



# UNIVERSITÀ DI PARMA

## ARCHIVIO DELLA RICERCA

University of Parma Research Repository

Probiogenomics analysis of 97 lactobacillus crispatus strains as a tool for the identification of promising next-generation probiotics

This is the peer reviewed version of the following article:

*Original*

Probiogenomics analysis of 97 lactobacillus crispatus strains as a tool for the identification of promising next-generation probiotics / Fontana, F.; Alessandri, G.; Lugli, G. A.; Mancabelli, L.; Longhi, G.; Anzalone, R.; Viappiani, A.; Ventura, M.; Turrone, F.; Milani, C.. - In: MICROORGANISMS. - ISSN 2076-2607. - 9:1(2021), pp. 1-14. [10.3390/microorganisms9010073]

*Availability:*

This version is available at: 11381/2886620 since: 2021-02-16T15:47:22Z

*Publisher:*

MDPI AG

*Published*

DOI:10.3390/microorganisms9010073

*Terms of use:*

Anyone can freely access the full text of works made available as "Open Access". Works made available

*Publisher copyright*

note finali coverpage

(Article begins on next page)

13 August 2025



## Article

# Probiogenomics Analysis of 97 *Lactobacillus crispatus* Strains as a Tool for the Identification of Promising Next-Generation Probiotics

Federico Fontana <sup>1,2</sup>, Giulia Alessandri <sup>1</sup>, Gabriele Andrea Lugli <sup>1</sup> , Leonardo Mancabelli <sup>1</sup>, Giulia Longhi <sup>1,2</sup>, Rosaria Anzalone <sup>2</sup>, Alice Viappiani <sup>2</sup>, Marco Ventura <sup>1,3</sup>, Francesca Turroni <sup>1,3,\*</sup> and Christian Milani <sup>1,3,\*</sup>

- <sup>1</sup> Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, 43124 Parma, Italy; federico.fontana1@unipr.it (F.F.); giulia.alessandri@unipr.it (G.A.); gabrieleandrea.lugli@unipr.it (G.A.L.); leonardo.mancabelli@unipr.it (L.M.); giulia.longhi@unipr.it (G.L.); marco.ventura@unipr.it (M.V.)
- <sup>2</sup> GenProbio s.r.l., 43124 Parma, Italy; rosaria.anzalone@genprobio.com (R.A.); alice.viappiani@genprobio.com (A.V.)
- <sup>3</sup> Microbiome Research Hub, University of Parma, 43124 Parma, Italy
- \* Correspondence: francesca.turroni@unipr.it (F.T.); christian.milani@unipr.it (C.M.); Tel.: +39-521-905666 (F.T.); +39-521-904785 (C.M.); Fax: +39-521-905604 (F.T.)
- † These authors contributed equally.

**Abstract:** Members of the genus *Lactobacillus* represent the most common colonizers of the human vagina and are well-known for preserving vaginal health and contrasting the colonization of opportunistic pathogens. Remarkably, high abundance of *Lactobacillus crispatus* in the vaginal environment has been linked to vaginal health, leading to the widespread use of many *L. crispatus* strains as probiotics. Nevertheless, despite the scientific and industrial relevance of this species, a comprehensive investigation of the genomics of *L. crispatus* taxon is still missing. For this reason, we have performed a comparative genomics analysis of 97 *L. crispatus* strains, encompassing 16 strains sequenced in the framework of this study alongside 81 additional publicly available genome sequences. Thus, allowing the dissection of the *L. crispatus* pan-genome and core-genome followed by a comprehensive phylogenetic analysis based on the predicted core genes that revealed clustering based on ecological origin. Subsequently, a genomics-targeted approach, i.e., probiogenomics analysis, was applied for in-depth analysis of the eight *L. crispatus* strains of human origin sequenced in this study. In detail their genetic repertoire was screened for strain-specific genes responsible for phenotypic features that may guide the identification of optimal candidates for next-generation probiotics. The latter includes bacteriocin production, carbohydrates transport and metabolism, as well as a range of features that may be responsible for improved ecological fitness. In silico results regarding the genetic repertoire involved in carbohydrate metabolism were also validated by growth assays on a range of sugars, leading to the selection of putative novel probiotic strains.

**Keywords:** comparative genomics; *Lactobacillus crispatus*; probiogenomics and vaginal microbiota



**Citation:** Fontana, F.; Alessandri, G.; Lugli, G.A.; Mancabelli, L.; Longhi, G.; Anzalone, R.; Viappiani, A.; Ventura, M.; Turroni, F.; Milani, C. et al. Probiogenomics Analysis of 97 *Lactobacillus crispatus* Strains as a Tool for the Identification of Promising Next-Generation Probiotics. *Microorganisms* **2021**, *9*, 73. <https://doi.org/10.3390/microorganisms9010073>

Received: 18 November 2020  
Accepted: 26 December 2020  
Published: 30 December 2020

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



**Copyright:** © 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 (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The host-associated microbiota is considered a primary contributor to the healthy status of many human compartments, such as the gastrointestinal tract (GIT) and the urogenital tract [1]. Regarding the latter, recent literature revealed that the vaginal microbiota is fundamental in both biological and chemical homeostasis [2]. In this context, it is well-known that bacteria harbored by the vaginal environment participate in maintaining a low pH, which on average should range between 3.8 and 4.5 in a healthy condition [3]. This supports the presence of a few dominant bacterial species resistant to low pH environments and concurs in reducing the colonization by pathogenic species [3]. Among the dominant bacterial genera observed in the vaginal microbiota, the most prevalent are LAB

bacteria, particularly the genus *Lactobacillus*, represented by the species *Lactobacillus iners*, *Lactobacillus crispatus*, *Lactobacillus jensenii*, and *Lactobacillus gasseri* [4]. Interestingly, in addition to the release of lactic acid and the acidification of the surrounding environment, these species have been shown to produce bacteriocins that counteract colonization of opportunistic pathogens [5].

An overview of the average vaginal microbiota composition in healthy women also revealed that the presence of specific dominant bacteria is associated with recurrent taxonomic profiles, known as community state types (CSTs) [6]. Five different CSTs have been identified in the literature, four of which (CST 1, 2, 3, and 5) showed a prevalent *Lactobacillus* composition, with the dominance of the species *L. crispatus*, *L. jensenii*, *L. gasseri*, and *L. iners*, respectively [6]. Remarkably, the CST 1, which is characterized by the dominance of the species *L. crispatus*, is known for its positive correlation with a healthy vaginal environment [6].

Nevertheless, the vaginal environment is often subject to dysbiosis, i.e., alteration of the taxonomic composition, due to the natural susceptibility of this environment to urogenital infections, or so-called bacterial vaginosis (BV). This condition is caused by opportunistic pathogens such as *Gardnerella vaginalis* or members of the *Prevotella* and *Atopobium* genera, which constitute the CST 4 [6]. Because of the relevance of *L. crispatus* in participating in vaginal health, many strains of this species have been industrially exploited as probiotics [7].

Concurrently, the species *L. crispatus* has been the subject of a broad body of literature focusing on dissecting their roles in human health through interaction with the host or competition with (opportunistic) pathogens, including interaction with VK2/E6E7 human vaginal epithelial cells and bacteriocin production [7–10]. Furthermore, in recent years, the degradative metabolism of *L. crispatus* toward glycogen, and the genetic differences between strains of human origin isolated from gut or vagina have also been investigated. The results of these studies highlighted some characteristics of *L. crispatus*, such as mechanisms of resistance and adaptation to the vaginal environment, leading to the growth of biotechnological and industrial interest toward the use of this species as a probiotic supplement [11,12].

Nevertheless, probiogenomics, i.e., the study of probiotic strains' genetic repertoire, still have not been applied to identify putative novel *L. crispatus* of probiotic relevance through screening of human vagina isolates and their genetic comparison with publicly available *L. crispatus* genomes.

For this reason, in this study we performed an in-depth comparative genomic analysis between 97 different *L. crispatus* strains, including eight novel isolates from the vagina of healthy women and eight additional strains isolated from other hosts. Thus, leading to a thorough pan-genome analysis of *L. crispatus* complemented by a detailed functional dissection of the strain-specific genetic features. Such investigations were further extended by predicting bacteriocins-encoding genes and by an *in-silico* analysis of the glycobiomes of *L. crispatus* genomes that were validated by growth assays.

## 2. Materials and Methods

### 2.1. Genome Sequencing

For the purpose of this study, a total of 16 *L. crispatus* isolates were submitted to genome sequencing. DNA extracted from these isolates was subjected to whole-genome sequencing using MiSeq (San Diego, CA, USA) at GenProbio srl (Parma, Italy) according to the supplier's protocol (San Diego, CA, USA). Fastq files of the paired-end reads obtained from targeted genome sequencing of isolated strains were utilized as input for genome assemblies through the MEGAnnotator pipeline [13]. SPAdes software was used for de novo assembly of each genome sequence [14], while protein-encoding ORFs were predicted using Prodigal [15].

Human participants gave their informed written consent before enrollment. All investigations were carried out following the principles of the Declaration of Helsinki.

## 2.2. Comparative Genomics

A pangenome calculation was performed using the pan-genome analysis pipeline PGAP [16], including each *L. crispatus* genome decoded in the framework of this study. Each predicted proteome of a given *L. crispatus* strain was screened for orthologues against the proteome of every collected *L. crispatus* strain by means of BLAST analysis (cutoff, E value of  $<1 \times 10^{-4}$  and 50% identity over at least 80% of both protein sequences) [17]. The resulting output was then clustered into protein families by means of MCL (graph theory-based Markov clustering algorithm, using the gene family method [17]). A pangenome profile was built using all possible BLAST combinations for each genome being sequentially added. Using this approach, unique protein families encoded by the analyzed *L. crispatus* genomes were also identified. Protein families shared between analyzed genomes allowed us to identify the core genome of the *L. crispatus* species. Each set of orthologous proteins, belonging to the core genome, was aligned using Mafft software [18], and phylogenetic trees were constructed using ClustalW [19]. Based on these comparative analyses, a *L. crispatus* supertree was constructed and visualized using FigTree (<http://tree.bio.ed.ac.uk/software/figtree/>).

Using InterProScan [20], RapSearch2 [21] and NCBI CDD analysis [22], we also evaluated the presence of functional domains conserved within the unique genes present in the eight strains of *L. crispatus* isolated from human vagina present in our *L. crispatus* bacterial stock.

## 2.3. Carbohydrate Growth Assays

*Lactobacillus crispatus* strains were cultivated on semisynthetic MRS medium without glucose supplemented with a 1% (wt/vol) concentration of a particular sugar, and the optical densities (measured at a wavelength of 600 nm) were recorded using a plate reader (BioTek, Winooski, VT, USA). Tested sugars are arabinose, cellobiose, fructose, fucose, galactose, glycogen, glucose, inulin, maltodextrin, mannitol, N-acetyl-D-glucosamine, N-acetyl-D-galactosamine, mannose, melibiose, maltose, raffinose, rhamnose, ribose, sorbitol, starch, sucrose, threulose, turanose, and xylose (Figure 3). The plate reader was read in intermittent mode, with absorbance readings performed at 3-min intervals for three times after 24 h and 48 h of growth, where each reading was ahead of 30 s of shaking at medium speed. Cultures were grown in biologically independent triplicates, and the resulting growth data were expressed as the means of these replicates. Carbohydrates were purchased from Sigma and Carbosynth (Berkshire, UK). Carbohydrate-active enzymes were identified based on similarity to the carbohydrate-active enZYmes (CAZy) database entries [23].

## 2.4. Data Availability

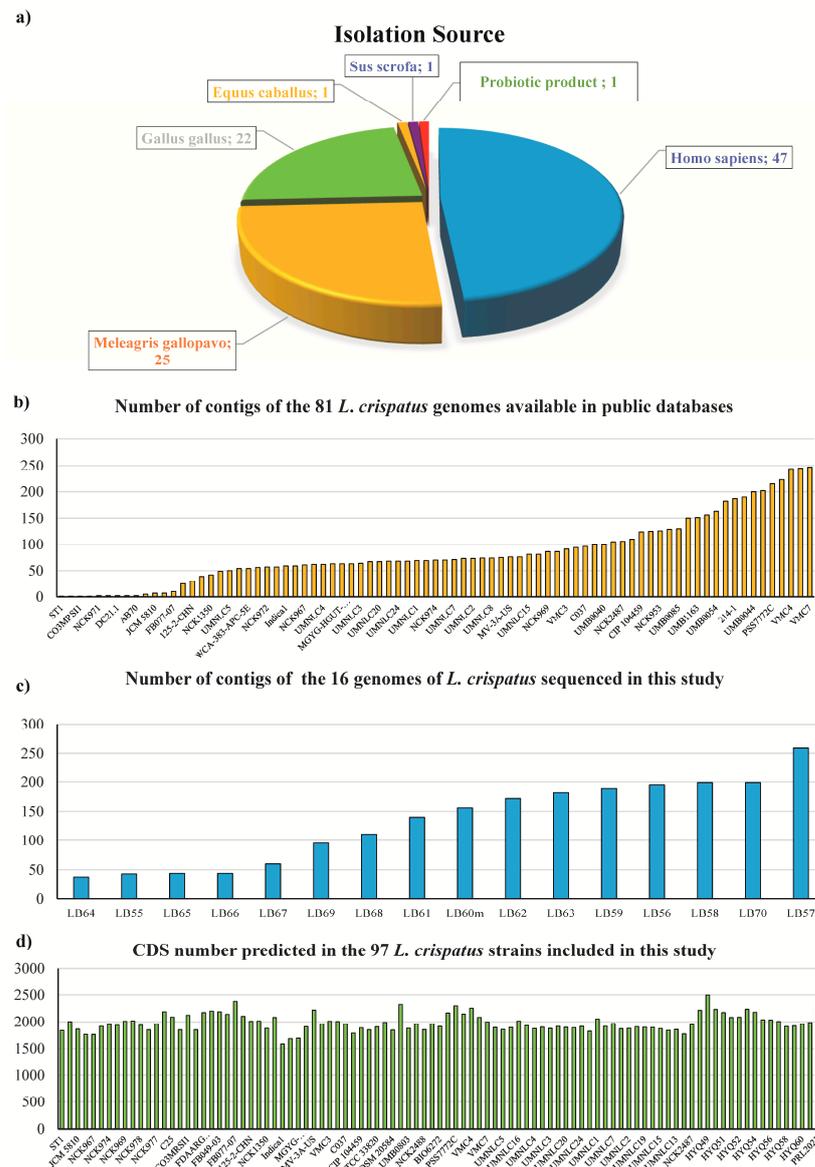
The accession numbers of the genomic sequence of the 16 strains sequenced in this study are reported in Table S1.

## 3. Results and Discussion

### 3.1. General Genome Features of *L. crispatus* Strains Included in This Study

A pool of 47 bacterial strains belonging to the *L. crispatus* species isolated from the human vaginal environment was used to perform a comparative genomics analysis. To dissect the specific genetic repertoire of strains isolated from the human vaginal tract, 50 additional *L. crispatus* strains isolated from other hosts, including *Sus scrofa*, *Equus caballus*, *Meleagris gallopavo* and *Gallus gallus*, were included in the analysis as an outgroup for comparative purposes (Table S1). As shown in Table S1, 16 of these strains were derived from our bacterial collection, of which eight were isolated from healthy women, one corresponding to a commercially available probiotic product, and seven isolated from poultry. Moreover, a total of 81 genomes were retrieved from those publicly deposited on NCBI after selecting those fragments in less than 250 contigs (Table S1, Figure 1). In addition, the average length of the genomes is 2120 Mb, with the smaller one

with length of 1.641 Mb, thus ensuring adequate coverage with respect to the reference genome GCA\_008694205.1, whose length consists of 2.2199 Mb (Table S1). Notably, poultry represents the most common ecological niche for publicly available genomes of *L. crispatus* along with human vagina (Table S1). In fact, 15 out of 81 genomes retrieved from NCBI are originated from the gut of chickens, 25 from the gut of turkeys, two derived respectively from pig and horse while 39 from the human vagina (Table S1, Figure 1). In order to obtain a comprehensive overview of the genetic variability of *L. crispatus* and identify the unique features of the strains of human origin, all the publicly available strains were included in this study.



**Figure 1.** Graphic representation of the general genome features of strains included in the comparative genomics analysis. Panel (a) shows the subdivision of all the analyzed samples, through a pie chart, based on the isolation source. Panel (b) reports a bar plot illustrating the number of contigs of the *L. crispatus* genomes downloaded from public databases. Panel (c) shows a bar plot reporting the number of contigs of the 16 genomes of *L. crispatus* sequenced in this study. Panel (d) illustrate, by means of a bar plot, the number of CDS detected within each of the 97 strains of *L. crispatus* analyzed in this study.

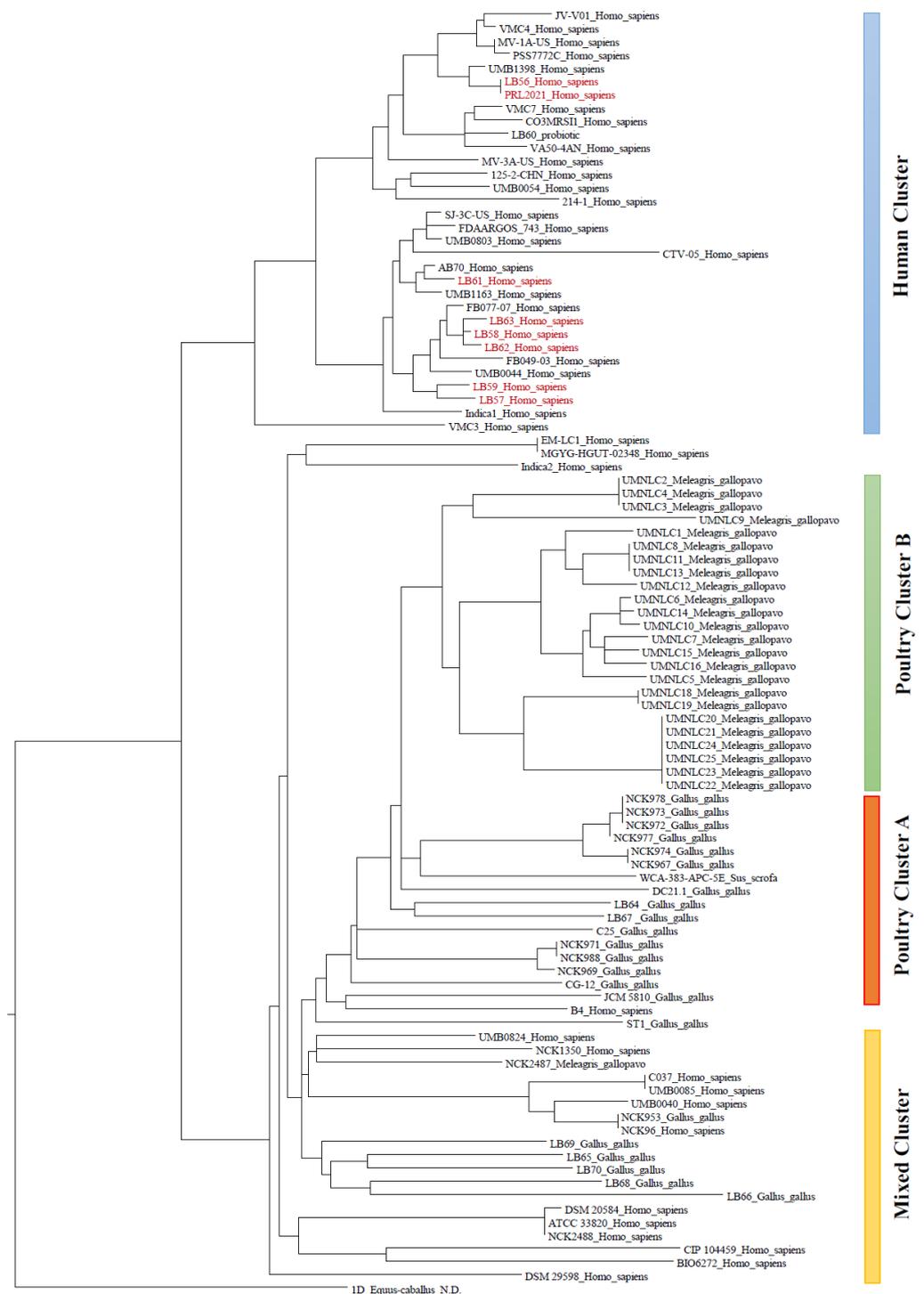
The 16 *L. crispatus* strains whose genomes were decoded in this study resulted in an average of 84 contigs per genome, ranging between 36 and 259 contigs, as well as an average number of CDS equal to 2118, ranging between 1921 and 2496 (Table S1, Figure 1). Moreover, *L. crispatus* genomes retrieved from NCBI showed an average of 84 contigs per genome, ranging between one and 249 contigs, as well as an average number of CDS of 1957, ranging between 1585 and 2378.

### 3.2. Comparative Genomics Analysis of *L. crispatus* Strains

We performed a comparative genomics analysis through the use of the Pan-Genomes Analysis Pipeline (PGAP) software [16]. Clusters of Orthologous Groups (COGs) were defined as genes sharing >50% identity from alignments with >80% coverage. This analysis resulted in the prediction of the *L. crispatus* pan-genome, i.e., the entire set of genes belonging to this species and encoded by the strains analyzed, constituted by a total of 8387 COGs (Figures S1 and S2). Furthermore, the core-genome analysis allowed to identify 581 COGs shared by all genomes, thus representing the core-genome of the *L. crispatus* species. The 97 genomes included in this comparative genomics analysis were also screened for strain-specific unique genes. This analysis revealed that the pool of unique genes ranges from 11 to 137, with an average of 40 (Supplementary Excel File 1). In addition, we extrapolated the curves of both the pan-genome and the core-genome in order to evaluate their trend. The achieved pan-genome curve demonstrates that the large majority of the genetic diversity of *L. crispatus* is represented by the pool of sequenced genomes that we included in the analysis (Figure S2). Moreover, the *L. crispatus* core-genome appears to be stable at circa 750 COGs across the last five iterations.

Furthermore, we performed a phylogenomics analysis based on the alignment of the core gene sequences, which is an approach that has been shown to provide a detailed overview of the phylogenetic relationships between both distantly and closely related strains [24] (Figure 2). The reconstructed phylogenetic tree revealed clustering based on the strains' ecological source, with a clear distinction between the genomes of human and animal origin (Figure 2). Interestingly, separate clusters were also observed between strains isolated from the poultry species *Gallus gallus* and *Meleagris gallopavo*, thus revealing peculiar host-specific adaptation of *L. crispatus* strains. In addition, a minor cluster with a mixed composition of strains of human and poultry origin was also identified, thus representing strains with putative multi-host specificity.

The phylogenetic tree obtained from the alignment of the *L. crispatus* core-genome allowed an in-depth exploration of the genetic features characterizing the strains of human origin (Figure 2). The 31 *L. crispatus* strains constituting the "Human" cluster observed in the phylogenetic tree (Figure 2) were used to identify the human-specific core-genome (HSCG), i.e., the set of COGs present in all human isolates and absent in the other genomes included in this study (Supplementary Excel File 1). Notably, no COG shared by all *L. crispatus* strains of human origin was observed. In fact, the analysis resulted in the identification of just 27 COGs shared by >50% of the 31 strains, of which the most shared COG was identified in only 84% (26/31) of the analyzed genomes. Thus, highlighting the high genomic variability of these strains despite their common ecological source, i.e., the human vagina, as also indicated by a higher average number of strain-specific (44) respect to that of the other genomes of strains isolated from non-human hosts (37) (Supplementary Excel File 1). Furthermore, we have screened the core-genome of strains isolated from chicken and turkey gut as for the strains derived from humans, i.e., focusing on the genes present only among the strains of a specific ecological niche. From this screening, we observed that the genomes of *L. crispatus* strains isolated from these two ecological niches are characterized by COGs shared by all the strains isolated from the same host. These results confirm the conclusions seen above for the human isolates, thus corroborating the high genetic variability of the whole *L. crispatus* species.



**Figure 2.** Phylogenetic core-genes tree of the 97 *L. crispatus* strains included in the comparative analysis. The figure shows the phylogenetic tree generated by alignment of the core genome previously detected among the 97 *L. crispatus* strains analyzed. The isolation source of each strain is also reported. Red text is used to highlight strains sequenced in this study and isolated from human.

### 3.3. Comprehensive Genomic Characterization of *L. crispatus* Strains Isolated from the Human Vaginal Tract

*In silico* surveys exploiting a genomic approach, i.e., probiogenomics, is becoming crucial for the rapid screening of bacterial collections to identify optimal potential probiotic candidates, which will be further characterized by *in vitro* and *in vivo* experimental

validations [25]. Following this approach, we performed a genomic dissection of the eight strains sequenced in the framework of this study that were previously isolated from the human vaginal environment of healthy women, in order to provide an example of how comparative genomics data can drive selection of new next-generation probiotics (Table S1).

Therefore, we analyzed the genomic repertoire occurring only and exclusively among the eight strains which were isolated from human sequenced in this study and not present in any other of the additional 89 *L. crispatus* strains left from the whole comparative genomics analysis, i.e., 81 publicly available genomes and eight strain of non-human origin sequenced in the framework of this project. Between the eight selected *L. crispatus* strains, an average of 53 genes (16%) were dispensable, a total of 278 (84%) were unique of a specific strain, ranging from 21 to 40, and no genes were shared exclusively between these eight strains (Figure S1, Table S2). Dispensable genes are those genes that are not included in the core-genome and that are also not unique genes of specific strains.

Subsequently, we performed a detailed analysis of the unique genetic repertoire of the eight selected strains by homology search in the NCBI RefSeq database and protein domain prediction based on a range of databases including PFAM, PANTHERA and CDD [22,26,27] (Table S2).

Of the 278 unique genes submitted to functional annotation, 168 resulted to encode hypothetical proteins, while among the 110 genes with a predicted function only 69 showed presence of a known protein domain. This functional prediction has been validated through the presence of specific functional domains, ranging from zero to 19 for each strain (Table S2, Figure 2). Moreover, of the 168 unique genes predicted to encode hypothetical proteins, only nine encompass a functional domain, seven of which correspond to the Domain of Unknown Function DUF4044-type domain (Table S2). Intriguingly, *L. crispatus* LB57 genome displays the highest number of unique genes, i.e., 49, of which 19 showing conserved functional domains, followed by the genomes of LB58 and LB56 with 48 and 40 unique genes, respectively (Table S2, Figure 2). Furthermore, the genome of LB57 presents intriguingly unique genes, such as transporters for sugars (HYQ49\_1573 and HYQ49\_2460), a transporter putatively involved in exporting the thiol-containing redox-active molecules cysteine and glutathione (HYQ49\_1855 and HYQ49\_1856) [28], two flippases (HYQ49\_1948 and HYQ49\_2461) and a heavy metal resistance efflux pump (HYQ49\_2249) as indicated by substrate specificity prediction by homology search in the Transporter Classification DataBase (TCDB) [29] (Table S2). Moreover, LB57 showed the presence of genes encoding for putative surface proteins that may be involved in the interaction with the surrounding environment (HYQ49\_2454, HYQ49\_2521, HYQ49\_2583 and HYQ49\_2591), a gene participating in D-alanyl-lipoteichoic acid biosynthesis (HYQ49\_2170) and a range of peptidases (HYQ49\_0166, HYQ49\_2407, HYQ49\_2520, and HYQ49\_2586) (Table S2). LB56 and LB58 strains also showed interesting unique genetic features. The unique gene set of LB56 is characterized by the presence of a type II toxin-antitoxin system (HYQ48\_2254 and HYQ48\_2255), a PTS transporter for sugars (HYQ48\_1404 and HYQ48\_1405) and a nitroreductase family protein (HYQ48\_0882), while LB58 encompass a putative listeriolysin S family TOMM bacteriocin (HYQ50\_0043), a type II toxin-antitoxin system (HYQ50\_2190) and three putative cell wall-anchored proteins (HYQ50\_2017, HYQ50\_2201, and HYQ50\_2202) (Table S2). Nevertheless, further functional analyses, e.g., transcriptomics and proteomics experiments, will be needed to corroborate these *in silico* analyses.

### 3.4. Prediction of Bacteriocin Production by *L. crispatus* Strains

Bacteriocins are peculiar proteins, produced by some bacterial strains, able to selectively inhibit the growth of a very limited range of bacterial targets [30]. In nature, they are produced by specific bacterial strains providing an ecological advantage in their niches toward specific competing strains [30]. Therefore, it is of high probiotic interest to study the presence of genes related to the synthesis of bacteriocins, as they can be fundamental in

defining potential superior colonization capabilities of a specific strain or in limiting the occurrence of pathogens.

We performed an analysis of the genes related to bacteriocin production using BAGEL4 tool on the 16 strains we isolated and sequenced [31] (Table S3, Figure 3). Through this analysis, we could identify a total of six putative bacteriocins genes/loci, named LCB (Lactobacillus Crispatus Bacteriocins) 1–6, distributed among the eight *L. crispatus* isolates of human origin and additional two in the strains isolated from poultry (LCB 7–8). Intriguingly, LCB 3 is constituted by a locus of six genes that include a predicted two-component regulatory system, a putative dedicated ABC transporter, a small peptide pheromone along with an immunity protein. This locus is absent in the isolates from poultry, while it is present in all the eight *L. crispatus* strains isolated from humans. In-depth analysis of this bacteriocin family based on PFAM data (<https://www.ebi.ac.uk/interpro/entry/pfam/PF10439/>) revealed that it may be implicated in inhibiting the growth of streptococci and closely related taxa [32].

Moreover, the genome analysis of the *L. crispatus* strains isolated from healthy human vagina revealed that LB57 is the one with the highest number of genes for bacteriocins synthesis, i.e., 13 genes, encoding for LCBs 1, 2, 3, 4, 5, and 6 (Table S3, Figure 3). Thus, representing the *L. crispatus* strain with the broader genetic potential in bacteriocin production. Furthermore, the *L. crispatus* LB62 and LB63 strains also showed the presence of LCBs 1, 2, 3, 4, 5, and 6, thus highlighting their potential interest for industrial applications, followed by LB58, LB59, LB60, and LB61 showing a total of five LCBs. In contrast, LB56 showed the presence of just three LCBs (Table S3).

### 3.5. Glycobiome Prediction of *L. crispatus* and Experimental Validations

The glycobiome is defined as the genetic repertoire related to sugar metabolism responsible for the simple and complex carbohydrate utilization by bacteria [33,34]. Based on this genetic background, each bacterial strain may exhibit unique growth characteristics in environments with specific carbohydrates availability, providing them specific ecological advantages when competing with other strains. Therefore, the prediction of the glycobiome in putative probiotic bacteria is of fundamental importance to define the possible colonization and persistence capabilities of a strain in relation to the availability of natural carbon sources in the environment [35].

So, we analyzed the glycobiome profiles of the eight *L. crispatus* strains of human origin sequenced in this study in order to obtain a comprehensive overview of their ability to metabolize different sugars (Figure 4). Notably, the results showed similar glycobiome profiles among the analyzed strains, with the only exception of the predicted glycobiome of *L. crispatus* LB57. (Figure 4). In fact, this strain showed the higher abundance of genes constituting the glycobiome, i.e., 294, followed by LB58 and LB56 with 273 and 267 genes, respectively. Remarkably, LB57 also showed the absence of genes encoding for 11 glycosyl hydrolases (GH) and four glycosyl transferases (GT) that are instead present in all other strains and displayed a concomitant high abundance of genes classified as a range of 16 GT families absent in the other genomes (Figure 4). Furthermore, the genome of LB57 encodes 33 genes classified as GT3, absent in the other isolates, involved in glycogen synthesis (Figure 4). This may indicate that this strain has evolved to use glycogen for energy reserve, as previously observed in other bacteria, resulting in higher resistance to starvation and improved ecological fitness [12,36]. GT3 genes have been previously observed in *L. crispatus* strains isolated from vaginal microbial populations characterized by lactobacilli dominance [12]. Notably, their presence in *L. crispatus* genomes has also been linked to the synthesis of cell surface glycoconjugates that may be implicated in interactions with the host's epithelial cells [12].

a)

Isolation Source	Human	Human	Human	Human	Probiotic	Human	Human	Human	Human	Chicken						
Strain	LB56	LB57	LB58	LB59	LB60	LB61	LB62	LB63	PRL2021	LB64	LB65	LB66	LB67	LB68	LB69	LB70
LCB_1	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2
LCB_2	1	2	1	1	1	1	1	1	1	2	1	1	2	1	1	1
LCB_3	6	6	6	6	6	5	6	6								6
LCB_4		1		1				1	1							
LCB_5		1	1	1	1	1	1	1								
LCB_6		1	1		1	1	1	1	1	1	1	1	1	1	1	1
LCB_7																1
LCB_8																1
<b>Total</b>	<b>8</b>	<b>13</b>	<b>11</b>	<b>11</b>	<b>11</b>	<b>10</b>	<b>12</b>	<b>12</b>	<b>9</b>	<b>5</b>	<b>4</b>	<b>4</b>	<b>5</b>	<b>4</b>	<b>4</b>	<b>12</b>

b)

Legend		
ID Code	OD	Colour Code
-	$X < 0.16$	Orange
+	$0.161 < X < 0.5$	Light Blue
++	$0.51 < X < 0.8$	Light Green
+++	$0.81 < X < 1.9$	Dark Green

	Human		Human		Human		Human		Human		Human		Human		Human	
	LB55		LB56		LB57		LB58		LB59		LB61		LB62		LB63	
	24h	48h														
Arabinose	0.138	0.160	0.140	0.172	0.201	0.237	0.333	0.426	0.120	0.129	0.251	0.274	0.136	0.179	0.162	0.205
Cellobiose	0.161	1.174	0.465	0.824	0.657	0.701	0.855	1.271	0.127	0.140	0.994	1.091	0.796	1.030	0.299	1.306
Fructose	1.449	1.589	0.973	1.289	0.695	0.696	1.060	1.177	0.133	0.152	0.951	1.462	1.059	1.150	1.088	1.082
Fucose	0.145	0.168	0.154	0.250	0.245	0.303	0.281	0.318	0.119	0.132	0.311	0.330	0.179	0.209	0.141	0.183
Galactose	0.256	0.995	0.379	1.030	0.390	0.548	1.189	1.309	0.375	0.768	0.952	1.238	0.954	1.054	0.863	1.101
Glucose	1.304	1.433	1.139	1.448	0.621	0.713	1.271	1.546	0.908	0.923	0.913	1.011	1.055	1.138	1.013	1.161
Glycogen	0.5383	0.5651	0.4558	0.4824	0.4598	0.4612	0.4874	0.5818	0.4483	0.4638	0.4292	0.4456	0.4407	0.4551	0.4527	0.4819
Inulin	0.187	0.263	0.325	0.343	0.691	0.768	0.651	0.840	0.134	0.143	0.390	0.449	0.240	0.294	0.264	0.338
Lactose	0.183	0.188	0.300	0.304	0.251	0.323	1.012	1.154	0.129	0.138	0.129	0.138	1.177	1.339	0.791	0.848
Maltodextrin	1.063	1.170	0.917	0.937	0.643	0.771	1.071	1.216	0.486	0.708	1.041	1.139	1.089	1.273	0.757	1.037
Maltose	1.347	1.582	0.608	1.312	0.481	0.512	1.328	1.413	1.008	1.074	1.472	1.759	1.045	1.080	1.278	1.559
Mannitol	0.147	0.292	0.148	0.282	0.229	0.289	0.841	1.034	0.113	0.123	0.236	0.251	0.138	0.182	0.438	0.679
Mannose	0.118	0.123	0.140	0.156	0.845	1.149	1.032	1.251	0.155	0.173	0.836	1.072	0.772	1.043	1.182	1.252
Melibiose	0.154	0.168	0.204	1.197	0.326	0.379	0.551	1.344	0.124	0.138	0.332	0.359	0.184	0.229	0.149	0.200
N-acetyl-D-galactosamine	0.164	0.174	0.207	0.230	0.221	0.280	0.403	0.402	0.123	0.142	0.278	0.280	0.176	0.209	0.159	0.159
N-acetyl-D-glucosamine	0.877	1.142	0.653	1.166	0.227	0.224	1.151	1.204	0.237	0.336	0.171	1.035	0.533	0.920	0.949	1.112
Raffinose	0.149	0.172	0.338	1.017	0.305	0.365	0.853	0.991	0.130	0.135	0.317	0.326	0.170	0.229	0.158	0.175
Rhamnose	0.135	0.161	0.179	0.281	0.245	0.304	0.390	0.406	0.120	0.124	0.303	0.317	0.193	0.207	0.131	0.153
Ribose	0.143	0.156	0.143	0.183	0.244	0.274	0.327	0.343	0.117	0.113	0.239	0.243	0.156	0.204	0.139	0.175
Sorbitol	0.159	0.174	0.200	0.253	0.301	0.353	0.339	0.392	0.129	0.143	0.359	0.386	0.189	0.217	0.164	0.197
Starch	0.182	0.199	0.199	0.279	0.359	0.411	0.388	0.424	0.120	0.136	0.456	0.482	0.197	0.265	0.179	0.205
Sucrose	0.162	0.516	0.293	0.961	0.440	0.508	1.156	1.170	0.131	0.143	1.253	1.346	0.924	0.968	1.357	1.373
Trehalose	0.163	1.451	0.307	1.227	0.834	0.866	0.404	1.147	0.128	0.138	1.293	1.546	0.228	0.751	0.253	0.755
Turanose	0.165	0.178	0.219	0.240	0.389	0.945	0.852	1.094	0.135	0.145	0.483	0.795	0.223	0.308	0.256	0.339
Xylose	0.157	0.161	0.168	0.292	0.305	0.352	0.376	0.408	0.114	0.123	0.384	0.414	0.189	0.212	0.139	0.170

**Figure 3.** Analysis of the distribution of genes for bacteriocins and growths on specific sugars. Panel (a) shows a table illustrating the distribution of genes involved in the synthesis of bacteriocins among the 16 strains of *L. crispatus* sequenced in this study, with subdivision by isolation source. Additionally, a color coding was used to further highlight the distribution of these genes, with blue indicating high presence and black indicating complete absence of the genes. Panel (b) reported the growth performances of the eight strains of vaginal origin sequenced in this study when cultivated on media enriched with a range of different carbon sources.

	251	267	294	273	258	240	264	261
	LB55	LB56	LB57	LB58	LB59	LB61	LB62	LB63
AA1	1	1	1	1	1	1	1	1
AA10	0	0	0	1	0	0	0	0
CBM0	2	2	2	2	2	2	2	2
CBM13	2	2	2	2	2	2	2	2
CBM2	0	0	0	1	0	0	0	0
CBM23	1	1	1	1	1	1	1	1
CBM34	1	1	1	1	1	1	1	1
CBM37	0	1	1	1	1	0	1	1
CBM41	1	1	2	1	1	3	1	1
CBM48	5	6	0	6	6	8	6	6
CBM5	1	1	1	1	1	1	1	1
CBM50	17	20	0	21	18	19	20	19
CBM6	1	1	1	1	1	1	1	1
CE0	1	1	3	1	1	1	1	1
CE1	1	1	1	1	1	1	1	1
CE7	0	2	2	2	2	0	2	2
CE8	1	1	1	1	1	1	1	1
CE9	6	6	0	6	6	4	6	5
CE11	0	0	0	2	2	0	2	2
CE12	0	0	1	1	0	0	0	0
CE14	1	1	0	1	1	1	1	1
GH0	1	1	3	1	1	1	1	1
GH1	16	17	0	17	18	11	17	17
GH2	6	6	0	5	5	4	5	5
GH3	2	2	1	1	1	1	1	1
GH4	1	1	2	2	2	2	2	2
GH6	3	3	3	3	3	3	3	3
GH8	2	2	3	1	1	1	1	1
GH13	2	2	1	2	1	1	2	2
GH13_3	5	5	0	5	5	5	5	5
GH13_11	4	5	0	5	5	5	5	5
GH13_14	1	1	2	1	1	3	1	1
GH13_18	1	1	0	1	0	0	0	0
GH13_20	1	1	1	1	1	1	1	1
GH13_26	1	1	1	1	1	1	1	1
GH13_29	4	3	0	4	3	4	4	4
GH13_30	1	1	1	1	1	1	1	1
GH13_31	2	2	3	2	2	2	2	2
GH17	0	1	0	1	0	0	0	0
GH19	1	1	1	1	1	1	1	1
GH23	10	11	0	12	12	10	14	13
GH24	0	0	0	0	0	1	0	0
GH25	5	8	0	5	6	7	7	6
GH28	1	1	1	1	1	1	1	1
GH31	3	3	2	2	3	3	2	2
GH32	6	6	0	4	4	4	6	6
GH36	1	1	1	1	0	0	0	0
GH37	3	3	3	3	3	3	3	3
GH38	0	0	0	2	2	0	0	0
GH42	3	5	0	3	4	3	3	4
GH43_11	4	4	2	2	1	1	2	1
GH43_18	1	1	0	0	1	1	1	1
GH55	1	1	3	2	2	1	2	2
GH65	3	3	0	3	3	3	3	3
GH68	0	0	0	1	1	0	0	0
GH70	0	1	0	0	0	0	0	0
GH72	1	1	1	1	1	1	1	1
GH73	8	8	0	7	6	6	6	7
GH91	0	0	0	1	0	0	0	0
GH93	1	1	0	0	0	0	0	0
GH95	1	1	1	1	1	1	1	1
GH125	0	0	0	1	1	0	0	0
GH130	1	1	1	1	1	1	1	1
GH133	1	1	0	1	0	0	1	1

Figure 4. Cont.

GT0	1	1	1	1	1	1	1	1
GT1	4	4	0	4	4	4	3	4
GT2	30	30	38	33	32	29	33	33
GT3	0	0	33	0	0	0	0	0
GT4	32	32	25	34	31	31	32	33
GT5	1	1	21	1	0	0	1	1
GT6	0	0	19	0	0	0	0	0
GT7	0	0	11	0	0	0	0	0
GT8	7	7	7	8	8	8	8	8
GT9	0	0	7	0	0	0	0	0
GT10	0	0	7	0	0	0	0	0
GT11	0	0	5	0	0	0	0	0
GT12	0	0	5	0	0	0	0	0
GT13	2	2	5	2	2	2	2	2
GT14	0	0	5	0	0	0	0	0
GT15	0	0	5	0	0	0	0	0
GT16	0	0	5	0	0	0	0	0
GT17	0	0	4	0	0	0	0	0
GT18	0	0	4	0	0	0	0	0
GT19	0	0	4	0	0	0	0	0
GT20	0	0	4	0	0	0	0	0
GT21	0	0	4	0	0	0	0	0
GT22	0	0	4	0	0	0	0	0
GT26	1	1	2	2	2	2	2	2
GT27	1	1	1	1	1	1	1	1
GT28	4	4	0	4	4	4	4	4
GT30	1	1	1	1	1	1	1	1
GT32	2	2	0	2	2	2	2	2
GT47	1	1	1	1	1	1	1	1
GT51	5	5	0	7	6	5	6	6
GT61	1	1	1	1	1	1	1	1
GT66	1	1	1	1	1	1	1	0
GT68	1	1	2	1	2	1	2	1
GT83	1	1	1	1	1	1	1	1
GT96	1	1	1	1	1	1	1	1
GT101	3	3	2	2	2	2	2	2
GT104	1	1	1	1	1	1	1	1

**Figure 4.** Predicted glycobiome of the eight strains of *L. crispatus* isolated from human vaginal sequenced in this study. The heat map shows the number of genes inherent to that specific class of the glycobiome, i.e., glycosyl hydrolases (GH), glycosyl transferases (GT), carbohydrate esterases (CE), carbohydrate binding module (CBM), or auxiliary activities (AA). The data was colored incrementally, starting from red to address the absence of related genes, up to green in cases where the number of hits, and therefore of genes, is high, with the maximum reported value being 38.

To validate these in silico data, we performed fermentation profiling analyses of these *L. crispatus* strains cultivated on 23 different sugars as the only carbon source, aiming to cross these data with the predicted glycobiome results (Supplementary Excel File 1, Figure 3). Notably, the cultivation of these strains on 25 carbohydrates, including both simple and complex sugars, corroborated the in silico glycobiome data (Figure 3). Remarkably, among the *L. crispatus* strain isolated from the vaginal environment, LB57 showed a peculiar growth profile, thus reflecting its glycobiome profile (Figures 3 and 4). Nevertheless, growth on glycogen resulted in performances of LB57 similar to the other strains, suggesting that the high number of GT3 genes identified by glycobiome analysis may not be involved in glycogen degradation and may participate in other biological processes such as the synthesis of cell surface glycoconjugates [12].

#### 4. Conclusions

We performed a comprehensive comparative genomics analysis of 97 strains of *L. crispatus* encompassing 81 publicly available genomes and 16 strain sequenced in the framework of this study resulting in the reconstruction of the pan-genome of this species. Data collected were used to define the core-genome of *L. crispatus* and to perform an in-depth phylogenomics analysis. Thus, revealing that strains isolated from the human vagina are characterized by high genetic variability despite their close phylogenetic relationship.

Moreover, a probiogenomics approach was used to investigate the genetic potential of the eight *L. crispatus* strains isolated from humans and sequenced in this study in order to identify putative novel probiotics for the treatment of vaginal dysbiosis. Interestingly, unique genes putatively involved in improved colonization capabilities and (opportunistic) pathogens competition have been detected. In detail, the strain LB57 showed the broader repertoire of unique strain-specific genes as well as of genes for the synthesis of bacteriocins along with a peculiar carbohydrate metabolism that was validated by growth assay. Notably, this strain also showed a high abundance of glycosyl transferases of class 3, which may be involved in glycogen metabolism and in the synthesis of cell surface glycoconjugates that may be implicated in interactions with the host's epithelial cells. Furthermore, while LB56 and LB58 also represented optimal candidates for further investigations due to their peculiar unique features involved in energy harvesting and host-microbe interaction, LB58 showed a broader repertoire of genes involved in the synthesis of bacteriocins that may provide superior ecological fitness.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/2076-2607/9/1/73/s1>.

**Author Contributions:** Conceptualization, C.M. and F.F.; methodology, C.M., F.F., G.A.L., G.L., R.A. and A.V.; formal analysis, F.F., G.A., G.A.L., G.L. and L.M.; investigation, C.M., F.F., M.V., F.T.; data curation, F.F., G.A.L., G.A., and R.A.; writing—original draft preparation, F.F.; writing—review and editing, C.M.; visualization, F.F., L.M., G.L., R.A. and A.V.; supervision, C.M. and F.T. All authors have read and agreed to the published version of the manuscript.

**Funding:** G.A. is supported by Fondazione Cariparma, Parma, Italy.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The accession numbers of the genomic sequence of the 16 strains sequenced in this study are reported in Table S1.

**Acknowledgments:** We furthermore thank GenProbio s.r.l. for financial support of the Laboratory of Probiogenomics. This research benefited from the HPC (High Performance Computing) Facility of the University of Parma, Italy.

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

## References

1. Barrientos-Duran, A.; Fuentes-Lopez, A.; de Salazar, A.; Plaza-Diaz, J.; Garcia, F. Reviewing the Composition of Vaginal Microbiota: Inclusion of Nutrition and Probiotic Factors in the Maintenance of Eubiosis. *Nutrients* **2020**, *12*, 419. [CrossRef]
2. Happel, A.U.; Kullin, B.; Gamielien, H.; Wentzel, N.; Zauchenberger, C.Z.; Jaspan, H.B.; Dabee, S.; Barnabas, S.L.; Jaumdally, S.Z.; Dietrich, J.; et al. Exploring potential of vaginal *Lactobacillus* isolates from South African women for enhancing treatment for bacterial vaginosis. *PLoS Pathog.* **2020**, *16*, e1008559. [CrossRef]
3. Kaur, H.; Merchant, M.; Haque, M.M.; Mande, S.S. Crosstalk Between Female Gonadal Hormones and Vaginal Microbiota Across Various Phases of Women's Gynecological Lifecycle. *Front Microbiol.* **2020**, *11*, 551. [CrossRef] [PubMed]
4. Chee, W.J.Y.; Chew, S.Y.; Than, L.T.L. Vaginal microbiota and the potential of *Lactobacillus* derivatives in maintaining vaginal health. *Microb. Cell Factories* **2020**, *19*, 203. [CrossRef]
5. Hassan, M.U.; Nayab, H.; Rehman, T.U.; Williamson, M.P.; Haq, K.U.; Shafi, N.; Shafique, F. Characterisation of Bacteriocins Produced by *Lactobacillus* spp. Isolated from the Traditional Pakistani Yoghurt and Their Antimicrobial Activity against Common Foodborne Pathogens. *BioMed Res. Int.* **2020**, *2020*, 8281623. [CrossRef]
6. Ravel, J.; Gajer, P.; Abdo, Z.; Schneider, G.M.; Koenig, S.S.; McCulle, S.L.; Karlebach, S.; Gorle, R.; Russell, J.; Tacket, C.O.; et al. Vaginal microbiome of reproductive-age women. *Proc. Natl. Acad. Sci. USA* **2011**, *108* (Suppl. S1), 4680–4687. [CrossRef]
7. Pan, M.; 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.* **2019**, *10*, 3146. [CrossRef]
8. Phukan, N.; Brooks, A.E.S.; Simoes-Barbosa, A. A Cell Surface Aggregation-Promoting Factor from *Lactobacillus gasseri* Contributes to Inhibition of *Trichomonas vaginalis* Adhesion to Human Vaginal Ectocervical Cells. *Infect. Immun.* **2018**, *86*. [CrossRef]

9. Takada, K.; Komine-Aizawa, S.; Kuramochi, T.; Ito, S.; Trinh, Q.D.; Pham, N.T.K.; Sasano, M.; Hayakawa, S. Lactobacillus crispatus accelerates re-epithelialization in vaginal epithelial cell line MS74. *Am. J. Reprod. Immunol.* **2018**, *80*, e13027. [[CrossRef](#)]
10. 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*, 1070. [[CrossRef](#)]
11. Zhang, Q.; Zhang, L.; Ross, P.; Zhao, J.; Zhang, H.; Chen, W. Comparative Genomics of Lactobacillus crispatus from the Gut and Vagina Reveals Genetic Diversity and Lifestyle Adaptation. *Genes* **2020**, *11*, 360. [[CrossRef](#)]
12. van der Veer, C.; Hertzberger, R.Y.; Bruisten, S.M.; Tytgat, H.L.P.; Swanenburg, J.; de Kat Angelino-Bart, A.; 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*, 49. [[CrossRef](#)] [[PubMed](#)]
13. Lugli, G.A.; Milani, C.; Mancabelli, L.; van Sinderen, D.; Ventura, M. MEGAnnotator: A user-friendly pipeline for microbial genomes assembly and annotation. *FEMS Microbiol. Lett.* **2016**, *363*. [[CrossRef](#)]
14. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)]
15. 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](#)]
16. Zhao, Y.; Wu, J.; Yang, J.; Sun, S.; Xiao, J.; Yu, J. PGAP: Pan-genomes analysis pipeline. *Bioinformatics* **2012**, *28*, 416–418. [[CrossRef](#)] [[PubMed](#)]
17. Lugli, G.A.; Milani, C.; Duranti, S.; Mancabelli, L.; Mangifesta, M.; Turrone, F.; Viappiani, A.; van Sinderen, D.; Ventura, M. Tracking the Taxonomy of the Genus Bifidobacterium Based on a Phylogenomic Approach. *Appl. Environ. Microbiol.* **2018**, *84*. [[CrossRef](#)]
18. Rozewicki, J.; Li, S.; Amada, K.M.; Standley, D.M.; Katoh, K. MAFFT-DASH: Integrated protein sequence and structural alignment. *Nucleic Acids Res.* **2019**, *47*, W5–W10. [[CrossRef](#)] [[PubMed](#)]
19. Larkin, M.A.; Blackshields, G.; Brown, N.P.; Chenna, R.; McGettigan, P.A.; McWilliam, H.; Valentin, F.; Wallace, I.M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **2007**, *23*, 2947–2948. [[CrossRef](#)]
20. Jones, P.; Binns, D.; Chang, H.Y.; Fraser, M.; Li, W.; McAnulla, C.; McWilliam, H.; Maslen, J.; Mitchell, A.; Nuka, G.; et al. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* **2014**, *30*, 1236–1240. [[CrossRef](#)]
21. Zhao, Y.; Tang, H.; Ye, Y. RAPSearch2: A fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics* **2012**, *28*, 125–126. [[CrossRef](#)] [[PubMed](#)]
22. Lu, S.; Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Marchler, G.H.; Song, J.S.; et al. CDD/SPARCLE: The conserved domain database in 2020. *Nucleic Acids Res.* **2020**, *48*, D265–D268. [[CrossRef](#)] [[PubMed](#)]
23. Consortium, C.A. Ten years of CAZypedia: A living encyclopedia of carbohydrate-active enzymes. *Glycobiology* **2018**, *28*, 3–8. [[CrossRef](#)] [[PubMed](#)]
24. Duranti, S.; Milani, C.; Lugli, G.A.; Mancabelli, L.; Turrone, F.; Ferrario, C.; Mangifesta, M.; Viappiani, A.; Sanchez, B.; Margolles, A.; et al. Evaluation of genetic diversity among strains of the human gut commensal Bifidobacterium adolescentis. *Sci. Rep.* **2016**, *6*, 23971. [[CrossRef](#)]
25. Langella, P.; Guarner, F.; Martin, R. Editorial: Next-Generation Probiotics: From Commensal Bacteria to Novel Drugs and Food Supplements. *Front. Microbiol.* **2019**, *10*, 1973. [[CrossRef](#)]
26. El-Gebali, S.; Mistry, J.; Bateman, A.; Eddy, S.R.; Luciani, A.; Potter, S.C.; Qureshi, M.; Richardson, L.J.; Salazar, G.A.; Smart, A.; et al. The Pfam protein families database in 2019. *Nucleic Acids Res.* **2019**, *47*, D427–D432. [[CrossRef](#)]
27. Thomas, P.D.; Kejariwal, A.; Campbell, M.J.; Mi, H.; Diemer, K.; Guo, N.; Ladunga, I.; Ulitsky-Lazareva, B.; Muruganujan, A.; Rabkin, S.; et al. PANTHER: A browsable database of gene products organized by biological function, using curated protein family and subfamily classification. *Nucleic Acids Res.* **2003**, *31*, 334–341. [[CrossRef](#)]
28. Holyoake, L.V.; Poole, R.K.; Shepherd, M. The CydDC Family of Transporters and Their Roles in Oxidase Assembly and Homeostasis. *Adv. Microb. Physiol.* **2015**, *66*, 1–53. [[CrossRef](#)]
29. Saier, M.H., Jr.; Reddy, V.S.; Tsu, B.V.; Ahmed, M.S.; Li, C.; Moreno-Hagelsieb, G. The Transporter Classification Database (TCDB): Recent advances. *Nucleic Acids Res.* **2016**, *44*, D372–D379. [[CrossRef](#)]
30. Simons, A.; Alhanout, K.; Duval, R.E. Bacteriocins, Antimicrobial Peptides from Bacterial Origin: Overview of Their Biology and Their Impact against Multidrug-Resistant Bacteria. *Microorganisms* **2020**, *8*, 639. [[CrossRef](#)]
31. de Jong, A.; van Hijum, S.A.; Bijlsma, J.J.; Kok, J.; Kuipers, O.P. BAGEL: A web-based bacteriocin genome mining tool. *Nucleic Acids Res.* **2006**, *34*, W273–W279. [[CrossRef](#)] [[PubMed](#)]
32. Dawid, S.; Roche, A.M.; Weiser, J.N. The blp bacteriocins of Streptococcus pneumoniae mediate intraspecies competition both in vitro and in vivo. *Infect. Immun.* **2007**, *75*, 443–451. [[CrossRef](#)] [[PubMed](#)]
33. Milani, C.; Turrone, F.; Duranti, S.; Lugli, G.A.; Mancabelli, L.; Ferrario, C.; van Sinderen, D.; Ventura, M. Genomics of the Genus Bifidobacterium Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment. *Appl. Environ. Microbiol.* **2016**, *82*, 980–991. [[CrossRef](#)] [[PubMed](#)]

34. Milani, C.; Lugli, G.A.; Duranti, S.; Turrone, F.; Mancabelli, L.; Ferrario, C.; Mangifesta, M.; Hevia, A.; Viappiani, A.; Scholz, M.; et al. Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. *Sci. Rep.* **2015**, *5*, 15782. [[CrossRef](#)] [[PubMed](#)]
35. Ventura, M.; O'Flaherty, S.; Claesson, M.J.; Turrone, F.; Klaenhammer, T.R.; van Sinderen, D.; O'Toole, P.W. Genome-scale analyses of health-promoting bacteria: Probiogenomics. *Nat. Rev. Microbiol.* **2009**, *7*, 61–71. [[CrossRef](#)]
36. Sekar, K.; Linker, S.M.; Nguyen, J.; Grunhagen, A.; Stocker, R.; Sauer, U. Bacterial Glycogen Provides Short-Term Benefits in Changing Environments. *Appl. Environ. Microbiol.* **2020**, *86*. [[CrossRef](#)]