enzyme genes, where different colors represent bacteria at different phylum levels; (**C**) the unique GH family enzymes in the intestine of Lantang pigs correspond to bacterial genus.

Furthermore, this investigation also applied identical sequencing and bioinformatics methodologies to analyze fecal samples from 30 Duroc pigs. Compared with the findings from the Duroc group, the analysis revealed that the Lantang pig group harbored 87 unique CAZY enzyme genes. This subset included 53 GHs, 14 GTs, 13 PLs, 4 CBMs, and 3 CEs. The differentiation based on GH enzyme classes was visually depicted in Figure 3C, highlighting genera such as Ruminococcus, Prevotella, Eubacterium_R, and CAG-353. These genera were noted for their higher representation and the structural domains of their associated GH families were biochemically characterized to predominantly act upon cellulose, specifically through enzymes GH5, GH6, GH9, GH12, GH44, GH45, and GH48.

Table 1. Distribution and quantity statistics of carbohydrate degradation genes at the bacterial phylum level.

Phylum	MAGs	GHs		CEs		PLs		Total	
		No.	Mean	No.	Mean	No.	Mean	No.	Mean
Actinomycetota	27	121	4.5	12	0.4	0	0.0	133	4.9
Bacillota	99	585	5.9	66	0.7	5	0.1	656	6.6
Bacillota_A	356	5837	16.4	625	1.8	92	0.3	6554	18.4
Bacillota_C	11	180	16.4	22	2.0	1	0.1	203	18.5
Bacteroidota	107	1229	11.5	106	1.0	62	0.6	1397	13.1
Campylobacterota	1	24	24.0	2	2.0	1	1.0	27	27.0
Chlamydiota	2	0	0.0	0	0.0	0	0.0	0	0.0
Cyanobacteriota	3	93	31.0	14	4.7	3	1.0	110	36.7
Desulfobacterota	1	6	6.0	0	0.0	0	0.0	6	6.0
Fibrobacterota	2	24	12.0	3	1.5	1	0.5	28	14.0
Fusobacteriota	1	62	62.0	9	9.0	2	2.0	73	73.0
Methanobacteriota	1	2	2.0	0	0.0	0	0.0	2	2.0
Patescibacteria	1	0	0.0	1	1.0	0	0.0	1	1.0
Pseudomonadota	2	11	5.5	3	1.5	0	0.0	14	7.0
Spirochaetota	20	447	22.4	41	2.1	11	0.6	499	25.0

4. Discussion

The gut microbiota comprises a complex community of microorganisms that colonize the gastrointestinal tract, existing in symbiosis with their host [52]. The host provides a habitat for the survival and evolution of microorganisms. In return, the microbial community benefits the host by supplying essential nutrients, aiding in the digestion of dietary components, and enhancing immune function [53–56]. Mammals typically lack the enzymes necessary to decompose complex carbohydrates, particularly plant-based dietary fibers [57,58]. The selection of dietary fibers exerts a significant impact on the gut microbiome of the host [59]. Microbial communities play a crucial role in the digestion of these complex carbohydrates. To better understand microbial degradation of feed carbohydrates in the gut ecosystem of Lantang pigs, this study conducted metagenomic sequencing on fecal samples from 30 Lantang and 30 Duroc pigs, and assembled 699 bacterial MAGs and 2 archaeal MAGs, including 97 species from uncultured microorganisms. These were predominantly found in the genus *Faecousia* (6), *Ventricola* (6), *Ruminococcus_E* (5), and *CAG-269* (5). These MAGs play various physiological and biochemical roles within the intestines of Lantang pigs. Notably, the Lachnospiraceae and Ruminococcaceae families contain 109 and 61 genomes, respectively, many of which are major butyrate producers with the metabolic capacity to degrade and utilize plant-derived fibers as nutrients [60,61]. There are 32 genus more prevalent in the fecal microbiota of Duroc pigs than in Lantang pigs, while 121 genus are more prevalent in Lantang pigs, with *Prevotella* being the most dominant, known for its potent fiber-degrading activity [62,63]. The enhanced fiber-digesting capability observed in Lantang pigs may be attributed to a higher abundance of fiber-digesting bacteria in the large intestine.

To compare the capabilities of the two pig breeds in terms of carbohydrate-degrading enzymes, we aligned sequences with the CAZy database using software to identify putative genes for CAZymes. A total of 911,124 protein-coding genes were predicted from 671 MAGs of Lantang pigs. Of these, 6232 carbohydrate-degrading enzyme families and 9703 degrading genes were identified. GHs are the most abundant enzymes in the gut microbiota of mammals, primarily hydrolyzing glycosidic bonds in glycosides, with substrate specificity similar to that of GTs [64]. Duroc pig samples contained 5197 GHs, while Lantang pig samples contained 8622 GHs, with unique carbohydrate-degrading enzymes present in Lantang pig samples. For example, enzymes such as GH5, GH11, GH13, GH43, GH45, and PL22 are capable of breaking down indigestible fibers into monosaccharides.

Concurrently, genus such as *Ruminococcus*, *Prevotella*, and *Hallerella* continuously ferment food, facilitating the production of these enzymes.

Pigs lack the enzymes necessary for the hydrolysis of dietary fibers. And the breakdown of dietary fiber in the pig's digestive system is primarily mediated by microbial fermentation. Soluble fibers typically ferment more rapidly than insoluble fibers [65]. When selecting diets for Lantang pigs, it is recommended to prioritize feeds enriched with soluble fibers. Compared to monogastric animals like pigs, ruminants possess a multi-chambered stomach structure and a unique rumination process that significantly enhances fiber degradation rates [66]. So, investigating the differences in microbial communities and fiber degradation capacities between monogastric and ruminant animals constitutes an intriguing research avenue.

5. Conclusions

Despite shotgun metagenomics offering a pathway to characterize the taxonomic and functional potential of carbohydrate processing in complex microbial communities, this technology still faces many limitations [67]. The use of high-quality, filtered metagenome-assembled genomes (MAGs) significantly enhances the credibility of research. In this study, compared to Duroc pigs, we found a significant and unique fibrolytic microbial community in the intestines of Lantang pigs, dominating the enzyme-driven fiber degradation process. The activity of these microbes and enzymes enables Lantang pigs to convert indigestible dietary fiber into essential energy. Furthermore, we identified 97 microbial species not listed in the GTDB database and 75 carbohydrate enzymes not recorded in the CAZy database. These discoveries not only broaden our understanding of microbial diversity but also provide valuable resources for further exploring the potential of microbes in environmental and industrial applications. Moreover, utilizing these data to enrich existing microbial and enzyme databases lays an important foundation for microbial ecology, phylogenetic analysis, and biotechnological development, potentially offering solutions for future food competition challenges between humans and livestock.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/fermentation10040207/s1, Table S1: Summary of 671 assembled genome data assembled and classified from the fecal microbiome of Lantang pigs. This includes detailed information on each genome, including integrity and contamination, GC% content, genome size, number of Contigs, Longest contig, N50 contig, and estimates of the predicted taxonomic lineage for GTDB-tk; Table S2: List of CAZymes families encoded in the genomes assembled from metagenomic datasets; Table S3: Statistics on the distribution and quantity of carbohydrate degradation genes at the family level.

Author Contributions: Conceptualization, J.Z. and Y.L.; methodology, J.Y. and Y.F.; software, J.Y. and Y.P.; validation, J.C., X.W. and Y.Z.; formal analysis, R.J. and F.D.; investigation, J.Z.; resources, Y.L.; data curation, J.Y.; writing—original draft preparation, J.Y. and Y.F.; writing—review and editing, F.D. and J.Z.; visualization, J.Y.; supervision, Y.L.; funding acquisition, J.Z. and Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Research and Development Program of China (2023YFE0124400), Youth Project of Guangdong Foshan joint fund of the Guangdong Natural Science Foundation (2022A1515110819), Foshan University Student Academic Fund (xsjj202309kjb02), Guangdong Provincial Key Laboratory of Animal Molecular Design and Precise Breeding (2019B030301010), and Key Laboratory of Animal Molecular Design and Precise Breeding of Guangdong Higher Education Institutes (2019KSYS011).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Raw sequencing files and associated metadata have been deposited at NCBI's Sequence Read Archive (accession PRJNA1079049).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Song, B.; Zheng, C.; Zheng, J.; Zhang, S.; Zhong, Y.; Guo, Q.; Li, F.; Long, C.; Xu, K.; Duan, Y. Comparisons of carcass traits, meat quality, and serum metabolome between Shaziling and Yorkshire pigs. *Anim. Nutr.* **2022**, *8*, 125–134. [CrossRef] [PubMed]
2. Zhang, X.; Wang, X.; Mu, L.; Ding, Z. Immune responses in pigs induced by recombinant DNA vaccine co-expressing swine IL-18 and membrane protein of porcine reproductive and respiratory syndrome virus. *Int. J. Mol. Sci.* **2012**, *13*, 5715–5728. [CrossRef] [PubMed]
3. Li, L.; Li, S.; Luo, J.; Chen, T.; Xi, Q.; Zhang, Y.; Sun, J. The difference of intestinal microbiota composition between Lantang and Landrace newborn piglets. *BMC Vet. Res.* **2023**, *19*, 174. [CrossRef] [PubMed]
4. Wassie, T.; Cheng, B.; Zhou, T.; Gao, L.; Lu, Z.; Xie, C.; Wu, X. Microbiome-metabolome analysis reveals alterations in the composition and metabolism of caecal microbiota and metabolites with dietary Enteromorpha polysaccharide and Yeast glycoprotein in chickens. *Front. Immunol.* **2022**, *13*, 996897. [CrossRef]
5. Wardman, J.F.; Bains, R.K.; Rahfeld, P.; Withers, S.G. Carbohydrate-active enzymes (CAZymes) in the gut microbiome. *Nat. Rev. Microbiol.* **2022**, *20*, 542–556. [CrossRef] [PubMed]
6. Kaddouch, E.; Cleveland, M.E.; Navarro, D.; Grisel, S.; Haon, M.; Brumer, H.; Lafond, M.; Berrin, J.-G.; Bissaro, B. A simple and direct ionic chromatography method to monitor galactose oxidase activity. *RSC Adv.* **2022**, *12*, 26042–26050. [CrossRef]
7. Kaoutari, A.E.; Armougom, F.; Gordon, J.I.; Raoult, D.; Henrissat, B. The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nat. Rev. Microbiol.* **2013**, *11*, 497–504. [CrossRef]
8. Kassan, M.; Kwon, Y.; Munkhsaikhan, U.; Sahyoun, A.M.; Ishrat, T.; Galán, M.; Gonzalez, A.A.; Abidi, A.H.; Kassan, A.; Ait-Aissa, K. Protective Role of Short-Chain Fatty Acids against Ang-II-Induced Mitochondrial Dysfunction in Brain Endothelial Cells: A Potential Role of Heme Oxygenase 2. *Antioxidants* **2023**, *12*, 160. [CrossRef] [PubMed]
9. Ma, L.; Tao, S.; Song, T.; Lyu, W.; Li, Y.; Wang, W.; Shen, Q.; Ni, Y.; Zhu, J.; Zhao, J. Clostridium butyricum and carbohydrate active enzymes contribute to the reduced fat deposition in pigs. *iMeta* **2024**, *3*, e160. [CrossRef]
10. Wang, Y.; LaPointe, G. Arabinogalactan Utilization by Bifidobacterium longum subsp. longum NCC 2705 and Bacteroides caccae ATCC 43185 in Monoculture and Coculture. *Microorganisms* **2020**, *8*, 1703. [CrossRef]
11. Cockburn, D.W.; Koropatkin, N.M. Polysaccharide degradation by the intestinal microbiota and its influence on human health and disease. *J. Mol. Biol.* **2016**, *428*, 3230–3252. [CrossRef]
12. Nieter, A.; Haase-Aschoff, P.; Kelle, S.; Linke, D.; Krings, U.; Popper, L.; Berger, R.G. A chlorogenic acid esterase with a unique substrate specificity from Ustilago maydis. *Appl. Environ. Microbiol.* **2015**, *81*, 1679–1688. [CrossRef] [PubMed]
13. Cavalier, D.M.; Lerouxel, O.; Neumetzler, L.; Yamauchi, K.; Reinecke, A.; Freshour, G.; Zabotina, O.A.; Hahn, M.G.; Burgert, I.; Pauly, M. Disrupting two Arabidopsis thaliana xylosyltransferase genes results in plants deficient in xyloglucan, a major primary cell wall component. *Plant Cell* **2008**, *20*, 1519–1537. [CrossRef] [PubMed]
14. Plakys, G.; Gasparavičiūtė, R.; Vaitekūnas, J.; Rutkienė, R.; Meškys, R. Characterization of *Paenibacillus* sp. GKG endo-β-1, 3-glucanase, a member of family 81 glycoside hydrolases. *Microorganisms* **2022**, *10*, 1930. [CrossRef] [PubMed]
15. Yang, J.C.; Madupu, R.; Durkin, A.S.; Ekborg, N.A.; Pedamallu, C.S.; Hostetler, J.B.; Radune, D.; Toms, B.S.; Henrissat, B.; Coutinho, P.M. The complete genome of Teredinibacter turnerae T7901: An intracellular endosymbiont of marine wood-boring bivalves (shipworms). *PLoS ONE* **2009**, *4*, e6085. [CrossRef] [PubMed]
16. Sidhu, C.; Kirstein, I.V.; Meunier, C.L.; Rick, J.; Fofonova, V.; Wiltshire, K.H.; Steinke, N.; Vidal-Melgosa, S.; Hehemann, J.-H.; Huettel, B. Dissolved storage glycans shaped the community composition of abundant bacterioplankton clades during a North Sea spring phytoplankton bloom. *Microbiome* **2023**, *11*, 77. [CrossRef] [PubMed]
17. Andaleeb, H.; Ullah, N.; Falke, S.; Perbandt, M.; Brognaro, H.; Betzel, C. High-resolution crystal structure and biochemical characterization of a GH11 endoxylanase from Nectria haematococca. *Sci. Rep.* **2020**, *10*, 15658. [CrossRef] [PubMed]
18. Schultz-Johansen, M.; Bech, P.K.; Hennessy, R.C.; Glaring, M.A.; Barbeyron, T.; Czjzek, M.; Stougaard, P. A novel enzyme portfolio for red algal polysaccharide degradation in the marine bacterium Paraglaciecola hydrolytica S66T encoded in a sizeable polysaccharide utilization locus. *Front. Microbiol.* **2018**, *9*, 839. [CrossRef] [PubMed]
19. Zhu, L.; Wu, Q.; Dai, J.; Zhang, S.; Wei, F. Evidence of cellulose metabolism by the giant panda gut microbiome. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17714–17719. [CrossRef]
20. Borbón-García, A.; Reyes, A.; Vives-Flórez, M.; Caballero, S. Captivity shapes the gut microbiota of Andean bears: Insights into health surveillance. *Front. Microbiol.* **2017**, *8*, 1316. [CrossRef] [PubMed]
21. Xu, J.; Li, Y.; Yang, Z.; Li, C.; Liang, H.; Wu, Z.; Pu, W. Yeast probiotics shape the gut microbiome and improve the health of early-weaned piglets. *Front. Microbiol.* **2018**, *9*, 2011. [CrossRef]
22. Fevrier, C.; Bourdon, D.; Aumaitre, A. Effects of level of dietary fibre from wheat bran on digestibility of nutrients, digestive enzymes and performance in the European Large White and Chinese Mei Shan pig. *J. Anim. Physiol. Anim. Nutr.* **1992**, *68*, 60–72. [CrossRef]
23. Lei, L.; Wang, Z.; Li, J.; Yang, H.; Yin, Y.; Tan, B.; Chen, J. Comparative microbial profiles of colonic digesta between ningxiang pig and large white pig. *Animals* **2021**, *11*, 1862. [CrossRef]
24. Rios-Covian, D.; Salazar, N.; Gueimonde, M.; de Los Reyes-Gavilan, C.G. Shaping the metabolism of intestinal Bacteroides population through diet to improve human health. *Front. Microbiol.* **2017**, *8*, 376. [CrossRef] [PubMed]

25. Pinnell, L.J.; Reyes, A.A.; Wolfe, C.A.; Weinroth, M.D.; Metcalf, J.L.; Delmore, R.J.; Belk, K.E.; Morley, P.S.; Engle, T.E. Bacteroidetes and firmicutes drive differing microbial diversity and community composition among micro-environments in the bovine rumen. *Front. Vet. Sci.* **2022**, *9*, 897996. [CrossRef] [PubMed]
26. Yu, K.; Shu, G.; Zhu, X.; Gao, P.; Wang, S.; Wang, L.; Xi, Q.; Zhang, S.; Zhang, Y.; Li, Y. Fatty acid and transcriptome profiling of longissimus dorsi muscles between pig breeds differing in meat quality. *Int. J. Biol. Sci.* **2013**, *9*, 108. [CrossRef] [PubMed]
27. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef] [PubMed]
28. Langdon, W.B. Performance of genetic programming optimised Bowtie2 on genome comparison and analytic testing (GCAT) benchmarks. *BioData Min.* **2015**, *8*, 1. [CrossRef] [PubMed]
29. Warr, A.; Affara, N.; Aken, B.; Beiki, H.; Bickhart, D.M.; Billis, K.; Chow, W.; Eory, L.; Finlayson, H.A.; Flicek, P. An improved pig reference genome sequence to enable pig genetics and genomics research. *Gigascience* **2020**, *9*, giaa051. [CrossRef] [PubMed]
30. Shen, W.; Le, S.; Li, Y.; Hu, F. SeqKit: A cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **2016**, *11*, e0163962. [CrossRef] [PubMed]
31. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Prjibelski, A.D. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [CrossRef] [PubMed]
32. 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] [PubMed]
33. 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] [PubMed]
34. Deng, F.; Han, Y.; Huang, Y.; Li, D.; Chai, J.; Deng, L.; Wei, M.; Wu, K.; Zhao, H.; Yang, G. A comprehensive analysis of antibiotic resistance genes in the giant panda gut. *iMeta* **2024**, e171. [CrossRef]
35. Deng, F.; Wang, C.; Li, D.; Peng, Y.; Deng, L.; Zhao, Y.; Zhang, Z.; Wei, M.; Wu, K.; Zhao, J. The unique gut microbiome of giant pandas involved in protein metabolism contributes to the host's dietary adaption to bamboo. *Microbiome* **2023**, *11*, 180. [CrossRef]
36. Olm, M.R.; Brown, C.T.; Brooks, B.; Banfield, J.F. dRep: A tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *ISME J.* **2017**, *11*, 2864–2868. [CrossRef] [PubMed]
37. Hyatt, D.; Chen, G.-L.; LoCascio, P.F.; Land, M.L.; Larimer, F.W.; Hauser, L.J. Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform.* **2010**, *11*, 119. [CrossRef]
38. Li, W.; Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **2006**, *22*, 1658–1659. [CrossRef] [PubMed]
39. Chaumeil, P.-A.; Mussig, A.J.; Hugenholtz, P.; Parks, D.H. GTDB-Tk: A toolkit to classify genomes with the Genome Taxonomy Database. *Bioinformatics* **2020**, *36*, 1925–1927. [CrossRef]
40. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [CrossRef]
41. Cantarel, B.L.; Coutinho, P.M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active EnZymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233–D238. [CrossRef] [PubMed]
42. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **2015**, *12*, 59–60. [CrossRef] [PubMed]
43. Yang, J.; Chen, R.; Peng, Y.; Chai, J.; Li, Y.; Deng, F. The role of gut archaea in the pig gut microbiome: A mini-review. *Front. Microbiol.* **2023**, *14*, 1284603. [CrossRef]
44. Chen, T.; Long, W.; Zhang, C.; Liu, S.; Zhao, L.; Hamaker, B.R. Fiber-utilizing capacity varies in Prevotella-versus Bacteroides-dominated gut microbiota. *Sci. Rep.* **2017**, *7*, 2594. [CrossRef] [PubMed]
45. Gorvitovskaia, A.; Holmes, S.P.; Huse, S.M. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome* **2016**, *4*, 15. [CrossRef]
46. Flint, H.J.; Bayer, E.A.; Rincon, M.T.; Lamed, R.; White, B.A. Polysaccharide utilization by gut bacteria: Potential for new insights from genomic analysis. *Nat. Rev. Microbiol.* **2008**, *6*, 121–131. [CrossRef]
47. Artzi, L.; Bayer, E.A.; Moraïs, S. Cellulosomes: Bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol.* **2017**, *15*, 83–95. [CrossRef] [PubMed]
48. Chai, J.; Zhuang, Y.; Cui, K.; Bi, Y.; Zhang, N. Metagenomics reveals the temporal dynamics of the rumen resistome and microbiome in goat kids. *Microbiome* **2024**, *12*, 14. [CrossRef]
49. Helbert, W.; Poulet, L.; Drouillard, S.; Mathieu, S.; Loiodice, M.; Couturier, M.; Lombard, V.; Terrapon, N.; Turchetto, J.; Vincentelli, R. Discovery of novel carbohydrate-active enzymes through the rational exploration of the protein sequences space. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6063–6068. [CrossRef]
50. Li, H.; Ma, L.; Li, Z.; Yin, J.; Tan, B.; Chen, J.; Jiang, Q.; Ma, X. Evolution of the gut microbiota and its fermentation characteristics of Ningxiang pigs at the young stage. *Animals* **2021**, *11*, 638. [CrossRef]
51. Xu, Q.; Yuan, X.; Gu, T.; Li, Y.; Dai, W.; Shen, X.; Song, Y.; Zhang, Y.; Zhao, W.; Chang, G. Comparative characterization of bacterial communities in geese fed all-grass or high-grain diets. *PLoS ONE* **2017**, *12*, e0185590. [CrossRef] [PubMed]

52. Liang, M.; Liu, J.; Chen, W.; He, Y.; Kahaer, M.; Li, R.; Tian, T.; Liu, Y.; Bai, B.; Cui, Y. Diagnostic model for predicting hyperuricemia based on alterations of the gut microbiome in individuals with different serum uric acid levels. *Front. Endocrinol.* **2022**, *13*, 925119. [CrossRef] [PubMed]
53. Jiminez, J.A.; Uwiera, T.C.; Abbott, D.W.; Uwiera, R.R.; Inglis, G.D. Impacts of resistant starch and wheat bran consumption on enteric inflammation in relation to colonic bacterial community structures and short-chain fatty acid concentrations in mice. *Gut Pathog.* **2016**, *8*, 67. [CrossRef] [PubMed]
54. Wielgosz-Grochowska, J.P.; Domanski, N.; Drywień, M.E. Efficacy of an irritable bowel syndrome diet in the treatment of small intestinal bacterial overgrowth: A narrative review. *Nutrients* **2022**, *14*, 3382. [CrossRef] [PubMed]
55. Sun, G.; Zhang, H.; Wei, Q.; Zhao, C.; Yang, X.; Wu, X.; Xia, T.; Liu, G.; Zhang, L.; Gao, Y. Comparative analyses of fecal microbiota in European mouflon (*Ovis orientalis musimon*) and blue sheep (*Pseudois nayaur*) living at low or high altitudes. *Front. Microbiol.* **2019**, *10*, 1735. [CrossRef] [PubMed]
56. Bevins, C.L.; Salzman, N.H. The potter's wheel: The host's role in sculpting its microbiota. *Cell. Mol. Life Sci.* **2011**, *68*, 3675–3685. [CrossRef]
57. Xiao, K.; Liang, X.; Lu, H.; Li, X.; Zhang, Z.; Lu, X.; Wang, H.; Meng, Y.; Roy, A.; Luo, W. Adaptation of gut microbiome and host metabolic systems to lignocellulosic degradation in bamboo rats. *ISME J.* **2022**, *16*, 1980–1992. [CrossRef] [PubMed]
58. Pu, G.; Li, P.; Du, T.; Niu, Q.; Fan, L.; Wang, H.; Liu, H.; Li, K.; Niu, P.; Wu, C. Adding appropriate fiber in diet increases diversity and metabolic capacity of distal gut microbiota without altering fiber digestibility and growth rate of finishing pig. *Front. Microbiol.* **2020**, *11*, 533. [CrossRef]
59. Meng, X.; Zheng, J.; Wang, F.; Zheng, J.; Yang, D. Dietary fiber chemical structure determined gut microbiota dynamics. *iMeta* **2022**, *1*, e64. [CrossRef]
60. Flint, H.J.; Scott, K.P.; Duncan, S.H.; Louis, P.; Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **2012**, *3*, 289–306. [CrossRef] [PubMed]
61. Milani, C.; Duranti, S.; Bottacini, F.; Casey, E.; Turroni, F.; Mahony, J.; Belzer, C.; Delgado Palacio, S.; Arboleya Montes, S.; Mancabelli, L. The first microbial colonizers of the human gut: Composition, activities, and health implications of the infant gut microbiota. *Microbiol. Mol. Biol. Rev.* **2017**, *81*, e00036-17. [CrossRef] [PubMed]
62. Li, D.; Wang, P.; Wang, P.; Hu, X.; Chen, F. Gut microbiota promotes production of aromatic metabolites through degradation of barley leaf fiber. *J. Nutr. Biochem.* **2018**, *58*, 49–58. [CrossRef] [PubMed]
63. Zhao, W.; Abdelsattar, M.M.; Wang, X.; Zhang, N.; Chai, J. In Vitro Modulation of Rumen Fermentation by Microbiota from the Recombination of Rumen Fluid and Solid Phases. *Microbiol. Spectr.* **2023**, *11*, e0338722. [CrossRef] [PubMed]
64. Nakamura, T.; Fahmi, M.; Tanaka, J.; Seki, K.; Kubota, Y.; Ito, M. Genome-wide analysis of whole human glycoside hydrolases by data-driven analysis in silico. *Int. J. Mol. Sci.* **2019**, *20*, 6290. [CrossRef] [PubMed]
65. Agyekum, A.K.; Nyachoti, C.M. Nutritional and Metabolic Consequences of Feeding High-Fiber Diets to Swine: A Review. *Engineering* **2017**, *3*, 716–725. [CrossRef]
66. Weimer, P.J. Degradation of Cellulose and Hemicellulose by Ruminal Microorganisms. *Microorganisms* **2022**, *10*, 2345. [CrossRef] [PubMed]
67. Nayfach, S.; Roux, S.; Seshadri, R.; Udwary, D.; Varghese, N.; Schulz, F.; Wu, D.; Paez-Espino, D.; Chen, I.-M.; Huntemann, M. A genomic catalog of Earth's microbiomes. *Nat. Biotechnol.* **2021**, *39*, 499–509. [CrossRef] [PubMed]