

Article

Comparative Genomics of *Lactobacillus crispatus* from the Gut and Vagina Reveals Genetic Diversity and Lifestyle Adaptation

Qiuxiang Zhang^{1,2,3}, Lili Zhang^{1,2}, Paul Ross^{3,4}, Jianxin Zhao^{1,2}, Hao Zhang^{1,2,5,6} and Wei Chen^{1,2,5,7,*}

¹ State Key Laboratory of Food Science and Technology, Jiangnan University, Wuxi 214122, China; zhangqx@jiangnan.edu.cn (Q.Z.); 17851318862@163.com (L.Z.); zhaojianxin@jiangnan.edu.cn (J.Z.); zhanghao61@jiangnan.edu.cn (H.Z.)

² School of Food Science and Technology, Jiangnan University, Wuxi 214122, China

³ International Joint Research Center for Probiotics & Gut Health, Jiangnan University, Wuxi 214122, China; p.ross@ucc.ie

⁴ APC Microbiome Ireland, University College Cork, Cork, Ireland

⁵ National Engineering Research Center for Functional Food, Jiangnan University, Wuxi 214122, China

⁶ Wuxi Translational Medicine Research Center and Jiangsu Translational Medicine Research Institute Wuxi Branch, Wuxi 214122, China

⁷ Beijing Innovation Centre of Food Nutrition and Human Health, Beijing Technology and Business University (BTBU), Beijing 100048, China

* Correspondence: chenwei66@jiangnan.edu.cn; Tel.: +86-510-859-12155

Received: 13 February 2020; Accepted: 25 March 2020; Published: 27 March 2020



Abstract: *Lactobacillus crispatus* colonizes the human feces, human vagina, and the crops and ceca of chicken. To explore the genetic characteristics and evolutionary relationships of *L. crispatus* isolated from different niches, we selected 37 strains isolated from the human vagina (n = 17), human feces (n = 11), and chicken feces (n = 9), and used comparative genomics to explore the genetic information of *L. crispatus* from the feces and vagina. No significant difference was found in the three sources of genomic features such as genome size, GC content, and number of protein coding sequences (CDS). However, in a phylogenetic tree constructed based on core genes, vagina-derived *L. crispatus* and feces-derived strains were each clustered separately. Therefore, the niche exerted an important impact on the evolution of *L. crispatus*. According to gene annotation, the *L. crispatus* derived from the vagina possessed a high abundance of genes related to acid tolerance, redox reactions, pullulanase, and carbohydrate-binding modules (CBMs). These genes helped *L. crispatus* to better adapt to the acidic environment of the vagina and obtain more nutrients, maintaining its dominance in the vagina in competition with other strains. In feces-derived bacteria, more genes encoding CRISPR/Cas system, glycoside hydrolases (GHs) family, and tetracycline/lincomycin resistance genes were found to adapt to the complex intestinal environment. This study highlights the evolutionary relationship of *L. crispatus* strains isolated from the vagina and feces, and the adaptation of *L. crispatus* to the host environment.

Keywords: *Lactobacillus crispatus*; comparative genomics; evolution; host adaptation; acid tolerance

1. Introduction

Lactobacillus crispatus is an obligately homofermentative, catalase-negative, Gram-positive bacterium [1]. Numerous studies have shown the considerable probiotic potential of *L. crispatus*, such as modulating the host's immune system [2], reducing allergic symptoms in mice [3], down-regulating

the expression of cancer genes, and inhibiting the proliferation of MDA-MB-231 breast cancer cell lines [4]. In addition, as one of the most detected *Lactobacillus* species in the healthy female vagina [5], *L. crispatus* can inhibit vaginal pathogenic microorganisms, and maintain the health of the female reproductive system [6–8]. Moreover, clinical trials have shown the therapeutic effect of *L. crispatus* on female vaginitis and recurrent urinary tract infection [9,10].

In recent years, with the rapid development of gene sequencing technology, we can conduct in-depth researches on the genetic and functional diversity of bacteria at the genome level. Comparative genomics provides favorable support for understanding the basic genetic features, species evolution, and host adaptation of strains [11,12]. Teija Ojala et al. compared the pan- and core genomes of 10 *L. crispatus* strains, and revealed that some core genes implicated in protecting the urogenital tract from *Gardnerella vaginalis* colonization [13]. Van der Veer et al. found that glycosyltransferase (GT) genes were more abundant in *L. crispatus* isolated from vaginal samples with dysbiotic vaginal microbiota than that from the vaginal tracts of women with *Lactobacillus*-dominated vaginal microbiota (LVM). The vaginal niche in a dysbiotic state was under some immune pressure. The GTs regulated the surface glycogen of *L. crispatus* through a phase change, helping it escape immunity [14].

Other species of *Lactobacillus* that have been subjected to comparative genomic analysis are *L. reuteri* [15], *L. salivarius* [16], *L. curvatus* [17], and *L. brevis* [18]. The results showed that niches affected the genetic characteristics and evolutionary orientations of the given species. Meanwhile, the strains exhibited a specific host adaptability during the evolution process. However, the above studies only compared *L. crispatus* isolated from the same niche (vagina). Therefore, the differences in the genetic information inherent in *L. crispatus* isolated from different niches remain unclear. Recently, comparative genomics of two *L. crispatus* strains from the vagina and feces showed that the two isolates differed slightly in terms of genomic size, exopolysaccharides, and CRISPR–Cas system [19]. However, this work did not reveal the genetic relationships of strains from different habitats, and only focused on limited number of strains.

In this study, we analyzed 37 *L. crispatus* strains isolated from the human vagina, human feces, and chicken feces. Comparative genomics was applied to analyze whether feces-derived strains and vagina-derived strains differed significantly in basic genetic features and evolutionary relationships. In addition, the genes of *L. crispatus* associated with intestinal or vaginal adaptation were highlighted.

2. Materials and Methods

2.1. *L. crispatus* Strain Isolation

The human and chicken feces, and the vaginal swabs from healthy people were plated on the *Lactobacillus*-selective medium (LBS) agar under anaerobic conditions for 48 h at 37 °C [20]. Colonies with morphology (white, round, central raised, irregular edges) similar to *L. crispatus* were picked and streaked on LBS agar. One single colony was chosen from each plate and amplified using *L. crispatus*-specific primers (forward: 5'-ATTGATCGGAAGCGCAGTCT-3', reverse: 5'-CAGTTGGAGTGCCTGAAAGG-3'). PCR was performed as follows: 95 °C for 5 min, 34 cycles (95 °C for 30 s, 58 °C for 40 s, 72 °C for 40 s), and 72 °C for 5 min. The products were then subjected to 1% agarose gel electrophoresis at 120 V for 30 min. The strains with the target band were 16S-rRNA-sequenced to further identify the species. In total, 37 strains of *L. crispatus* were separated from different samples, including 17 strains from human vaginal samples, 11 from human feces, and 9 from chicken feces (Table 1). These strains were incubated in de Man, Rogosa, and Sharpe (MRS) broth for 24 h at 37 °C. After centrifugation at 8000 rpm for 5 min, the bacterial cells were washed three times with sterilized water and stored at –80 °C before sequencing. Genomic DNA was extracted using a bacterial DNA extraction kit (OMEGA, Norcross, GA, USA) according to the manufacturer's instructions. The quality of the extracted genomic DNA was determined by agarose gel electrophoresis (1% gel concentration), and the purity of the DNA was examined using an ultra-micro ultraviolet

spectrophotometer. The concentration of DNA was determined using a QubitTM4 fluorometer and Qubit[®] DNA Assay Kit (Life Technologies, USA).

Table 1. General genome features of the 38 *L. crispatus* strains.

Strains	Source	Genome Size (Mb)	GC Content (%)	CDS No.	tRNA No.	rRNA No.	Accession No.
ST1	Chicken Feces	2.04	36.9	2023	64	12	SAMEA2272191
FGSYC18L1	Human Feces	2.18	36.47	2177	37	3	SAMN12869300
FGuxi14M5	Human Feces	2.19	36.47	2191	42	1	SAMN12869301
FHNFQ15L11	Human Feces	2.26	36.48	2241	44	3	SAMN12869302
FHNFQ29L3	Human Feces	2.03	36.75	2053	56	3	SAMN12869303
FHNFQ37L2	Human Feces	2.09	36.60	2091	45	1	SAMN12869304
FHNFQ45L4	Human Feces	2.13	36.64	2190	44	3	SAMN12869305
FHNFQ56M4	Human Feces	2.23	36.52	2218	40	3	SAMN12869306
FHNXY35M5	Human Feces	2.11	36.61	2121	40	3	SAMN12869307
FHNXY45L1	Human Feces	2.06	36.62	2066	35	0	SAMN12869308
FSCDJY67L3	Human Feces	2.21	36.61	2190	45	1	SAMN12869309
FZJHZD27M9	Human Feces	2.22	36.47	2243	41	1	SAMN12869310
FAHXC2L1	Chicken Feces	2.04	36.81	2023	46	4	SAMN12869311
FCQJJ9M2	Chicken Feces	1.97	36.93	1988	44	3	SAMN12869312
FHNXY70M14	Chicken Feces	1.98	36.92	1985	45	3	SAMN12869313
FHNXY73M2	Chicken Feces	2.02	36.69	1989	49	3	SAMN12869314
FJLHD56L3	Chicken Feces	2.02	36.77	1970	38	3	SAMN12869315
FJXPY62L2	Chicken Feces	2.02	36.92	1977	41	3	SAMN12869316
FSCPS86L2	Chicken Feces	2.22	36.53	2203	49	3	SAMN12869317
QGSYC5L3	Chicken Feces	2.02	36.82	2003	45	3	SAMN12869318
QGSYC6L6	Chicken Feces	1.98	36.77	2004	43	3	SAMN12869319
QJSWX106M5	Human Vagina	2.15	36.82	2194	72	4	SAMN12869320
QJSWX109M4	Human Vagina	2.18	36.72	2231	53	3	SAMN12869321
QJSWX110M1	Human Vagina	2.17	36.72	2227	41	4	SAMN12869322
QJSWX112M2	Human Vagina	2.17	36.73	2241	53	3	SAMN12869323
QJSWX113M3	Human Vagina	2.22	36.77	2309	46	3	SAMN12869324
QJSWX120M1	Human Vagina	2.15	36.77	2206	68	4	SAMN12869325
QJSWX128M5	Human Vagina	1.94	36.73	1972	41	0	SAMN12869326
QJSWX148M3	Human Vagina	1.87	37.05	1869	43	2	SAMN12869327
QJSWX149M5	Human Vagina	1.96	37.08	1981	53	3	SAMN12869328
QJSWX150M37	Human Vagina	1.90	36.98	1914	40	3	SAMN12869329
QJSWX154M12	Human Vagina	2.03	37.1	2108	42	2	SAMN12869330
QJSWX166M13	Human Vagina	1.92	36.75	1962	42	1	SAMN12869331
QJSWX169M3	Human Vagina	2.21	36.74	2325	66	7	SAMN12869332
QJSWX174M9	Human Vagina	1.93	37.09	1969	49	3	SAMN12869333
QJSWX175M3	Human Vagina	2.12	36.79	2112	50	2	SAMN12869334
QJSWX178M1	Human Vagina	2.19	36.84	2243	71	6	SAMN12869335
QJSWX197M1	Human Vagina	2.20	36.94	2275	43	3	SAMN12869336

The study (Ethical No. KS202006) was approved by the clinical new technology and scientific research ethics committee of the Wuxi People's Hospital, Wuxi, Jiangsu province, China.

2.2. Genome Sequencing, Assembly, and Annotation

The draft genomes of *L. crispatus* were sequenced using Illumina HiSeq×10 platform (Majorbio BioTech Co, Shanghai, China). Each sample provided an amount of raw sequencing data with no less than 100× coverage depth of the genome. The genome sequences were assembled using the software SOAP denovo2 (v2.0) [21]. The protein-coding sequences (CDS) were predicted using Glimmer according to Delcher et al. [22]. The tRNAscan-SE (v2.0) [23] and Barrnap were used to predict the tRNA and rRNA of genomes, respectively. The CDS of predicted genes were annotated across the Kyoto Encyclopedia of Genes and Genomes (KEGG) database to obtain functional information [24].

2.3. Average Nucleotide Identity Calculation

The average nucleotide identity (ANI) value between any two genomes was calculated using ANI perl script. The resulting matrices were clustered and visualized using R-packages' heat map software (v1.0.8).

2.4. Pangenome and Core Genome Analysis

The pangenome and core genome of the 37 *L. crispatus* genomes were calculated using the PGAP1.21 [25], and the pangenome model was analyzed according to Heap's law.

2.5. Phylogenetic Analysis

The protein sequences of the 37 *L. crispatus* were analyzed by OrthoMCL1.4 [26]. The protein families with the same function were then clustered using the Markov cluster algorithm (MCL). Core genes and specific genes were shown in a Venn plot. The core genes were aligned using MAFFT-7.313 [27], clustering analysis was performed using the neighbor-joining (NJ) method [28], and phylogenetic trees were constructed using PHYLIP (v3.6).

2.6. Carbohydrate Active Enzyme Analysis

The strain's carbohydrate metabolism genes were annotated using the Carbohydrate Active Enzyme Database (CAZy) [29].

2.7. Statistical Analysis

Difference between two groups was evaluated using an independent samples *t*-test (SPSS 19.0), and a *p* value of < 0.05 was considered to indicate a significant difference.

3. Results

3.1. General Genome Features of the *L. crispatus* Strains

The 37 *L. crispatus* genomes ranged in size from 1.87 Mb to 2.26 Mb, with an average length of 2.09 Mb (Figure 1). Genomic sizes of vaginal-source bacteria (range 1.87–2.22 Mb, ~2.08 Mb) were similar to those of strains isolated from chicken (range 1.97–2.22 Mb, ~2.03 Mb). The genomes of the human-fecal-origin *L. crispatus* (range 2.03–2.26 Mb, ~2.16 Mb) were the largest. The number of CDS varied from 1869 to 2325, with an average of 2110 (Table 1). The GC content of the genomes was on average 36.76%, and no significant difference was found among the three sources (Figure 1). The ANI value of the 37 strains compared with the reference strain *L. crispatus* ST1 from NCBI and between any two strains was 96.27% to 99.96% (Figure 2), i.e., greater than 96%, indicating that they belonged to the same species [30]. As shown in Figure 2, the genomes of vaginal *L. crispatus* strains clustered separately from the feces isolates based on hierarchical clustering of ANI values. However, the human fecal-origin *L. crispatus* did not form a separate cluster from the chicken isolates. The feces-derived strains had a closer genetic relationship.

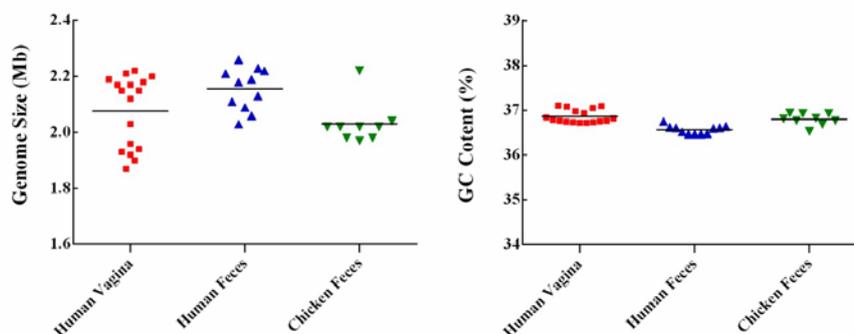


Figure 1. Genome size and GC content of *L. crispatus*. The red square, blue triangle, and green inverted triangle represent *L. crispatus* isolated from the human vagina, human feces, and chicken feces, respectively.

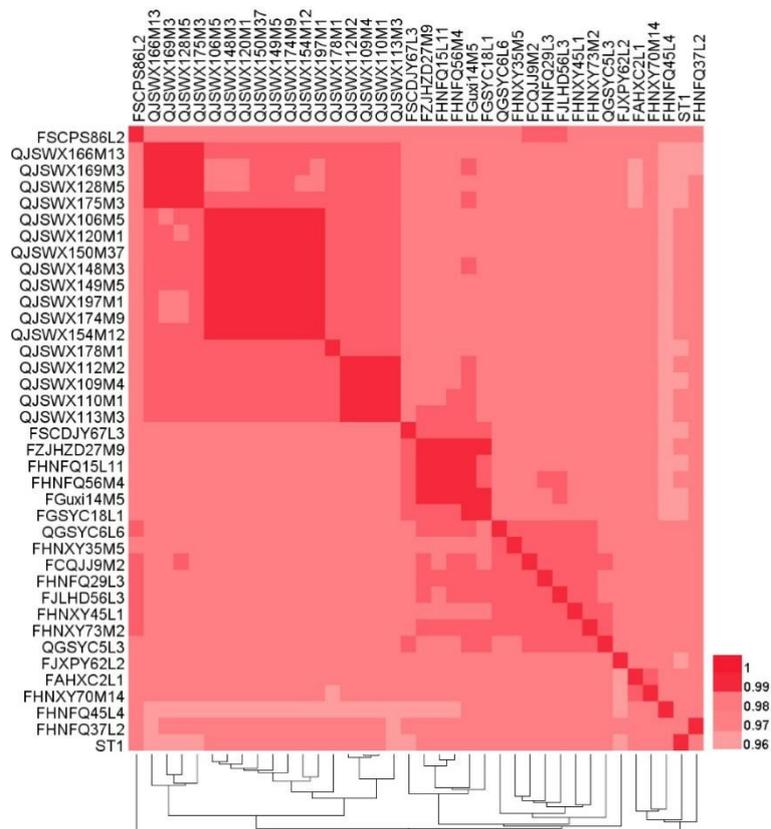


Figure 2. Heatmap showing the Average Nucleotide Identity (ANI) value among 38 *L. crispatus* strains.

3.2. Pan-genome and Core Genome Analysis of *L. crispatus*

The pan-genome of 37 strains contained 5014 genes. As the number of strains increased, the pan-genome curve gradually converged. The derived mathematical function had an index value of less than 0.5, indicating that *L. crispatus* has a closed pan-genome [16]. However, the core genome only consisted of 1250 genes. The number of core genes was saturated and the gene curve trend was stable after the first 30 strains were analyzed (Figure 3a). The results showed that the 37 strains of *L. crispatus* in this study fully represented this species. According to the result of MCL clustering, the core genes, accessory genes, and specific genes of 37 strains were obtained. An accessory gene is defined as a gene shared by at least two strains. A specific gene is a gene that only exists in a certain strain. The number of specific genes in *L. crispatus* varied from 4 to 134 (Figure 3b). The average number of specific genes of *L. crispatus* isolated from chicken and human feces was 61 and 57 respectively. However, *L. crispatus* strains isolated from vaginal samples had an average number of specific genes of only 21.

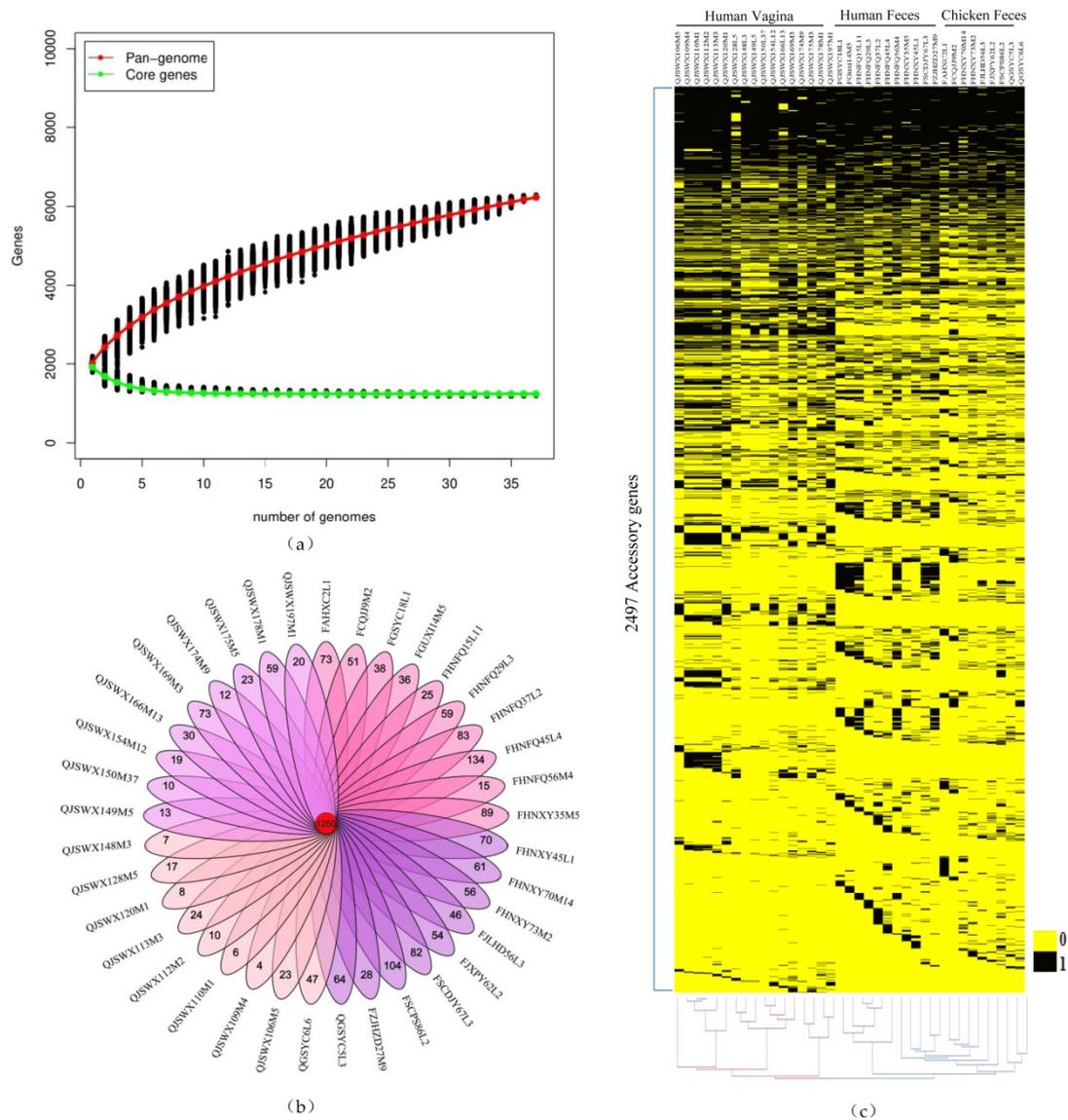


Figure 3. Pan-genome and core genome of *L. crispatus*. (a) The number of new genes accumulated in the *L. crispatus* pan-genome was plotted against the number of new genomes added and the number of accumulated genes attributed to the core genome relative to the number of added genomes. (b) Venn diagram representing the core and specific genes of 37 *L. crispatus* strains obtained by MCL clustering. The number in the middle red circle represents the number of core genes of the 37 strains. The numbers on the petals indicate the specific genes of each strain. (c) A heatmap representation of accessory genes against sequenced *L. crispatus* genomes. Each row represents an accessory gene. Each column represents a strain.

Accessory genes contribute to species diversity. They are not important for bacterial growth, but confer selective advantages such as adaptation to different niches, drug resistance, or colonization of new hosts [31]. The distribution of the 2497 accessory genes was different in the 37 strains, as visualized by heat map (Figure 3c). It was shown that the 37 *L. crispatus* strains were divided into two large branches; one branch was vaginal-derived strains and the other was strains from feces. This indicated that the accessory genes were affected by the habitat. Meanwhile, strains isolated from chicken or human feces were similar in the composition of their accessory genes.

3.3. Phylogenetic Analysis of *L. crispatus* Strains

To explore the effects of different niches on the phylogeny of the *L. crispatus* genome, 1091 single-copy orthologous genes were used to construct a phylogenetic tree (Figure 4) against the 37 strains via the neighbor-joining method. The feces-derived strains and the vagina-derived strains were each clustered into different clusters. The phylogenetic tree was divided into four branches, including Branches A, B, C, and D. All 17 vagina-isolated strains were clustered into Branch D. However, chicken and human isolates were distributed in either Branch A or Branch B. Branch C contained only six *L. crispatus* strains, all from human feces. Therefore, different habitats might relate to differences in the genetic evolution of the strains.

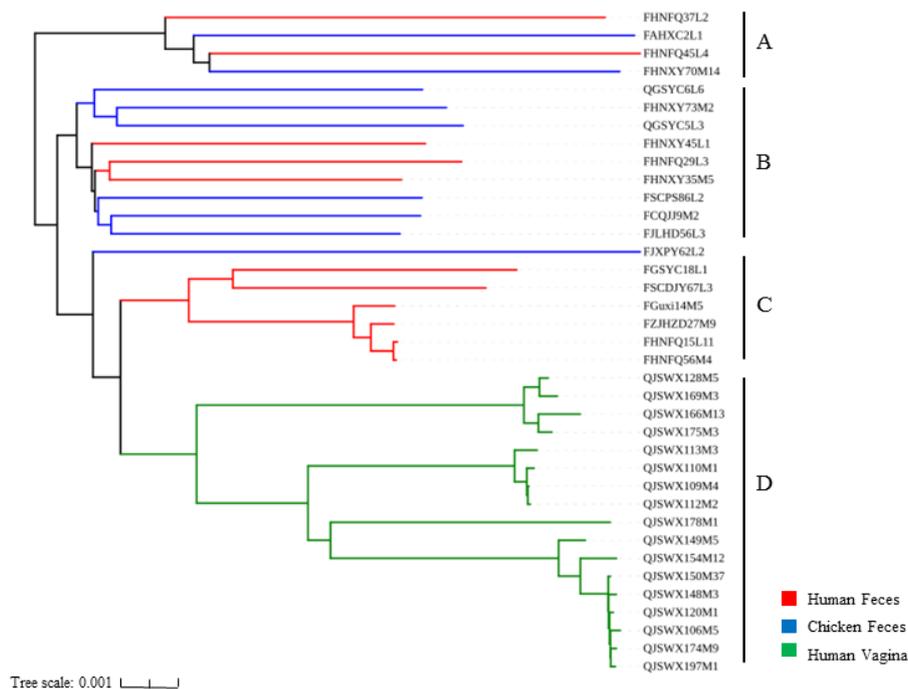


Figure 4. Phylogenetic tree of 37 *L. crispatus* strains. Based on 1091 single-copy core genes of 37 strains, the phylogenetic tree was constructed using the neighbor-joining method. According to the clustering results of the phylogenetic tree, the phylogenetic tree is divided into 4 branches, namely Branch A, B, C and D.

3.4. Evolution and Adaptation to Environment

To investigate whether *L. crispatus* showed some adaptability to different habitats in terms of gene function, we performed KEGG annotation for all 37 strains. However, based on the results of ANI value, phylogenetic tree, and the distribution of the accessory genes, we found that the feces-derived (human/chicken) strains had a high degree of similarity. Therefore, in terms of subsequent habitat adaptation analysis, we mainly focused on the adaptation of fecal-derived strains and vagina-derived strains to their respective niches. Across the two niches, genes associated with carbohydrate metabolism (~13%) and membrane transport (~11%) occupied a large proportion. The genes encoding transport and catabolism, cell motility, global and overview maps, xenobiotics biodegradation and metabolism, lipid metabolism, metabolism of cofactors and vitamins, metabolism of other amino acids, amino acid metabolism, glycan biosynthesis and metabolism, metabolism of terpenoids and polyketides, and drug resistance were annotated. Significant differences ($p < 0.05$) were observed between the two niches in genes involved with antimicrobial and environmental adaptation (Figure 5).

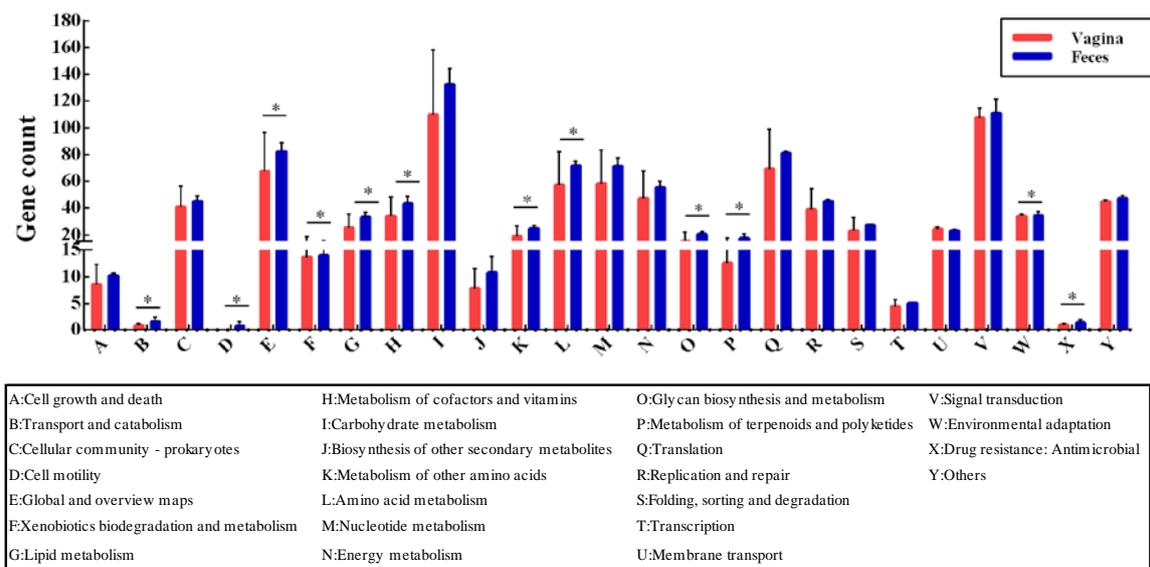


Figure 5. Annotation for genes of *L. crispatus* isolated from vagina and gut using the KEGG database. * $p < 0.05$.

To further explore genes that might be adaptive to the vaginal and intestinal environment, we analyzed the genes with significant differences ($p < 0.05$) in gene abundance of each KEGG orthology between the two sources. Meanwhile, if a gene was present in more than half of the fecal strains, but in fewer than half in the vaginal strains, this gene was considered adapted to the gut. The genes associated with vaginal adaptation were also analyzed according to the same method. Based on this criterion, 69 important genes were found. Among the feces-derived *L. crispatus* isolates, 31 genes were analyzed, including 11 fecal-specific genes, (i.e., only present in feces). Of 38 genes found in the vagina-derived strains, 7 were vagina-specific genes.

Of the 31 genes in the feces-derived strains, 22% were related to the CRISPR–Cas system, and 19% to carbohydrate metabolism (deoxyribose-phosphate aldolase, sucrose phosphorylase, α -galactosidase). Other genes were linked to biosynthesis of other secondary metabolites and to amino acid metabolism. In the 38 genes of vagina-derived *L. crispatus*, 31% were related to carbohydrate metabolism (galactitol-specific IIC component, pullulanase, hexulose-6-phosphate isomerase). In addition, there were many genes involved in redox reactions and acid tolerance in the vagina, such as iron–sulfur protein and manganese ion transporters (Tables S1 and S2).

3.5. Active Carbohydrate Enzymes

The carbohydrate-utilization-related genes occupied the largest proportion in each genome by KEGG annotation. Therefore, we further compared the number of genes encoding carbohydrate-active enzymes in strains from feces and vagina.

The 37 strains had genes encoding 33 glycoside hydrolases (GHs) families. The GH family was the most important type of carbohydrate enzyme in *L. crispatus*, accounting for 55.9%. The abundance of GH1, GH13_18, GH2, GH20, GH25, GH3, GH36, GH43_14, GH73, GH78, and GH92 differed significantly ($p < 0.05$) between strains isolated from the feces and vagina. *L. crispatus* strains of fecal origin contained more types of carbohydrate enzyme than did *L. crispatus* of vaginal origin. Twelve GH families were only present in strains derived from feces. The GH140 and GH43_4 only existed in human feces isolates, whereas GH105 were only present in strains isolated from chicken feces (Figure 6a).

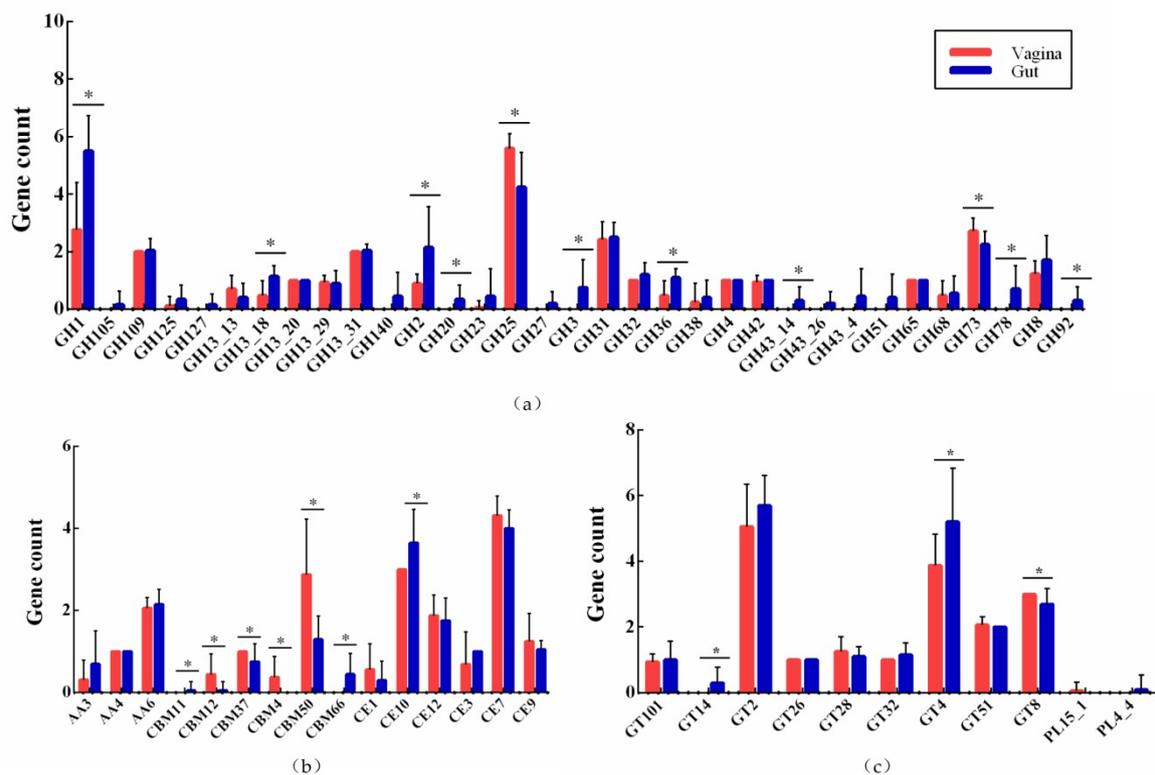


Figure 6. Distribution and abundance of active carbohydrate enzyme family genes in *L. crispatus* strains isolated from feces and vagina. * $p < 0.05$.

Additionally, nine glycosyl transferases (GTs), three auxiliary activities (AAs), six carbohydrate-binding modules (CBMs), six carbohydrate esterases (CEs), and two polysaccharide lyases (PL) were detected in 37 strains. The abundances of GT14, GT4, GT8, CE10, and CBM families were significantly different between the two groups. The CBM4 and PL15_1 were present only in vaginal strains (Figure 6b,c).

4. Discussion

Comparative genomic analysis of 37 *L. crispatus* strains from fecal and vaginal sources revealed the host's influence on the genetic characteristics and evolution of the strain, as well as the adaptability of the strain to the intestine and vagina.

The general genomic characteristics in terms of size, GC content, and number of CDS (Table 1) showed that there was no significant difference between the *L. crispatus* strains derived from the three sources. However, when compared with chicken feces and human vaginal isolates, the genome size of strains isolated from human feces was larger. A possible explanation is that the human gut placed more environmental pressure on the bacteria [32]. To survive in a complicated environment, the *L. crispatus* could have obtained additional genes through horizontal transfer [33].

The phylogenetic tree results showed that the fecal isolates and the vaginal isolates were each clustered separately (Figure 4). This distinction indicated that the niche exerted an important impact on the evolution of the *L. crispatus* [15]. This result provided the insight that we might be able to determine the niche of an unknown source *L. crispatus* based on the clustering results of strains in the phylogenetic tree. A previous study revealed that Bifidobacteria isolated from the vagina and feces had no tendency to host in the phylogenetic tree [34]. In this case, it might have been that Bifidobacteria were not the main strain in the vagina. Therefore, the effect of the source on the evolutionary relationship of Bifidobacteria was not obvious.

Compared with the gut, the vagina has a lower pH (4 ± 0.5) [35]. Low pH helps reduce the risk of gynecological diseases and maintain female reproductive health [36]. Therefore, *L. crispatus* from the vagina require more defense mechanisms to respond to the low pH. The gene encoding manganese transport protein was presented only in vagina-sourced strains. This protein can take up Mn^{2+} and expel protons out of the cell to take part in the acid response, maintaining intracellular pH homeostasis [35]. Meanwhile, Mn^{2+} can also help in protecting bacteria against oxidative stress [37].

Lactobacilli in the vagina produce high levels of hydrogen peroxide [38,39]. Hydrogen peroxide is an important antibacterial substance. It can inhibit such pathogens as *Gardnerella vaginalis* and *Neisseria gonorrhoeae* in the vagina, and help maintain the normal flora of the vagina [40]. Meanwhile, hydrogen peroxide also causes oxidative stress in *L. crispatus*. Therefore, we observed that genes associated with oxidative stress were significantly enriched in vaginally derived bacteria, as described in Table S1. The two-component system is a broad signal transduction pathway in bacteria. It plays a major role in adapting to changing environmental conditions such as cell envelope stress response, phosphate regulation, and oxidative stress [37,41]. The OmpR family significantly increases the resistance of bacteria to hydrogen peroxide [42]. The genes encoding Fe-S cluster assembly protein were more abundant in the vagina (Table S1), and these genes are critical in the catalysis of electron transfer or metabolic support reactions [43]. Meanwhile, the Suf protein complex helps to assemble or repair oxygen-labile Fe-S clusters under oxidative stress [44,45]. In addition, hydrogen peroxide could further react with some Fe^{2+} iron to result in high-activity oxidants through the Fenton reaction [46]. This highly active oxidant might have a stronger inhibitory effect on pathogens in the vagina and a more positive effect on reproductive health, but this requires further testing to verify.

Pullulanase is an enzyme that breaks down glycogen [14]. In the vagina, glycogen is the main source of carbohydrates for microorganisms [47]. Genes encoding pullulanase were more abundant in the vagina-derived *L. crispatus* strains than in strains of fecal origin (Table S1). *L. crispatus* uses this enzyme to break down glycogen and produce lactic acid, which could maintain the low pH and inhibit pathogenic bacteria in the vagina [48]. The genes coding clumping factor A and A GntR family transcription factor were only present in the vagina-derived *L. crispatus* strains (Table S1).

Due to the variety and high density of bacteria, the intestine is an ideal niche for phage survival [49]. The *L. crispatus* isolated from feces were more susceptible to phage infection. The CRISPR/Cas system is known as the prokaryotic immune system and resists the invasion of foreign genetic material such as phage viruses and foreign plasmids [50]. To protect against damage to cells by foreign DNA, more genes related to the CRISPR–Cas system were observed in feces-derived *L. crispatus* (Table S2).

Tetracycline and lincomycin resistance genes were present only in the fecal strains (Table S2). Antibiotics have been used extensively to treat diseases caused by bacterial infections in humans and animals. Correspondingly, bacteria have acquired certain resistance genes through horizontal transfer to improve antibiotic resistance and their own survival ability [51]. Gene transfer at the level of drug-resistance genes often occurs in the human gastrointestinal tract [52]. In addition, genes related to streptomycin biosynthesis, such as *rfbA*, *rfbB*, and *rfbC*, were significantly enriched in the feces-derived bacteria (Table S2). Meanwhile, some genes associated with galactose metabolism, cysteine and methionine metabolism, and arginine biosynthesis were found only in the fecal strains.

To increase our understanding of the ability of *L. crispatus* to use host-derived glycogen fermentation, the genomes of 37 *L. crispatus* were analyzed and compared using CAZymes. The feces-derived strain was enriched in various GH family enzymes, including GH140, GH20, GH3, GH43_14, GH78, and GH92, which are involved in the utilization of a variety of carbon sources (Figure 6a). This might be because the gut contains more types of carbohydrate than the vagina. *L. crispatus* would need more enzymes to break down these carbohydrates to provide them with energy. The abundance of GH2 (β -galactosidase) and GH 36 (α -galactosidase) isolates in human feces was significantly higher than human vagina and chicken feces. Milk and other dairy products are an important part of human diet. GH2 and GH36 can hydrolyze galactose in dairy products, which is beneficial to the growth of strains in the intestine [53]. However, the abundance of CBM12, CBM37,

CBM4, CBM66, GH25, and GH73 family in the vagina was higher than in the gut. CBM is a non-catalytic member of the cellulolytic enzymes. It is thought to assist in the synergy of enzymes through proximity and targeting, and plays an important role in the degradation of insoluble substrates [54,55]. Cellulose is the main component of plant cell walls. Humans and chickens consume cellulose-containing vegetables, but the feces isolates had a lower CBM abundance than the vagina (Figure 6b). This might require more in-depth exploration of the specific role of CBM in vagina-derived *L. crispatus*. GH25 has a role in facilitating cell division and in defense, as a lysozyme involved in cell wall and peptidoglycan catabolism [56]. GH73 encodes β -N-acetylglucosaminidases that cleave the β -1,4 glycosidic linkage between the N-acetylglucosamine (GlcNAc) and N-acetylmuramic acid (MurNAc) residues of bacterial cell wall peptidoglycan (PG) [57]. These two GH family lysozymes might provide a bacterial strategy to improve the competitive advantage of *L. crispatus* by destroying the cell walls of other pathogenic bacteria in the vagina and lysing them [58]. In summary, the above results demonstrate that *L. crispatus* is adaptable to different environments to improve its ability to survive.

5. Conclusions

In this study, we selected the genomes of 37 *L. crispatus* strains isolated from the human vagina, human feces, and chicken feces. The results of comparative genomics demonstrated that niche affected the evolution of *L. crispatus* to a large extent. Moreover, the *L. crispatus* strains showed adaptability to different environments to improve their own survival ability during the evolution of the strain.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4425/11/4/360/s1>, Table S1: Higher abundance of 38 genes in the vagina, Table S2: Higher abundance of 31 genes in the feces.

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

Funding: This work was supported by the National Natural Science Foundation of China (No. 31820103010, 31530056), National First-class Discipline Program of Food Science and Technology (JUFSTR20180102), and Collaborative innovation center of food safety and quality control in Jiangsu Province.

Acknowledgments: We would like to thank Weiyan Wu (Wuxi People's hospital, Wuxi, Jiangsu Province) for collecting samples and processing data.

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

References

1. Ruan, J. Bergey's manual of systematic bacteriology (second edition) volume 5 and the study of Actinomycetes systematic in China. *Wei Sheng Wu Xue Bao Acta Microbiol. Sin.* **2013**, *53*, 521–530.
2. Eslami, S.; Hadjati, J.; Motevaseli, E.; Mirzaei, R.; Bonab, S.F.; Ansaripour, B.; Khoramizadeh, M.R. Lactobacillus crispatus strain SJ-3C-US induces human dendritic cells (DCs) maturation and confers an anti-inflammatory phenotype to DCs. *Apms Acta Pathol. Microbiol. Immunol. Scand.* **2016**, *124*, 697–710. [[CrossRef](#)]
3. Tobita, K.; Yanaka, H.; Otani, H. Anti-allergic effects of Lactobacillus crispatus KT-11 strain on ovalbumin-sensitized BALB/c mice. *Anim. Sci. J.* **2010**, *81*, 699–705. [[CrossRef](#)] [[PubMed](#)]
4. Azam, R.; Ghafouri-Fard, S.; Tabrizi, M.; Modarressi, M.H.; Ebrahimzadeh-Vesal, R.; Daneshvar, M.; Mobasheri, M.B.; Motevaseli, E. Lactobacillus acidophilus and Lactobacillus crispatus culture supernatants downregulate expression of cancer-testis genes in the MDA-MB-231 cell line. *Asian Pac. J. Cancer Prev.* **2014**, *15*, 4255–4259. [[CrossRef](#)] [[PubMed](#)]
5. Humphries, C. Detecting diversity. *Nature* **2017**, *550*, S12–S14. [[CrossRef](#)] [[PubMed](#)]
6. Wang, S.; Wang, Q.Y.; Yang, E.C.; Yan, L.; Li, T.; Zhuang, H. Antimicrobial compounds produced by vaginal lactobacillus crispatus are able to strongly inhibit candida albicans growth, hyphal formation and regulate virulence-related gene expressions. *Front. Microbiol.* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]

7. Antonio, M.A.; Petrina, M.A.; Meyn, L.A.; Hillier, S.L. Women colonised by *Lactobacillus crispatus* have a lower risk of acquisition of bacterial vaginosis (Bv) than women colonised by other lactobacilli. *Sex. Transm. Infect.* **2013**, *89*, A232. [[CrossRef](#)]
8. Li, T.; Liu, Z.; Zhang, X.; Chen, X.; Wang, S. Local probiotic *Lactobacillus crispatus* and *Lactobacillus delbrueckii* exhibit strong antifungal effects against vulvovaginal candidiasis in a rat model. *Front. Microbiol.* **2019**, *10*, 1033. [[CrossRef](#)]
9. Bohbot, J.M.; Darai, E.; Bretelle, F.; Brami, G.; Daniel, C.; Cardot, J.M. Efficacy and safety of vaginally administered lyophilized *Lactobacillus crispatus* IP 174178 in the prevention of bacterial vaginosis recurrence. *J. Gynecol. Obstet. Hum. Reprod.* **2018**, *47*, 81–86. [[CrossRef](#)]
10. Stapleton, A.E.; Au-Yeung, M.; Hooton, T.M.; Fredricks, D.N.; Roberts, P.L.; Czaja, C.A.; Yarova-Yarovaya, Y.; Fiedler, T.; Cox, M.; Stamm, W.E. Randomized, placebo-controlled phase 2 trial of a *Lactobacillus crispatus* probiotic given intravaginally for prevention of recurrent urinary tract infection. *Clin. Infect. Dis.* **2011**, *52*, 1212–1217. [[CrossRef](#)]
11. Binnewies, T.T.; Motro, Y.; Hallin, P.F.; Lund, O.; Dunn, D.; La, T.; Hampson, D.J.; Bellgard, M.; Wassenaar, T.M.; Ussery, D.W. Ten years of bacterial genome sequencing: Comparative-genomics-based discoveries. *Funct. Integr. Genom.* **2006**, *6*, 165–185. [[CrossRef](#)] [[PubMed](#)]
12. Field, D.; Wilson, G.; van der Gast, C. How do we compare hundreds of bacterial genomes? *Curr. Opin. Microbiol.* **2006**, *9*, 499–504. [[CrossRef](#)] [[PubMed](#)]
13. Ojala, T.; Kankainen, M.; Castro, J.; Cerca, N.; Edelman, S.; Westerlund-Wikstrom, B.; Paulin, L.; Holm, L.; Auvinen, P. Comparative genomics of *Lactobacillus crispatus* suggests novel mechanisms for the competitive exclusion of *Gardnerella vaginalis*. *BMC Genom.* **2014**, *15*. [[CrossRef](#)] [[PubMed](#)]
14. van der Veer, C.; Hertzberger, R.Y.; Bruisten, S.M.; Tytgat, H.L.P.; Swanenburg, J.; Angelino-Bart, A.D.; Schuren, F.; Molenaar, D.; Reid, G.; de Vries, H.; et al. Comparative genomics of human *Lactobacillus crispatus* isolates reveals genes for glycosylation and glycogen degradation: Implications for in vivo dominance of the vaginal microbiota. *Microbiome* **2019**, *7*. [[CrossRef](#)] [[PubMed](#)]
15. Yu, J.; Zhao, J.; Song, Y.; Zhang, J.; Yu, Z.; Zhang, H.; Sun, Z. Comparative genomics of the herbivore gut symbiont *Lactobacillus reuteri* reveals genetic diversity and lifestyle adaptation. *Front. Microbiol.* **2018**, *9*, 1151. [[CrossRef](#)] [[PubMed](#)]
16. Lee, J.Y.; Han, G.G.; Kim, E.B.; Choi, Y.J. Comparative genomics of *Lactobacillus salivarius* strains focusing on their host adaptation. *Microbiol. Res.* **2017**, *205*, 48–58. [[CrossRef](#)]
17. Eisenbach, L.; Janssen, D.; Ehrmann, M.A.; Vogel, R.F. Comparative genomics of *Lactobacillus curvatus* enables prediction of traits relating to adaptation and strategies of assertiveness in sausage fermentation. *Int. J. Food Microbiol.* **2018**, *286*, 37–47. [[CrossRef](#)]
18. Feyereisen, M.; Mahony, J.; Kelleher, P.; Roberts, R.J.; O’Sullivan, T.; Geertman, J.M.A.; van Sinderen, D. Comparative genome analysis of the *Lactobacillus brevis* species. *BMC Genom.* **2019**, *20*. [[CrossRef](#)]
19. Pan, M.C.; Hidalgo-Cantabrana, C.; Goh, Y.J.; Sanozky-Dawes, R.; Barrangou, R. Comparative analysis of *Lactobacillus gasseri* and *Lactobacillus crispatus* isolated from human urogenital and gastrointestinal tracts. *Front. Microbiol.* **2020**, *10*. [[CrossRef](#)]
20. Beasley, S.S.; Takala, T.M.; Reunanen, J.; Apajalahti, J.; Saris, P.E.J. Characterization and electrotransformation of *Lactobacillus crispatus* isolated from chicken crop and intestine. *Poult. Sci.* **2004**, *83*, 45–48. [[CrossRef](#)]
21. Luo, R.B.; Liu, B.H.; Xie, Y.L.; Li, Z.Y.; Huang, W.H.; Yuan, J.Y.; He, G.Z.; Chen, Y.X.; Pan, Q.; Liu, Y.J.; et al. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler. *Gigascience* **2012**, *1*. [[CrossRef](#)] [[PubMed](#)]
22. Delcher, A.L.; Bratke, K.A.; Powers, E.C.; Salzberg, S.L. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **2007**, *23*, 673–679. [[CrossRef](#)] [[PubMed](#)]
23. Besemer, J.; Borodovsky, M. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res.* **2005**, *33*, W451–W454. [[CrossRef](#)] [[PubMed](#)]
24. Kanehisa, M.; Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **2000**, *28*, 27–30. [[CrossRef](#)]
25. Zhao, Y.B.; Wu, J.Y.; Yang, J.H.; Sun, S.X.; Xiao, J.F.; Yu, J. PGAP: Pan-genomes analysis pipeline. *Bioinformatics* **2012**, *28*, 416–418. [[CrossRef](#)]
26. Li, L.; Stoeckert, C.J.; Roos, D.S. OrthoMCL: Identification of ortholog groups for eukaryotic genomes. *Genome Res.* **2003**, *13*, 2178–2189. [[CrossRef](#)]

27. Katoh, K.; Rozewicki, J.; Yamada, K.D. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* **2019**, *20*, 1160–1166. [[CrossRef](#)]
28. Kumar, S.; Gadagkar, S.R. Efficiency of the neighbor-joining method in reconstructing deep and shallow evolutionary relationships in large phylogenies. *J. Mol. Evol.* **2000**, *51*, 544–553. [[CrossRef](#)]
29. Wang, S.; Yang, B.; Ross, R.P.; Stanton, C.; Zhao, J.X.; Zhang, H.; Chen, W. Comparative genomics analysis of lactobacillus ruminis from different niches. *Genes* **2020**, *11*, 70. [[CrossRef](#)]
30. Ciufu, S.; Kannan, S.; Sharma, S.; Badretdin, A.; Clark, K.; Turner, S.; Brover, S.; Schoch, C.L.; Kimchi, A.; DiCuccio, M. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI. *Int. J. Syst. Evol. Microbiol.* **2018**, *68*, 2386–2392. [[CrossRef](#)]
31. Medini, D.; Donati, C.; Tettelin, H.; Massignani, V.; Rappuoli, R. The microbial pan-genome. *Curr. Opin. Genet. Dev.* **2005**, *15*, 589–594. [[CrossRef](#)] [[PubMed](#)]
32. O'Sullivan, O.; O'Callaghan, J.; Sangrador-Vegas, A.; McAuliffe, O.; Slattery, L.; Kaleta, P.; Callanan, M.; Fitzgerald, G.F.; Ross, R.P.; Beresford, T. Comparative genomics of lactic acid bacteria reveals a niche-specific gene set. *BMC Microbiol.* **2009**, *9*, 50. [[CrossRef](#)] [[PubMed](#)]
33. Broadbent, J.R.; Neeno-Eckwall, E.C.; Stahl, B.; Tandee, K.; Cai, H.; Morovic, W.; Horvath, P.; Heidenreich, J.; Perna, N.T.; Barrangou, R.; et al. Analysis of the Lactobacillus casei supragenome and its influence in species evolution and lifestyle adaptation. *BMC Genom.* **2012**, *13*, 533. [[CrossRef](#)] [[PubMed](#)]
34. Freitas, A.C.; Hill, J.E. Bifidobacteria isolated from vaginal and gut microbiomes are indistinguishable by comparative genomics. *PLoS ONE* **2018**, *13*, e0196290. [[CrossRef](#)] [[PubMed](#)]
35. Shabayek, S.; Spellerberg, B. Acid stress response mechanisms of group B streptococci. *Front. Cell. Infect. Microbiol.* **2017**, *7*, 395. [[CrossRef](#)] [[PubMed](#)]
36. Miller, E.A.; Beasley, D.E.; Dunn, R.R.; Archie, E.A. Lactobacilli dominance and vaginal pH: Why is the human vaginal microbiome unique? *Front. Microbiol.* **2016**, *7*, 1936. [[CrossRef](#)]
37. Cotter, P.D.; Hill, C. Surviving the acid test: Responses of gram-positive bacteria to low pH. *Microbiol. Mol. Biol. Rev. MMBR* **2003**, *67*, 429–453. [[CrossRef](#)]
38. Xu, H.Y.; Tian, W.H.; Wan, C.X.; Jia, L.J.; Wang, L.Y.; Yuan, J.; Liu, C.M.; Zeng, M.; Wei, H. Antagonistic potential against pathogenic microorganisms and hydrogen peroxide production of indigenous lactobacilli isolated from vagina of Chinese pregnant women. *Biomed. Environ. Sci.* **2008**, *21*, 365–371. [[CrossRef](#)]
39. Balkus, J.E.; Mitchell, C.; Agnew, K.; Liu, C.; Fiedler, T.; Cohn, S.E.; Luque, A.; Coombs, R.; Fredricks, D.N.; Hitti, J. Detection of hydrogen peroxide-producing Lactobacillus species in the vagina: A comparison of culture and quantitative PCR among HIV-1 seropositive women. *BMC Infect. Dis.* **2012**, *12*, 188. [[CrossRef](#)]
40. Pascual, L.M.; Daniele, M.B.; Pajaro, C.; Barberis, L. Lactobacillus species isolated from the vagina: Identification, hydrogen peroxide production and nonoxynol-9 resistance. *Contraception* **2006**, *73*, 78–81. [[CrossRef](#)]
41. Stock, A.M.; Robinson, V.L.; Goudreau, P.N. Two-component signal transduction. *Annu. Rev. Biochem.* **2000**, *69*, 183–215. [[CrossRef](#)] [[PubMed](#)]
42. Zhao, Z.; Peng, T.; Oh, J.I.; Glaeser, J.; Weber, L.; Li, Q.; Klug, G. A response regulator of the OmpR family is part of the regulatory network controlling the oxidative stress response of Rhodobacter sphaeroides. *Environ. Microbiol. Rep.* **2019**, *11*, 118–128. [[CrossRef](#)]
43. Kiley, P.J.; Beinert, H. The role of Fe-S proteins in sensing and regulation in bacteria. *Curr. Opin. Microbiol.* **2003**, *6*, 181–185. [[CrossRef](#)]
44. Selbach, B.P.; Chung, A.H.; Scott, A.D.; George, S.J.; Cramer, S.P.; Dos Santos, P.C. Fe-S cluster biogenesis in Gram-positive bacteria: SufU is a zinc-dependent sulfur transfer protein. *Biochemistry* **2014**, *53*, 152–160. [[CrossRef](#)] [[PubMed](#)]
45. Nachin, L.; Loiseau, L.; Expert, D.; Barras, F. SufC: An unorthodox cytoplasmic ABC/ATPase required for [Fe-S] biogenesis under oxidative stress. *EMBO J.* **2003**, *22*, 427–437. [[CrossRef](#)]
46. Maresca, D.; Zotta, T.; Mauriello, G. Adaptation to aerobic environment of lactobacillus johnsonii/gasseri strains. *Front. Microbiol.* **2018**, *9*, 157. [[CrossRef](#)]
47. Mirmonsef, P.; Hotton, A.L.; Gilbert, D.; Gioia, C.J.; Maric, D.; Hope, T.J.; Landay, A.L.; Spear, G.T. Glycogen levels in undiluted genital fluid and their relationship to vaginal pH, estrogen, and progesterone. *PLoS ONE* **2016**, *11*, e0153553. [[CrossRef](#)]
48. Tachedjian, G.; Aldunate, M.; Bradshaw, C.S.; Cone, R.A. The role of lactic acid production by probiotic Lactobacillus species in vaginal health. *Res. Microbiol.* **2017**, *168*, 782–792. [[CrossRef](#)]

49. Stern, A.; Mick, E.; Tirosh, I.; Sagy, O.; Sorek, R. CRISPR targeting reveals a reservoir of common phages associated with the human gut microbiome. *Genome Res.* **2012**, *22*, 1985–1994. [[CrossRef](#)]
50. Barrangou, R.; Marraffini, L.A. CRISPR-Cas systems: Prokaryotes upgrade to adaptive immunity. *Mol. Cell* **2014**, *54*, 234–244. [[CrossRef](#)]
51. Zhang, F.; Gao, J.; Wang, B.; Huo, D.; Wang, Z.; Zhang, J.; Shao, Y. Whole-genome sequencing reveals the mechanisms for evolution of streptomycin resistance in *Lactobacillus plantarum*. *J. Dairy Sci.* **2018**, *101*, 2867–2874. [[CrossRef](#)] [[PubMed](#)]
52. McKeegan, K.S.; Borges-Walmsley, M.I.; Walmsley, A.R. The structure and function of drug pumps: An update. *Trends Microbiol.* **2003**, *11*, 21–29. [[CrossRef](#)]
53. Miguez Amil, S.; Jimenez-Ortega, E.; Ramirez-Escudero, M.; Talens-Perales, D.; Marin-Navarro, J.; Polaina, J.; Sanz-Aparicio, J.; Fernandez-Leiro, R. The cryo-EM structure of *thermotoga maritima* beta-galactosidase: Quaternary structure guides protein engineering. *ACS Chem. Biol.* **2020**, *15*, 179–188. [[CrossRef](#)] [[PubMed](#)]
54. Novy, V.; Aissa, K.; Nielsen, F.; Straus, S.K.; Ciesielski, P.; Hunt, C.G.; Saddler, J. Quantifying cellulose accessibility during enzyme-mediated deconstruction using 2 fluorescence-tagged carbohydrate-binding modules. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 22545–22551. [[CrossRef](#)] [[PubMed](#)]
55. Duan, C.J.; Huang, M.Y.; Pang, H.; Zhao, J.; Wu, C.X.; Feng, J.X. Characterization of a novel theme C glycoside hydrolase family 9 cellulase and its CBM-chimeric enzymes. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 5723–5737. [[CrossRef](#)] [[PubMed](#)]
56. O’Riordan, N.; O’Callaghan, J.; Butto, L.F.; Kilcoyne, M.; Joshi, L.; Hickey, R.M. Bovine glycomacropeptide promotes the growth of *Bifidobacterium longum* ssp. *infantis* and modulates its gene expression. *J. Dairy Sci.* **2018**, *101*, 6730–6741. [[CrossRef](#)]
57. Zaloba, P.; Bailey-Elkin, B.A.; Derksen, M.; Mark, B.L. Structural and biochemical insights into the peptidoglycan hydrolase domain of FlgJ from *Salmonella typhimurium*. *PLoS ONE* **2016**, *11*. [[CrossRef](#)]
58. Morais, S.; Cockburn, D.W.; Ben-David, Y.; Koropatkin, N.M.; Martens, E.C.; Duncan, S.H.; Flint, H.J.; Mizrahi, I.; Bayer, E.A. Lysozyme activity of the *Ruminococcus champanellensis* cellulosome. *Environ. Microbiol.* **2016**, *18*, 5112–5122. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).