UNIVERSITÀ DI PARMA

UNIVERSITY OF PARMA

Ph.D. in Biotechnology and Biosciences

XXXIII cycle (2017-2020)

Genomic and Ecological Studies of key Human Microbiota Members

Ph. D. Coordinator: Prof. Marco Ventura

Ph. D. Tutor: Prof. Francesca Turroni

Ph. D. Student: Walter Mancino

Table of contents

SUMMARY

The human microbiota is defined as the set of microorganisms residing on or within human tissues and biofluids. In last decades, the scientific community showed a huge interest in the study of the human microbiota, that represents one of the most complex bacterial community known and that have an important implication on human health. In this context, the genus *Bifidobacterium* represents one of the most important players of the human gut microbiota, especially in the early stages of life. In this Ph. D. thesis, different ecological and genomic aspects of this genus have been investigated. The first part of this thesis investigates genetic and ecological speciation between two important *Bifidobacterium* (sub)species, i.e. *Bifidobacterium animalis* subsp. *lactis* and *Bifidobacterium animalis* subsp. *animalis* taxa. Moreover, the *Bifidobacterium* inheritance from mothers to newborns is examined, through an *in vivo* study to evaluate the intriguing possible vertical transmission before delivery. Since different members of the *Bifidobacterium* genus are considered health-promoting bacteria, they should be totally safe to avoid the possible transmission of antibiotic resistance (AR) genes. Then, the resistome of bifidobacteria is explored in order to investigate the possible insurgence of AR genes in mobile genetic elements or close to them. Furthermore, a complete screening for the bifidobacterial Amoxicillin-Clavulanic acid (AMC) specific resistance is evaluated with the aim to understand the impact of this antibiotic in the human gut microbiota and especially versus this genus. Finally, the last part of this thesis concerns a detailed study regarding the vaginal microbiota, and more in specific the role of the *Lactobacillus crispatus* species in helping to preserve the healthy status of the host.

Chapter 1

General introduction

1.1 The human Microbiota.

The term “microbiota” means to the complex community of microorganisms that colonize a specific ecological niche (1, 2), whereas the word “microbiome” describes the collective genetic content of a determined microbiota (2-5). In this regard, the human microbiota is defined as the set of microorganisms residing on or within human tissues and biofluids (2). In last decades, the scientific community showed a huge interest for the study of the human microbiota, that represents one of the most complex bacterial community known and that have an important implication on human health (1). The long-term co-evolution between bacterial communities and the host contributed to the formation of different trophic interactions. In particular, the mutualistic relationship with the host ensures feeds and a niche for the growth of microorganisms, and at the same time, the microorganisms community is crucial for different physiological, metabolic and defensive functions.

In this context, several bacteria are able to colonize different human districts, such as the Gastro-Intestinal Tract (GIT), the vaginal environment, the skin tissue and the lung cavities (3, 6-10). Specifically, the environmental characteristics present in each human district play a selection for the colonization of different bacterial species. In fact, starting from the proximal GIT niches, the oral microbiota encompasses more than 700 different commonly occurring phylotypes (11). The most bacterial genus present in the human oral cavity is the *Streptococcus* genus, in particular at species level, *Streptococcus salivarius, Streptococcus mitis* and *Streptococcus saguinis* resulted to be the more abundant species in this human district (12). In addition, the nasal microbiota is strictly correlated to the oral microbiota, even if its composition is influenced by the humidity presents in this human district. In particular, the most prevalent genera are *Streptococcus*, *Prevotella*, *Veillonella* and *Fusobacterium* (9). In recent years, many studies focused their attention on the lower respiratory tract microbiota, which has been considered sterile until few years ago (13-15). In particular, it was demonstrated that the lung cavities have been colonized from bacteria that belonged to Bacteroidetes, Firmicutes, Proteobacteria and Actinobacteria phyla (13-15). In addition, there is increasing evidence that indicated the role of pulmonary microbiota in lung homeostasis and disease (15). Moreover,

several studies revealed that the human intestine is colonized principally by anaerobic bacteria with the prevalence of members of Bacteroidetes, Firmicutes, Proteobacteria and Actinobacteria phyla (16), furthermore, this composition change in based on the healthy status, the diet and the age of the host. Another very important niche is represented by human skin. This is a very huge space that covers our body with an important number of bacteria, which resulted to be principally aerobic bacteria (7). The latter belong to Firmicutes and Actinobacteria phyla, including members of *Staphylococcus*, *Corynebacterium* and *Propionibacterium* genera (7). More in specific, *Propionibacterium* spp. dominate the sebaceous sites, whereas *Corynebacterium* spp. and *Staphylococcus* spp. are preferentially abundant in moist areas, such as the bends of the elbows and the feet (7). Furthermore, the female vagina is an important hub of bacterial colonization. Regarding the composition of the human vagina microbiota, it is important underlined that bacteria that could resist to acidic environment (17), as well as members of Firmicutes phylum, such as *Lactobacillus* genus (18), dominate it. Finally, the investigation of the human niches, such as vagina and intestine, revealed the impact that their microbiota has on the health of the host. In fact, both vaginal and gut microbiota were demonstrated to be crucial for the healthy status of humans, including protection from viral, fungal and bacterial infections (18-21).

1.2 The human vaginal microbiota.

The human vaginal microbiota represents the bacterial community inhabiting the female vaginal tract. In last years, the study of this microbiota has become increasingly important in order to evaluate the central role of microorganisms for the correct balance of them to maintain a healthy status. Some of these studies have been demonstrated a variable composition of vaginal microbiota during the stages of life, in fact, it is strictly correlated to health and woman hormones (22-25). In particular, after birth, a high number of bacteria colonize the vaginal tract and it is characterized by an extremely low abundance of *Lactobacillus* species (26, 27). Subsequently, following the sexual maturation, lactobacilli become the dominant bacteria of the vaginal tract, reaching up to 90 % (18, 21, 26-30).

Recently, Ravel *et al.*, proposed a classification of vaginal microbiota based on their composition in five Community State Type (CST) (18). Specifically, four CST, known as CTS I, II, III and V, are characterized by the predominant presence of a *Lactobacillus* species. In particular, CST I is dominated by members of *Lactobacillus crispatus* species, CST II is dominated by bacteria belonging to *Lactobacillus gasseri* species, in CST III members of *Lactobacillus iners* species dominate the microbiota and, finally, CST V is dominated by microorganisms belong to *Lactobacillus jensenii* species (18). In this context, it is important to underline that CSTs I, II and III are the most frequent and typical of healthy women in reproductive age (18), whereas the CST V is frequently present in menopausal women and in subjects with dysfunctionality or pathologies (6). Lastly, CST IV is not dominated by a single bacterial species but is composed of different bacteria, such as *Gardnerella vaginalis* species and members of *Lachnospiraceae*, *Leptotrichiaceae* and *Prevotellaceae* families (18, 31, 32). Moreover, the CST IV is characterized by low *Lactobacillus*-level and it is associated with bacterial vaginosis condition and with bacterial, fungal or viral infections (18-21). Furthermore, this study highlights a correlation between the human vaginal microbiota composition and the ethnicity of women (18). In this context, the lactobacilli-dominant vaginal bacterial community is frequently present in vaginal microbiota of Asian and Caucasian women, meanwhile is less frequent in the vaginal tract of Hispanic and African individuals, which are characterized by CST IV in their urogenital tract (18).

Furthermore, the *L. crispatus* species is strictly correlated with a healthy status of the vaginal tract, protecting this human district from pathogens attachment and from bacterial infections (19-21). In fact, this species is able to promote the vaginal homeostatic (eubiotic) environment and immune barrier functions (33, 34). For these reasons, recently this species was considered an interesting new probiotic candidate for the prevention of urogenital tract infections and to restore a healthy vaginal environment (35-37).

1.3 The human intestinal microbiota.

The GIT represents the human organ system that digest food, extract and absorbs energy and nutrients, and expels the remaining waste as feces. The GIT is composed of all structures between the mouth and the anus, constituting a continuous way that concerns digestion organs, such as the stomach, small intestine and large intestine. It is important underlined that the GIT is colonized from trillions of bacteria and resulting to be the major human district colonized from microorganisms (38-40). All these microorganisms, which live into the human gut in a complex balanced community, have an important role in the healthy status of the host (41, 42). A serious variation of the balanced community could cause several metabolism dysfunction and different pathologies, as well as Irritable Bowel Syndrome (IBS) (43), Chron's disease, ulcerative colitis (44) and in the worst cases, could have a contribution to the development of colon cancer (45). Interestingly, the GIT microbiota composition changes along this tract, becoming gradually more complex and encompassing a major number of bacteria up to the most distal area of the gut, that seems colonized from more than 1000 different bacterial species (41). In this context, the gut microbiota plays a significant role in the regulation of immune responses and host defense against invading pathogens, as well as it is able to contribute to intestinal cell proliferation and differentiation, regulate the intestinal pH, have an effect on the absorption of certain nutrients, which are indigestible from human cells (42). Moreover, the intestinal microbiota composition is also highly variable in the life course and it is influenced by different factors, i.e., host factors (e.g., pH, digestive enzymes, mucus and transit time), environmental factors (e.g., diet-compounds and antibiotic treatments) and bacterial factors (e.g., metabolic capabilities, bacterial enzymes and adhesion strategies). Recently, several studies have been highlighted that the human GIT microbiota is specific to each individual and it is influenced by the physiological state, genetic background and by the host age (41, 42, 46-49). The different bacterial composition has a very important impact on the host health, influencing the immunity system (50), metabolisms (51) and the hormonal system (43). For these reasons, the formation of a stable GIT microbiota is a complex process starting immediately after the delivery (52). In fact, the human

microbiota is considered sterile before birth and it is colonized from several bacteria next to the delivery (52). After weaning and with the introduction of solid food in the diet, the microbiota community changes and becomes more complex and at three years of age the gut microbiota reaches a composition similar to adult and remains more or less stable over time in adulthood (53, 54). Each individual possesses specific microbial profiling of several bacterial species and it is different from the profiling of other people (55). Nowadays, different studies have been demonstrated that four phyla constituted about 97 % of the adult microbial intestinal population, i.e. Bacteroidetes (20-25 %), Firmicutes (60-65 %), Proteobacteria (5-10 %) and Actinobacteria (1-3 %) (16). Interestingly, in the infant intestine the Actinobacteria phylum, and in particular the *Bifidobacterium* genus, is numerically predominant (56).

1.4 The Actinobacteria phylum.

The Actinobacteria phylum represents one of the largest taxonomic unit among the Bacteria domain (57, 58). This phylum includes Gram-positive bacteria with a high content of guanine and cytosine (G+C) in the genome. They present several morphologies, as well as, coccoid morphology (*Micrococcus*), rod-coccoid form (*Arthrobacter*), fragmenting hyphal form (*Nocardia*) or permanent and highly differentiated branched mycelium (*Streptomyces*) (Fig. 1).

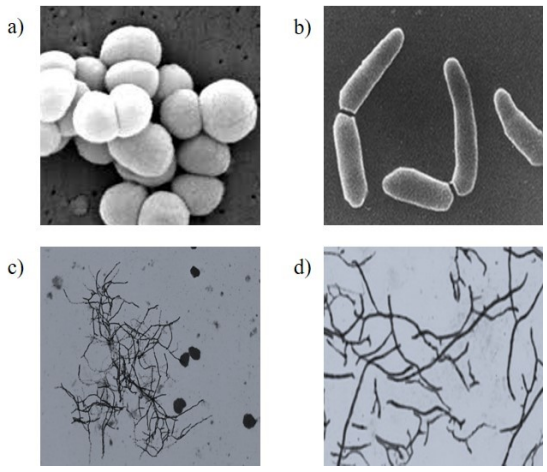


Figure 1. Morphological diversity of Actinobacteria captured with optical microscopy.
a) Colonies of *Micrococcus* spp.; b) *Arthrobacter* spp.; c) *Nocardia* spp.; d) *Streptomyces* spp.

The majority of Actinobacteria are widely distributed in aquatic, including marine habitats, and terrestrial ecosystems (59) and they exhibit different metabolic and physiological proprieties, such as the formation of different secondary metabolites and the production of extracellular enzymes (58). Interestingly, antibiotic molecules represent much of the secondary metabolites and the production of different antibiotics (e.g., neomycin, chloramphenicol and streptomycin) is a species-specific feature, which is largely exploited by pharmaceutical industries (60). For example, actinomycetes produce two-third known antibiotics (57), being very important microorganisms for the emerging multidrug-resistant pathogens (61-63). Furthermore, the Actinobacteria phylum includes very heterogeneous bacteria that have different impacts on humans. In fact, different members of *Mycobacterium*, *Nocardia*, *Tropheryma*, *Corynebacteria* and *Propionibacterium* genera are considered pathogens, whereas other bacteria, such as members of *Leifsonia* genus, are plant commensals and other species, as well as members of *Franchia* genus, are nitrogen-fixing symbionts. Moreover, few species are considered health-promoting bacteria, among these different members of *Bifidobacterium* genus are very important human commensals (58).

1.5 The *Bifidobacterium* genus.

The *Bifidobacterium* genus belongs to the family of *Bifidobacteriaceae*, encompassing non-motile, non-sporulating, non-gas producing, anaerobic or microaerophilic and saccharolytic bacteria. They were isolated for the first time by Henry Tissier in 1899 from infant stool samples and they were classified as a new species of Lactic Acid Bacteria (LAB) called *Bacillus bifidus* (Tissier, 1900). About 50 years later, they were taxonomically separated from the *Lactobacillus* genus in a new genus belonging to the *Bifidobacteriaceae* family. They present a typical bifid form or a rod-shape morphology. Nowadays, the *Bifidobacterium* genus encompasses 84 (sub)species, involving 72 species and 12 subspecies (Table 1).

Table 1. *Bifidobacterium* (sub)species recognized as reference strains (type strains).

***Bifidobacterium* strains**	**Isolation**	**References**
Bifidobacterium actinocoloniiforme DSM 22766	Bumblebee digestive tract	Killer 2011
Bifidobacterium adolescentis ATCC 15703	Intestine of human adult	Reuter 1963
Bifidobacterium aemilianum XV10	Carpenter bee digestive tract	Alberoni 2019
Bifidobacterium aerophilum DSM 100689	Feces of cotton-top tamarin	Michelini 2016
Bifidobacterium aesuclapii DSM 26737	Feces of baby common marmoset	Toh 2015
Bifidobacterium angulatum LMG 11039	Feces of human	Scardovi 1974
Bifidobacterium animalis subsp. *animalis* LMG 10508	Feces of rat	Mitsouka 1969
Bifidobacterium animalis subsp. *lactis* DSM 10140	Fermented milk	Masco 2004
Bifidobacterium anseris LMG 30189	Feces of domestic goose	Lugli 2018
Bifidobacterium apri DSM 100238	Digestive tract of wild pig	Pechar 2017
Bifidobacterium aquikefiri LMG 28769	Water Kefir	Laureys 2016
Bifidobacterium asteroides LMG 10735	Hindgut of honeybee	Scardovi 1969
Bifidobacterium avesanii DSM 100685	Feces of cotton-top tamarin	Michelini 2016
Bifidobacterium biavatii DSM 23969	Feces of tamarin	Endo 2012
Bifidobacterium bifidum LMG 11041	Feces of Brest-fed infant	Tissier 1900
Bifidobacterium bhoemicum DSM22767	Bumblebee digestive tract	Killer 2011
Bifidobacterium bombi DSM 19703	Bumblebee digestive tract	Killer 2009
Bifidobacterium boum LMG 10736	Rumen of bovine	Scardovi 1979
Bifidobacterium breve LMG 13208	Infant stool	Reuter 1963
Bifidobacterium callimiconis LMG 30938	Feces of Goeldi's marmoset	Duranti 2019
Bifidobacterium callitrichidarum DSM 103152	Feces of emperor tamarin	Modesto 2018
Bifidobacterium callitrichos DSM 23973	Feces of common marmoset	Endo 2012
Bifidobacterium castoris LMG 30937	Feces of beaver	Duranti 2019
Bifidobacterium catenulatum LMG 11043	Adult intestine	Scardovi 1974
Bifidobacterium catenulatum subsp. *kashiwanohense* DSM21854	Infant feces	Morita 2011
Bifidobacterium catulorum DSM103154	Feces of common marmoset	Modesto 2018
Bifidobacterium cebidarum LMG31469	Feces of golden-headed feces tamarin	Duranti 2020
Bifidobacterium choerinum LMG 10510	Feces of piglet	Scardovi 1979
Bifidobacterium commune LMG28292	Bumble bee gut	Praet 2015
Bifidobacterium coryneforme LMG 18911	Hindgut of honeybee	Scardovi 1969
Bifidobacterium criceti LMG 30188	Feces of European hamster	Lugli 2018
Bifidobacterium crudilactis LMG 23609	Raw cow milk	Delcenserie 2007
Bifidobacterium cuniculi LMG 10738	Feces of rabbit	Scardovi 1979
Bifidobacterium dentium LMG 11045	Oral cavity	Scardovi 1974
Bifidobacterium colichotidis LMG 30941	Feces of Patagonian mara	Duranti 2019
Bifidobacterium eulemuris DSM 100216	Feces of black lemur	Michelini 2016
Bifidobacterium felsineum DSM103139	Feces of cotton-top tamarin	Modesto 2018
Bifidobacterium gallicum LMG 11596	Adult intestine	Lauer 1900
Bifidobacterium pullorum subsp. *gallinarum* LMG 11586	Caecum of chicken	Watabe 1983
Bifidobacterium goeldii LMG 30939	Feces of Goeldi's marmoset	Duranti 2019
Bifidobacterium hapali DSM 100202	Feces of baby common marmoset	Michelini 2016
Bifidobacterium imperatoris LMG 30297	Feces of emperor tamarin	Lugli 2018

Bifidobacterium indicum LMG 11587	Insect	Scardovi 1969
Bifidobacterium italicum LMG 30187	Feces of European rabbit	Lugli 2018
Bifidobacterium jacchi DSM 103362	Feces of baby common marmoset	Modesto 2019
Bifidobacterium lemurum DSM 28807	Feces of ting-tailed lemur	Modesto 2015
Bifidobacterium leontopitechi LMG 31471	Feces of Goeldi's monkey	Duranti 2020
Bifidobacterium longum subsp. *infantis* ATCC 15697	Intestine of infant	Reuter 1963
Bifidobacterium longum subsp. *longum* LMG 13197	Adult intestine	Reuter 1963
Bifidobacterium longum subsp. suis LMG 21814	Feces of pig	Matteuzzi 1971
Bifidobacterium magnum LMG 11591	Feces of rabbit	Scardovi 1974
Bifidobacterium margollesii LMG 30296	Feces of pygmy marmoset	Lugli 2018
Bifidobacterium meryciucm LMG 11341	Rumen of bovine	Biavati 1991
Bifidobacterium minimum LMG 11592	Sewage	Scardovi 1974
Bifidobacterium mongoliense DSM 21395	Fermented mare's milk	Watanabe 2009
Bifidobacterium moukabalense DSM 27321	Feces of gorilla	Tsuchida 2014
Bifidobacterium myosotis DSM 100196	Feces of common marmoset	Michelini 2016
Bifidobacterium parmae LMG 30295	Feces of pygmy marmoset	Lugli 2018
Bifidobacterium primatium DSM 100687	Feces of cotton-top tamarin	Modesto 2018
Bifidobacterium pseudocatenulatum LMG 10505	Infant feces	Scardovi 1979
Bifidobacterium pseudolongum subsp. *globosum* LMG 11596	Rumen of bovine	Scardovi 1969
Bifidobacterium pseudolongum subsp. *pseudolongum* LMG 11571	Feces of swine	Mitsouka 1969
Bifidobacterium psychraerophilum LMG 21775	Caecum of pig	Simpson 2004
Bifidobacterium pullorum subsp. *pullorum* LMG 21816	Feces of chicken	Trovatelli 1974
Bifidobacterium ramosum DSM 100688	Feces of cotton-top tamarin	Michelini 2016
Bifidobacterium reuteri DSM 23975	Feces of common marmoset	Endo 2012
Bifidobacterium rousetti BCRC 81136	Feces of Egyptian fruit bat	Modesto 2019
Bifidobacterium ruminantium LMG 21811	Rumen of bovine	Biavati 1961
Bifidobacterium pullorum subsp. *saeculare* LMG 14934	Feces of rabbit	Biavati 1961
Bifidobacterium saguini LMG 23967	Feces of tamarin	Endo 2012
Bifidobacterium saimirii LMG 30940	Feces of Bolivian saimiri	Duranti 2019
Bifidobacterium scaligerum DSM 103140	Feces of cotton-top tamarin	Modesto 2018
Bifidobacterium scardovii LMG 21589	Blood	Hoyles 2002
Bifidobacterium simiarum DSM 103153	Feces of emperor tamarin	Modesto 2018
Bifidobacterium stellenboschense DSM 23968	Feces of tamarin	Endo 2012
Bifidobacterium subtile LMG 11597	Sewage	Scardovi 1974
Bifidobacterium porcinum LMG 21689	Feces of piglet	Zhu 2003
Bifidobacterium thermacidophilum LMG 21395	Anaerobic digester	Dong 2000
Bifidobacterium termophilum JCM 7027	Rumen of bovine	Mitsuoka 1969
Bifidobacterium tibiigranuli LMG 31086	Water Kefir	Eckel 2019
Bifidobacterium tissieri DSM 100201	Feces of baby common marmoset	Michelini 2016
Bifidobacterium tsurumiense JCM 13495	Hamster dental plaque	Okamoto 2008
Bifidobacterium vansinderenii LMG 30126	Feces of emperor tamarin	Duranti 2017
Bifidobacterium vespertilionis DSM 106025	Feces of Egyptian fruit bat	Modesto 2019
Bifidobacterium xylocopae DSM104955	Carpenter bee digestive tract	Alberoni 2019

In last years, the phylogeny of the *Bifidobacterium* genus has been explored using different methods based on the sequencing of the 16S rRNA gene or through multi-locus approach, as well as the sequencing of several housekeeping genes, i.e. *clpC*, *dnaJ*, *rpoC*, *xpf*, *dnaB* and *purF* (64, 65). Recently, a comparative genomics analysis based on 72 sequenced bifidobacterial type strains revealed the presence of 261 *Bifidobacterium*-specific clusters of orthologous genes (COGs) shared by these genomes, called bifidobacterial core-genome (66). In the same study, a phylogenetic tree constructed with the amino acid concatenation of the core- genome revealed the existence of 10 different phylogenetic groups, encompassing *Bifidobacterium adolescentis*, *Bifidobacterium boum*, *Bifidobacterium pullorum*, *Bifidobacterium asteroides*, *Bifidobacterium longum*, *Bifidobacterium psychraerophilum*, *Bifidobacterium bifidum*, *Bifidobacterium pseudolongum*, *Bifidobacterium bombi* and *Bifidobacterium tissieri* groups (66) (Fig. 2).

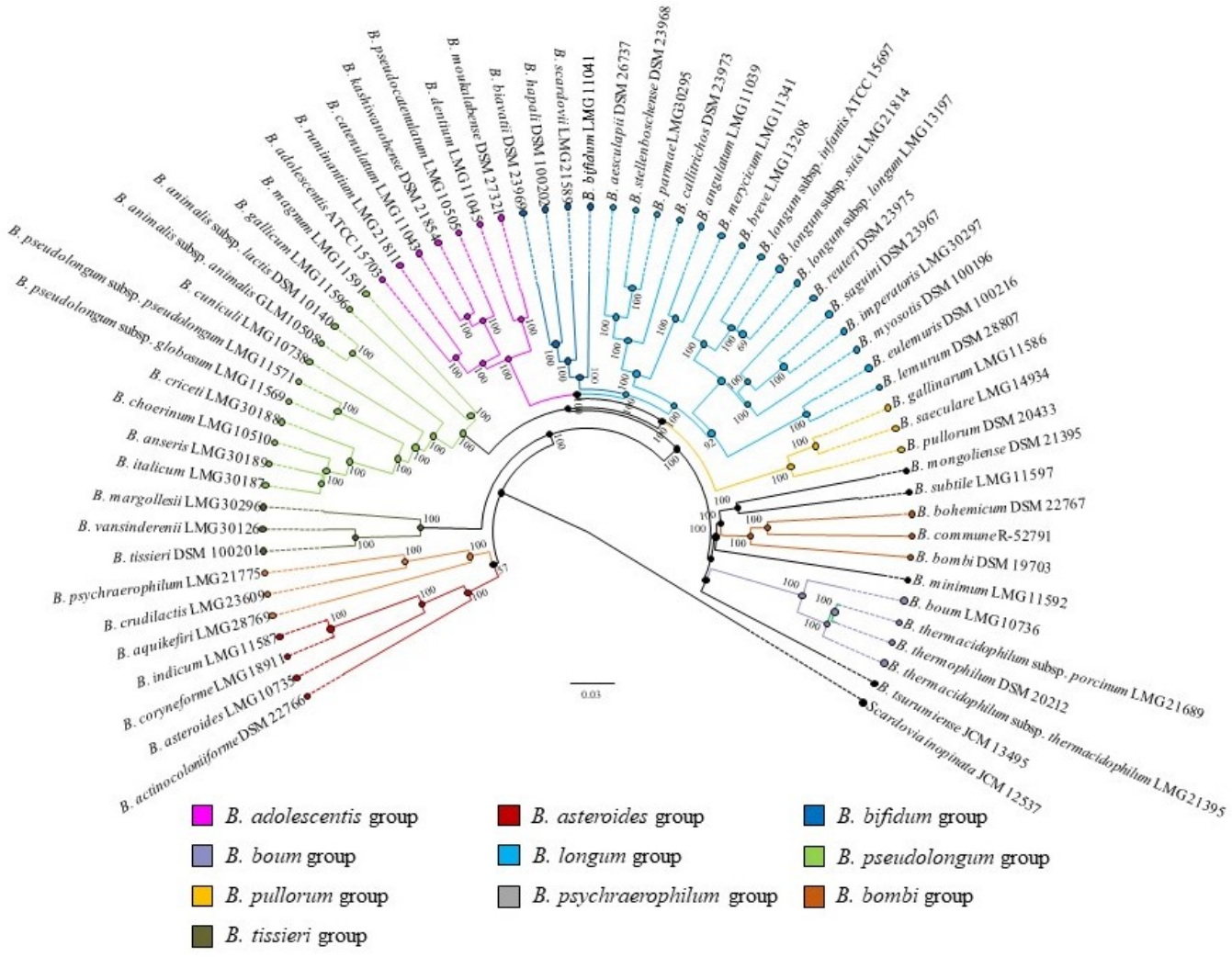


Figure 2. Phylogenomic tree of the *Bifidobacterium* genus based on the concatenation of the amino acid sequences of 232 core genes. The amino-acid deduced core gene-based tree shows the division into 10 phylogenetic groups of the *Bifidobacterium* genus as represented by different colors. The phylogenetic tree was constructed by the neighbor-joining method, with genome sequence of *Scardovia inopinata* JCM 12537 as outgroup (67).

1.6 Ecological origins of the *Bifidobacterium* genus and their adaptation to specific ecological niches.

Focusing on the ecological origin of bifidobacteria, it is important to underline that they have been isolated from many ecological niches such as human GIT (68), the GIT of non-human mammals,

birds and insect (69-72), but also from fermented milk (66), sewage (73), human blood (74) and the oral cavity (74) (Fig. 3).

Interestingly, despite this wide ecological distribution, the adaptation to ecological-niches is a species-specific feature. For example, several *Bifidobacterium* species, i.e. *B. longum*, *B. adolescentis*, *B. pseudolongum* and *B. bifidum*, display a cosmopolitan adaptation, whereas other species, such as *Bifidobacterium angulatum*, *Bifidobacterium cuniculi* and *B. pullorum* subsp. *gallinarum*, show to be adapted to a specific gut ecosystem of specific animals, as cows, rabbits and chickens (68, 75). Interestingly, within typical human GIT bifidobacterial species, there is also a specific adaptation due to the age and the diet of the host. In fact, the most frequently *Bifidobacterium* species individuates in the infant gut are *B. longum* subsp. *infantis*, *B. bifidum* and *Bifidobacterium breve* (56). Whereas, in adulthood, bifidobacteria are numerically less abundant and are limited to few taxa, including *B. longum* subsp. *longum*, *B. adolescentis*, *Bifidobacterium catenulatum* and *Bifidobacterium pseudocatenulatum* (76). The different ecological adaptation to specific environments is due to species-specific genetic features of each bifidobacterial species (77-80). An example of adaptation to the human gut is represented from the ability of *B. bifidum* species to degrade mucin and utilize it as sole carbon source (80-84). Mucin is the main component of the mucus gel layer in the epithelial surface of the GIT (85). Recent studies revealed that this taxon presents a small number of genes coding for carbohydrates carriers and metabolism compared to other species, as well as *B. breve*, *B. longum* subsp. *longum* and *B. longum* subsp. *infantis* (56, 86). In this context, a comparative genomic analysis among 15 members of *B. bifidum* species revealed that most Glycosyl Hydrolases (GHs) were specifically involved in mucin breakdown, which is an important strategy for target host-derived glycans (84, 87).

Another example of adaptation is represented by the ability of bifidobacteria to hydrolyze several oligosaccharides, including plant-derived dietary fibers (88-90). A recent comparative genomics study on 73 strains belonging to *B. longum* species revealed the presence within this taxon of 22 different GH families, including GH families assigned to ferment-plant derived carbohydrates. These

data suggested that members of this species might metabolize “non-digestible” plant polymers or host-derived glycoproteins and glycoconjugates (88). Moreover, the sequencing of *Bifidobacterium dentium* Bd1 genome, a strain isolated from human dental caries, revealed the presence of genes encoding for enzymes involved in the metabolism of sugars, such as fucose, galactose, N-acetylglucosamine and N-acetylgalactosamine, deriving from the degradative action of amylase and lipase present in saliva (79). Furthermore, the genome of the strains *B. dentium* Bd1 shows several genes encoding for proteins associated with acid tolerance, adhesion and defense toward substances, suggesting a genetic adaptation of this strain to colonize the oral cavity (79).

Figure 3. Ecological niches of *Bifidobacterium* genus.

1.7 The genomic of *Bifidobacterium* genus.

The first *Bifidobacterium* genome was sequenced in the early 2000s, corresponding to *B. longum* subsp. *longum* NCC 2705 genome (77). In last years, the number of *Bifidobacterium* genomes has increased exponentially for their importance in promoting health and for the decrease of the NGS techniques cost (66, 91, 92). Nowadays, more than 1000 bifidobacterial genomes were publicly available, of which more than 150 are completely sequenced (NCBI database). The genome length ranged from 1.24 Mb of the strain *B. pseudocatenulatum* 125.1 to 3.28 Mb of the strain *Bifidobacterium myosotis* RST17. The tRNA genes ranged from 91 to 44 and the rRNA operon numbers ranged from two to five, reflecting the adaptation of this genus to different ecological niches (93, 94). The high phylogenetic relationship among different *Bifidobacterium* species is due to the Horizontal Gene Transfer (HGT) events that occurred during the evolution of this genus (92). An important study, investigating on the genome analysis of 47 *Bifidobacterium* (sub)species, revealed that the 81,9 % of identified Open Reading Frames (ORFs) had a specific assigned function, representing the pan-genome of this genus (92). Conversely, 18.1 % of the identified ORFs seemed to be at unknown function (92). Interestingly, the authors of this study showed the identification of 18,181 BifCOGs (*Bifidobacterium*-specific Cluster of Orthologous Genes) among analyzed genomes, of which the majority codified for carbohydrate metabolism enzymes, for proteins involved in amino acids and nucleotides transport and biosynthesis and for housekeeping functions (92). Subsequently, the analysis of 55 members of *Bifidobacterium* genus reconstructed the pan-genome and the core genome (67). As expected, the number of BifCOGs identified in this study was higher respect previous study, i.e. 26,201 BifCOGs (67). Notably, the authors individuated the existence of 262 genes shared from all *Bifidobacterium* strains analyzed, representing the *Bifidobacterium* core genome (67).

Furthermore, the genomic analysis revealed also in this genus the presence of extrachromosomal elements, as well as plasmids, that are able to replicate independently through bacterial replication enzymes. Plasmids are not present in all *Bifidobacterium* strains and their size ranged from 2,1 kb to

11,1 kb (NCBI database). Nowadays, 36 *Bifidobacterium* plasmids were identified from different *Bifidobacterium* species, including *B. asteroids*, *B. breve*, *B. bifidum*, *B. catenulatum*, *Bifidobacterium choerinum*, *B. pullorum* subsp. *gallinarum*, *B. longum, B. pseudocatenulatum* and *B. pseudolongum*. These plasmids seemed not encode for phenotypic features, except for the plasmid NCFB 1454 which was identified in the strain *B. bifidum* NCFB 1454. In fact, this plasmid presents a bacteriocin-encoding gene, specifically it encodes for the bacteriocin bifidocin B (95).

Moreover, bifidobacteria genomes have been also investigating for the presence of prophage-like sequences. Among the Gram-positive bacteria, the identification and analysis of *Bifidobacterium* prophage, also known as bifidoprophages, revealed that phage infections do play a role in the genetic make-up of this genus (96-99). The bifidoprophages genetic architecture was shown to be like that of lambdoid phages, including genes for integration, immunity, as well as DNA replication, DNA packaging, morphogenesis and lysis (100). A recent study based on an extensive *in silico* analysis of 48 sub(species) *Bifidobacterium* genomes catalogued the recognized bifidoprophages (99). In this study, the comparative genomic analysis of 90 bifidoprophages classified, including complete and remnant phages, shows the existence of five bifidoprophage homology groups (99). These prophage-like elements seemed to be acquired through HGT events, allowing bifidobacteria to adapt to different ecological niches (94, 99, 101).

Focusing on transposable elements, different studies have been demonstrated the presence of Insertion Sequence (IS) elements and transposase in *Bifidobacterium* genomes. IS elements are short DNA sequences able to move within genomes, redesigning the genome sequence, whereas transposases are enzymes that act by a cut and paste mechanism or by a replicative mechanism. For example, an IS element was identified in the genome of the strain *B. longum* subsp. *longum* F8 upstream in respect to the *tetW* gene, conferring to this strain resistance toward tetracycline antibiotic and with the possibility to transfer this resistance to other microorganisms or in other portions of the genome. Interestingly, the same IS elements was identified in the genome of the strain *B. longum*

subsp. *longum* M21, but it is placed in a different genome position and this strain resulted sensible toward tetracycline (102, 103).

Finally, additional putative mobile elements identified in *Bifidobacterium* genus are represented by CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) loci (92). These loci are associated to Cas (CRISPR associated) proteins, forming a molecular system that provides adaptive immunity against exogenous genetic elements in bacteria and archaea (104). In this context, DNA from invasive elements is captured in CRISPR loci and transcribed in small interference RNAs that guide Cas nucleases for sequence-specific targeting and cleavage of cDNA (105). In bifidobacteria three main types of CRISPR-Cas system are identified, named type I, type II and type III, with the high occurrence for the CRISPR-Cas type I (92).

As some whole, mobile elements, i.e. plasmids, bacteriophage, IS elements and CRISPR loci, constitute the mobilome of the *Bifidobacterium* genus, which plays an important role in HGT events and in the evolution of these microorganisms.

1.8 Carbohydrates metabolism of the *Bifidobacterium* genus.

Bifidobacteria are able to utilize a widespread type of diet-derived carbohydrates, which can be derived from plants, such as starch, glycogen, cellulose, galactan, xylan, or can be derived from the host, including mucin and Human Milk Oligosaccharides (HMOs). Different carbohydrate-modifying/transporting enzymes have been identified in bifidobacteria, such as glycosyl hydrolases (GH), PEP-phosphoenolpyruvate-PTS-phosphotransferase systems (PEP-PTS systems) and sugar ABC transporters. Notably, all bifidobacteria metabolize hexose sugars through the fructose-6-phosphate phosphoketolase (F6PPK) enzyme in a phosphoketolase pathway called “bifid shunt” (86). Several studies revealed that different *Bifidobacterium* species could degrade different complex sugars, which may help these microorganisms to adapt to their ecological niches (86, 106-108). Comparative analysis of *Bifidobacterium* genus, as previously described, revealed that approximately 14 % of bifidobacterial genome is involved in carbohydrates metabolism, including GHs,

glycosyltransferases (GTs) and esterase's enzymes (92, 107, 108). In particular, GHs are considered key enzymes that allow bifidobacteria to adapt and colonize the host environment through the hydrolysis of different complex carbohydrates (86). Specifically, according to the Carbohydrate Active Enzyme (CAZy) system, GH3, GH13, GH43 and GH51 families resulted to be the highly abundant GH families recognized in bifidobacteria and they are widely distributed in all *Bifidobacterium* species (106, 109). Among these GH families, the family GH13, which is predicted to be involved in the degradation of starch and starch-derivatives, is most commonly found in bifidobacterial genomes, especially for those species isolated from the mammalian gut (106). Furthermore, different studies revealed that bifidobacterial glycobiome encompassed also GH families involved in host glycan breakdowns, such as those belonging to GH33 and GH34 (exo-sialidases), GH29 and GH30 (fucosidases) and GH20 (hexosaminidases) (106, 107). Notably, different *in silico* analyses revealed that the genetic arsenal involved in the breakdown of host-indigestible carbohydrates, derived from the host (i.e., mucin and HMOs) or from the diet, had a significant role on the composition of the gut microbiota (42). Regarding the degradation of mucin, *in silico* and *in vitro* studies showed that different *Bifidobacterium* species have been shown to degrade mucin, including *B. bifidum*, *B. longum* subsp. *infantis*, *Bifidobacterium biavatii*, *B. crudilactis*, *B. catenulatum* subsp. *kashiwanohense*, *Bifidobacterium stellenboschense* and *Bifidobacterium mongoliense* species (81, 106, 110, 111). Among these species, *B. bifidum* resulted to degrade mucin more efficiently than the other species (106). *In silico* analysis revealed the existence of genes involved in mucin metabolism presented uniquely in genomes of members of *B. bifidum* species, suggesting a strictly co-evolution of this species to the human gut (80, 82, 87). These genes seemed to codify for *N*-acetyl-β-hexosaminidases, β-galactosidases and exo-α-sialidases enzymes, all involved in mucin-metabolism (80, 82). Another important example of genetic adaptation to the human gut is the HMOs degradation capability of *Bifidobacterium* species. HMOs are host-glycans constituted by oligosaccharides present in human milk and that are not utilized by the infant host. In this context, *B. longum* subsp. *infantis*, *B. bifidum* and *B. breve* species are shown

to metabolize HMOs (78, 112). Interestingly, the genome of *B. longum* subsp. *infantis* species was shown to encompass GHs and carbohydrate transporters necessary for the import and the degradation of HMOs, including fucosidases, sialidases, β-hexosaminidases and β-galactosidases enzymes, extracellular solute binding proteins and permeases necessary for the transport of HMOs (78, 113-115). An additional considerable feature of genetic adaptation of members of *Bifidobacterium* genus to the human gut is the specific utilization of starch and its derivatives, such as maltodextrin, maltotriose and maltose. Different *in silico* analysis revealed that *B. adolescentis* and *B. breve* species presented in their genomes a complete gene set for the degradation of starch (116-118). Interestingly, the glycobiome prediction of *B. adolescentis* taxon revealed that the GH13 family is the most present. This GH family includes amylase, pullulanase and cyclomaltodextrine enzymes that are involved in plant-derived carbohydrate metabolism (107, 119). Moreover, also the fermentation profiles of several strains of *B. adolescentis* species highlighted the preference of this taxon for the utilization of plant-derived glycans that are typically present in the adult human diet (107, 117). All these different host-derived glycans metabolism abilities suggest the niche-specific adaptation among different bifidobacterial species (86).

1.9 Antibiotic resistance (AR) of the *Bifidobacterium* genus.

Bacterial Antibiotic Resistance (AR) is a widespread relevant phenomenon that is due by the increasingly extensive use of antibiotics, which facilitate growth and development of resistant microorganisms, removing sensible bacteria (120). AR could be an intrinsic or an acquire feature, with the latter that derives from the alteration of the antibiotic function through different mechanisms, such as the modification of the antimicrobial target, the decreased absorption of the drug, the activation of efflux mechanisms or the expression of enzymes counteracting antibiotic molecules (121). Microorganisms could acquire ARs through mutation in genes encoded for proteins involved in the antibiotic action mechanism, or through HGT events with the acquisition of exogenous DNA

codifying for AR determinants (121). A clear example of AR was investigated in *Bifidobacterium* genus being a typical commensal genus (122-124). In this context, very important is the resistance of this genus to the mupirocin antibiotic that is produced by *Pseudomonas fluorescens* (Fig. 4) (125).

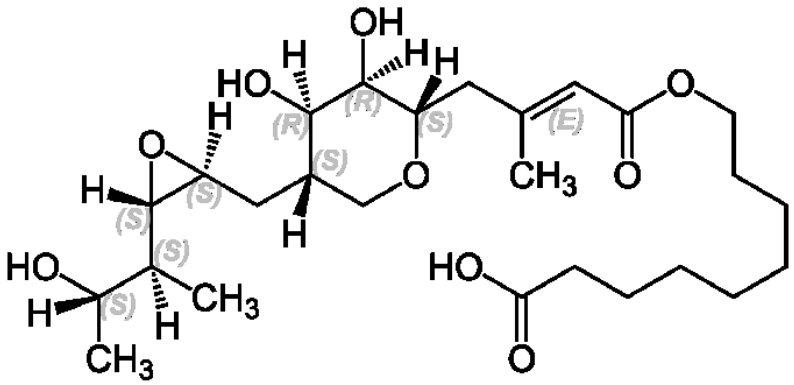


Figure 4. Molecular structure of mupirocin antibiotic.

The antibacterial activity of mupirocin is due to the competitive inhibition as substrate of the isoleucyl-tRNA enzyme with isoleucine amino acid, inhibiting the correct enzyme function (126). In fact, the mupirocin, that has a molecular structure very similar to that of the isoleucyl adenylate complex, is able to inhibit the amino acylation process and the synthesis of the isoleucyl adenylate (123). The mupirocin-resistance mechanism involved also another gene, known as *IleS* that encodes for a modified isoleucyl-tRNA synthetase, possessing a RG1 region that is able to reduce the mupirocin bond stability in the active site (123). This modification prevents the mupirocin competitive inhibition activity, making bifidobacteria resistant toward this antibiotic (123).

Another example of AR is the resistance of members of *Bifidobacterium* genus toward tetracycline antibiotic (124). In this context, different strains of *Bifidobacterium animalis* subsp. *lactis* (sub)species presented in their genomes a *tetW* gene, encoding for a GTPase that modify the tetracycline target size, protecting the ribosome (124). Interestingly, the strain *B. animalis* subsp. *lactis* Bl12 resulted sensible toward tetracycline antibiotic, differently than several other strains of this species (127). The genome sequencing allowed to identify the presence of the *tetW* gene also in *B. animalis* subsp. *lactis* Bl12 strain, but it resulted more sensitive to tetracycline in respect other

strains of the same species (127). Notably, this different behavior was due to a single polymorphism in the *miaA* gene, which resulted to influence the *tetW* mechanism. In particular, this polymorphism regards the substitution of a thymine leading to the formation of a glutamine residue instead of a lysine. This single polymorphism in the *miaA* sequence led to the coding of a different protein, decreasing the tetracycline resistance associated with *tetW* gene (127). Moreover, others genes resulted overexpressed in *B. animalis* subsp. *lactis* genomes when these strains were grown in the presence of tetracycline, as well as, *tetP* gene, codifying for an ulterior GTPase and the *tet(36)* gene, codifying for a specific tetracycline efflux pump (124).

1.10 Interaction between bifidobacteria and the human GIT.

As mentioned above, the GIT microbiota consists of a complex ecosystem with different metabolic and physiological activities. Recently, several studies showed various mechanisms by which bifidobacteria interact with the host, including the presence of genes encoding for extracellular structures, as well as exopolysaccharides (EPS), serine protease inhibitors and cell surface-encoding proteins, called pili (128-131).

EPS are carbohydrate polymers present as an extracellular layer on the surface of different microorganisms, including Gram-positive and Gram-negative bacteria (132). EPS can have two functions, the first one is to act as virulence factors in several diseases and the second is promoting health benefits, supporting the colonization of gut commensals (133-137). The EPS synthesized by members of *Bifidobacterium* genus are formed by different monosaccharides, i.e. D-glucose, D-galactose and L-rhamnose, and they are classified as hetero-polysaccharides (138). Several studies showed that the immune response is conjugate to the production of EPS (138, 139). Moreover, it is demonstrated that EPS can act as antioxidants to decrease the damage caused by reactive oxygen species in tissues (140) and they play a role in the modulation, in the composition and in the activities of the gut microbiota (138). Recently, the molecular characterization of the bifidobacterial EPS-ome

(e.g. all identified bifidobacterial genes that represent all predicted functions for EPS production), showing the existence of nine different EPS-coding genes genetic clusters conserved in all *Bifidobacterium* species and the presence of 44 species-specific clusters (130). Moreover, it was demonstrated that members of *Bifidobacterium* genus modulated the EPS production when they are cultivated in a simulating-infant gut environment (130).

Additionally, another important feature used by bifidobacteria to interact with the host is represented by pili-like structures. In particular, there are two types of pili in bifidobacteria genus: the sortase-dependent (SD) pili (types I and II) and the Tight Adherence Pili (Tad – types IV). These structures are involved in bacterial adhesion to the surface environments and they stimulate the immune response of the host (128). A recent study investigating the bifidobacterial SD pili revealed the existence of 15 different clusters of pili-encoding genes and the presence of species-specific and strain-specific clusters (131). Moreover, it is demonstrated that these pili-like structures are crucial in adhesion to glycans and extracellular matrix proteins, suggesting the ecological ability of bifidobacteria to colonize the human gut (131). Furthermore, the first locus encoded for Tad pili in bifidobacteria was individuated in the strain *B. breve* UCC2003 (141). Transcriptome analysis of the *B. breve* UCC2003 upon colonization in mice revealed an up-regulation of genes encompassing the *tad* locus, which does not result to be expressed when the strain was cultivated *in vitro* (141). Additionally, experiments based on mutational analysis revealed the importance of *tad* locus in the colonization and persistence of *B. breve* UCC2003 in the mammalian gut (141). Notably, the *tad* locus is highly conserved among all the bifidobacterial species, supporting the role of this locus in the host colonization (42).

Other bifidobacterial encoded proteins involved in the interaction between bacteria and the host are represented by the serine protease inhibitors, which are members of serpin superfamily (142, 143). For example, the strain *B. longum* NCC 2705 encodes for a serpin-like protease inhibitor that contributes to the interaction with the host in the GIT (142). This enzyme seemed to play an important role in the interaction between this strain and the host, inhibiting human neutrophil and pancreatic

elastases (142). Another study featured a serpin-encoding gene in the genome of the human mucosal-isolated strain *B. breve* 210B (129). Different analysis based on microarray and quantitative real-time PCR (qRT-PCR) approaches showed that the polycistronic mRNA, including the *ser* operon, is induced following treatment of this strain with gut proteases, as well as pancreatic elastase, human neutrophil elastase, thrombin, papain, kallikrein, trypsin, α-antitrypsin, chymotrypsin or plasmin (129). In this study, *Turroni et al.* highlighted that there was an anti-inflammatory response due to the activity of the serine protease inhibitor, lowering the negative effect of this protein in the intestinal inflammation (129).

1.11 The vertical transmission of bacteria from mother to newborns.

Until few years ago, the human infant intestine is considered sterile before the birth (144), but recent studies suggested the presence of bacterial colonization before the delivery (43, 145-147). A recent metagenomics study, based on NGS techniques, investigated the newborns microbiota immediately after the birth (148). These data highlighted that in newborns intestinal secretions, in the meconium and in the amniotic fluid, there was the presence of some microorganisms belonging to *Streptococcus*, *Lactococcus*, *Shigella*, *Escherichia*, *Enterococcus* and *Leuconostoc* genera (148). In this context, the period immediately after the birth is considered crucial for the correct development of the gut microbiota (148). It is shown a correlation between the type of the birth, i.e. natural birth or Caesarian Section (CS), and the gut microbiota composition of newborns (41, 149, 150). In fact, the passage of newborns through the vaginal tract and subsequent breastfeeding are considered primary factors that contributed to the microbial colonization of the GIT (148, 151). In this context, the type of the birth is very important because influence the composition of the newborn gut microbiota. In fact, different studies have been demonstrated that the microbial composition of children born with natural delivery is initially colonized from bacteria residing in the mother's urogenital tract, as well as members of *Lactobacillus* and *Bifidobacterium* genera (148, 151-153). Conversely, children born through CS are not exposed to maternal bacteria presenting a low number of microorganisms belonged to

Bifidobacterium genus and to Bacteroidetes phylum, whereas they are colonized from bacteria presents on the skin or in the environment, as members of *Streptococcus*, *Corynebacterium* and *Propionibacterium* genera (56, 154). The type of the feeding represents another important feature influencing the composition of the newborns gut microbiota. The human milk is a complex biological fluid that is necessary for the nutritional needs and for the immune system development in newborns (155). These important characteristics are due to the molecular composition of human milk, in fact this biological fluid includes immunoglobulins, fatty acids, polyamines, oligosaccharides, lysozyme, lactoferrin, other glycoproteins and antimicrobial peptides (155). Interestingly, the human milk microbiota reaches a very important complexity in the gestation period and after the delivery, remaining constantly during the breastfeeding (155). The microbial composition of human milk is predominated from two phyla: Proteobacteria and Firmicutes (156). In this context, *Staphylococcus*, *Pseudomonas*, *Streptococcus* and *Lactobacillus* genera result to be the most abundant genera in human milk (156). Furthermore, recent studies highlighted the presence of members of *Staphylococcus* and *Streptococcus* genera in the gut microbiota of breastfed newborns in respect the gut microbiota of newborns artificially feeding (157, 158). Notably, after the first week of life the number of bacteria belonged to these genera decrease for the low oxygen level, building an environment that promotes the colonization of anaerobic microorganisms (159).

One of the most important bacterial component of the human milk is represented by the *Bifidobacterium* genus, which are able to metabolize HMOs. HMOs are one of the most abundant molecular classes composing the human milk together with fatty acids and lactose (160, 161). Bifidobacteria are considered among the first colonizer of human gut and they are able to reach high abundant level in the first two weeks of life together with members of the Firmicutes phylum (162). Several recent studies revealed that the taxonomic classification of bifidobacteria in the mother gut microbiota is strictly correlated with bifidobacteria present in the newborns gut, suggesting the hypothesis of a Vertical Transmission (VT) of these microorganisms from mother to child (43, 153). In this context, recent studies based on Internal Transcribed Spacer (ITS)-profiling investigated on

Bifidobacterium species shared between mothers and newborns (148, 151). In these studies, the high level of members of *Bifidobacterium* species shared from mothers and newborns was investigated, considering different samples-sets made up of mothers and their newborns (148, 151). In particular, *B. adolescentis*, *B. angulatum*, *B. breve*, *B. dentium, B. longum*, *B. pseudolongum* and *Bifidobacterium thermacidophilum* species are demonstrated to be shared between mothers and newborns (148, 151). Combining NGS approaches, i.e., ITS-bifidobacterial profiling and shotgun metagenomics, with qRT-PCR, these studies confirmed the existence of a VT route of bifidobacterial communities from mothers to children (148, 151). Furthermore, they highlighted how bifidobacteria are inherited from the mother, implying human milk as vehicle to facilitate this acquisition, also for the ability of some *Bifidobacterium* species to utilize HMOs and HMO-derived glycans as carbon sources (148). In addition, the authors highlighted the existence of a bifidoprophage VT, facilitating their spread from mothers to newborns (148).

Chapter 2

Outline of the thesis

The aim of this Ph.D. thesis research is to investigate the biology of members of the *Bifidobacterium* genus. In particular, comparative genomic studies and Horizontal Gene Transfer (HGT) events among bifidobacteria have been investigated taking into account fully sequenced genomes. Since bifidobacteria is one of the first colonizers of the infants' gut, the vertical transmission (VT) between mothers to puppies of *Bifidobacterium* species, has been also examined. Furthermore, since bifidobacteria are intestinal microorganisms, the impact of antibiotic molecules has been explored through *in silico* and *in vitro* experiments. Finally, the last section of the thesis focuses on the investigation of the vaginal microbiota composition and the role of one of its main colonizer, such as *Lactobacillus crispatus* species.

Chapter 3 describes a comparative genomics analysis of members of the *Bifidobacterium animalis* species, including both subspecies, i.e., *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis*. In this chapter, the evolutionary differentiation between these two taxa was investigated through *in silico* analysis and experimentally evaluated by physiological *in vitro* experiments.

Chapter 4 dissects the VT of different *Bifidobacterium* strains through an *in vivo* animal model (*Rattus norvegicus*). Moreover, this chapter investigated the hypothesis of the possible colonization of the puppies' gut before the natural birth and on the existence of a pre-birth microbiota.

Chapter 5 illustrates an *in silico* analysis of the resistome and the mobilome of the *Bifidobacterium* genus. In addition, the association of mobilome and resistome data has allowed the characterization of genetic insertion signature that could be involved in HGT events.

Chapter 6 focuses on the effect of the Amoxicillin - Clavulanic acid (AMC) antibiotic on the gut microbiota. Furthermore, a screening involving different *Bifidobacterium* strains was performed to evaluate their resistance toward this antibiotic. Finally, the role of AMC resistance *Bifidobacterium* strains was investigated through *in vitro* batch-experiments.

Chapter 7 discusses the vaginal microbiota composition, underlining the healthy effect of *L. crispatus* species on this environment. In this chapter, a genomic analysis of *L. crispatus* species was

performed, i.e., pan-genome analys of different *L. crispatus* strains, and the role of *L. crispatus* on human vaginal microbiota was investigated through *in-vitro* batch experiments.

Chapter 3

Dissecting the evolutionary development of the *Bifidobacterium animalis* species through comparative genomics analyses

Lugli GA*, Mancino W*, Milani C, Duranti S, Mancabelli L, Napoli S, Mangifesta M, Viappiani A, Anzalone R, Longhi G, van Sinderen D, Ventura M, Turroni F.

The results of this chapter were published on Applied and Environmental Microbiology, 2019. 85:e02806-18. https://doi.org/10.1128/AEM.02806-18.

*These authors contribuited equally.

ABSTRACT

Bifidobacteria are members of the gut microbiota of animals, including mammals, birds and social insects. In this study, we analyzed and determined the pan-genome of *Bifidobacterium animalis* species, encompassing *Bifidobacterium animalis* subsp. *animalis* and the *Bifidobacterium animalis* subsp. *lactis* taxon, which is one of the most intensely exploited probiotic bifidobacterial species. In order to reveal differences within the *B. animalis* species, detailed comparative genomics and phylogenomics analyses were performed, indicating that these two subspecies recently arose through divergent evolutionary events. A subspecies-specific core genome was identified for both *B. animalis* subspecies, revealing the existence of subspecies-defining genes involved in carbohydrate metabolism. Notably, these *in silico* analyses coupled with carbohydrate profiling assays suggest genetic adaptations toward a distinct glycan milieu for each member of the *B. animalis* subspecies, resulting in a divergent evolutionary development of the two subspecies.

IMPORTANCE

The majority of characterized *B. animalis* strains have been isolated from human fecal samples. In order to explore genome variability within this species, we isolated 15 novel strains from the GIT of different animals, including mammals and birds. The current study allowed us to reconstruct the pan-genome of this taxon, including the genome contents of 56 *B. animalis* strains. Through careful assessment of subspecies-specific core-genes of the *B. animalis* subsp. *animalis*/*lactis* taxon, we identified genes encoding enzymes involved in carbohydrate transport and metabolism, while unveiling specific gene-acquisition and -loss events that caused the evolutionary emergence of these two subspecies.

For Supplementry Materials see the article published on *Applied and Environmental Microbiology*.

MATERIALS AND METHODS

Bifidobacterial selection.

In order to explore genome variability of the *B. animalis* species, 15 novel strains were isolated from fecal samples collected from different animals. Samples were composed of 10 g of fresh fecal material, which is a sufficient quantity to represent the overall biodiversity of the fecal microbiota as reported in a previously published study (163). One gram of fecal sample from each collected animal was mixed with nine ml of phosphate-buffered saline (PBS; pH 6.5). Serial dilution and subsequent plating were performed using de Man–Rogosa–Sharpe (MRS) agar, supplemented with 50 μg/ml mupirocin (Delchimica, Italy) and 0.05 % (wt/vol) L-cysteine hydrochloride. Agar plates were incubated for 48 h at 37 °C in a chamber (Concept 400; Ruskin) with an anaerobic atmosphere (2.99 % H_2, 17.01 % CO_2 and 80 % N_2). Morphologically different colonies that developed on MRS plates were randomly picked and re-streaked in order to isolate purified bacterial strains. All isolates were subjected to DNA isolation and characterized as previously described by Turroni *et al.* (164). The *B. animalis* strains isolated in this study are listed in Table 1, together with other strains used for *in silico* analyses.

Bifidobacterial ITS profiling.

Partial ITS sequences were amplified from extracted DNA using the primer pair Probio-bif_Uni/Probio-bif_Rev (165). Resulting reads were analyzed by means of an updated bifidobacterial ITS database encompassing all publicly available bifidobacterial genomes and a custom bioinformatics script as previously described (165). ITS bifidobacterial profiling of mammalian species and birds were coupled with data of mammalian bifidobacterial communities as previously determined by Milani *et al.* (75).

Genome sequencing and assemblies.

DNA extracted from bifidobacterial isolates was subjected to whole genome sequencing using MiSeq (Illumina, UK) at GenProbio srl (Parma, Italy) following the supplier's protocol (Illumina, UK). 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 (166). SPAdes software was used for *de novo* assembly of each bifidobacterial genome sequence (167, 168), while protein-encoding open reading frames (ORFs) were predicted using Prodigal (169). The coverage depth of these newly isolated 15 *B. animalis* chromosomes ranged from 85- to 278-fold, which upon assembly generated 47 to 12 contigs (Table 1). The number of predicted ORFs ranged from 1556 of *B. animalis* subsp. *lactis* 1808B to 1935 of *B. animalis* subsp. *animalis* 2022B (Table 1). In order to ensure data consistency, *B. animalis* chromosomes retrieved from public databases were re-annotated using the same bioinformatics pipeline applied for the 15 *B. animalis* strains isolated in the current study.

.

Comparative genomics.

A pan-genome calculation was performed using the pan-genome analysis pipeline PGAP (170), including each *B. animalis* genome collected from this study (Table 1). Each predicted proteome of a given *B. animalis* strain was screened for orthologues against the proteome of every collected *B. animalis* strain by means of BLAST analysis (171) (cutoff: E value of $< 1\times10^{-4}$ and 50 % identity over at least 80 % of both protein sequences). The resulting output was then clustered into protein families by means of MCL (graph theory-based Markov clustering algorithm) (172), using the gene family (GF) method. A pan-genome profile was built using all possible BLAST combinations for each genome being sequentially added. Using this approach, unique protein families encoded by the analyzed *B. animalis* genomes were also identified. Protein families shared between analyzed genomes allowed us to identify the core genome of the *B. animalis* species. Each set of orthologous proteins, belonging to the core genome, was aligned using Mafft software (173), and phylogenetic

trees were constructed using ClustalW (174). Based on these comparative analyses, a *B. animalis* supertree was constructed and visualized using FigTree (http://tree.bio.ed.ac.uk/software/figtree/).

Carbohydrate growth assays.

Bifidobacterial strains were cultivated on semi-synthetic MRS medium supplemented with 1 % (wt/vol) of a particular sugar and optical densities (OD measurements at a wave length of 600 nm) were recorded using a plate reader (BioTek, Winooski, VT, USA). The plate reader was read in intermittent mode, with absorbance readings performed at 3 minute intervals for 3 times after 48 h of growth, where each reading was ahead of 30 s shaking at medium speed. Cultures were grown in biologically independent triplicates and the resulting growth data were expressed as the mean of these replicates. Carbohydrates were purchased from Sigma and Carbosynth (Berkshire, UK). Carbohydrate-active enzymes were identified based on similarity to the carbohydrate-active enzyme (CAZy) database entries.

Single Nucleotide Polymorphism (SNP) identification.

Multiple alignment of conserved genomic sequence with rearrangements (Mauve) software (175) was employed to perform whole-genome sequence alignments between bifidobacterial genome sequences. SNPs reported by Mauve were manually evaluated to identify polymorphisms between subspecies.

Gene gain/loss through evolutionary reconstruction.

Identification of genes that are predicted to be acquired by Horizontal Genes Transfer (HGT) event was performed using COLOMBO v4.0 (176). Evolution-driven acquisition and loss of GH-encoding genes among members of the *B. pseudolongum* phylogenetic group was performed with Count (177) software using Wagner's parsimony.

Statistical analyses.

SPSS software (IBM, Italy) was used to perform statistical analysis between *B. animalis* subsp. *animalis* strains group and *B. animalis* subsp. *lactis* group by T-student test. Furthermore, t-test assumption was verified using the unequal variances Welch t-test analysis to validate samples that exhibit unequal variance in the sample size (Table S1).

Data Deposition.

Newly isolated *B. animalis* genomes were sequenced and deposited at DDBJ/ENA/GenBank under the accession numbers reported in Table 1 (BioProject PRJNA506409).

RESULTS AND DISCUSSION

Isolation and genetic characterization of the *B. animalis* species.

To investigate the occurrence of *B. animalis* in the gut of animals, we screened the Internally Transcribed Spacer (ITS) sequence profiling data derived from fecal samples of four mammalian and bird species together with the bifidobacterial community data previously determined by Milani *et al.* (75) (Fig. 1). In this context, *B. animalis* was detected in 55 % of such fecal samples, with a higher occurrence in the fecal samples of dogs (*Canis lupus*), onagers (*Equus hemionus kulan*), monkeys (*Chlorocebus pygerytrhus, Macaca fuscata, Macaca Sylvanus* and *Pan Troglodytes*) and mice (*Mus musculus*) (Fig. 1). These data revealed a cosmopolitan lifestyle of this taxon, underlining the potential high genetic adaptation of *B. animalis* strains to different (host) environments.

In order to investigate the genetic contents of the *B. animalis* species, including representatives of both *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* taxa, we applied a bifidobacterial isolation protocol on fecal samples of those animal species displaying a high abundance of these taxa. The above-mentioned analyses (see Materials and Methods) (178, 179) allowed the isolation of 15 novel *B. animalis* strains from birds (*Phasianus colchicus*) and various Mammalia, such as canine breeds, i.e., German shepherd, Pomeranian, Alaskan malamute and Flat coated retriever, and three different non-human primates, i.e., *Pan troglodytes, Chlorocebus pygerythrus* and *Macaca Sylvanus*. Moreover, *B. animalis* subsp. *lactis/animalis* strains were isolated from fecal samples of rabbits (*Oryctolagus cuniculus*), beavers (*Castor fiber*) and pigs (*Sus scrofa domesticus*) (Table 1 and Fig. 1). Interestingly, we were also able to isolate different *B. animalis* strains from stool samples of animals in which the ITS bifidobacterial profiling analysis indicated a low relative abundance of this species, i.e. *Sus scrofa* (0.06 %) and *Oryctolagus cuniculis* (0.002 %). This may be due to the better growth performance (e.g., high tolerance to environmental stresses) of members of the *B. animalis* species compared to other bifidobacteria (180-182). A comparative genomic analysis between newly isolated strains was complemented with the inclusion of publicly available genomic repertoire of 41 *B. animalis* strains, thereby exemplifying a broad ecological representation, including the GIT of

human and other animals (e.g., rat and chicken) (183), as well as different food matrices (e.g., milk and yogurt) and human vaginal swab (127, 184) (Table 1). This information further validates the notion that *B. animalis* seems to be genetically adapted to a large number of habitats. Notably, the ORFeome of *B. animalis* subsp. *animalis* strains, defined as the complete set of open reading frames in genomes of the same species, was shown to be substantially larger when compared to that of *B. animalis* subsp. *lactis* strains, suggesting that members of the *B. animalis* subsp. *animalis* taxon exhibit a more extensive level of genetic diversity.

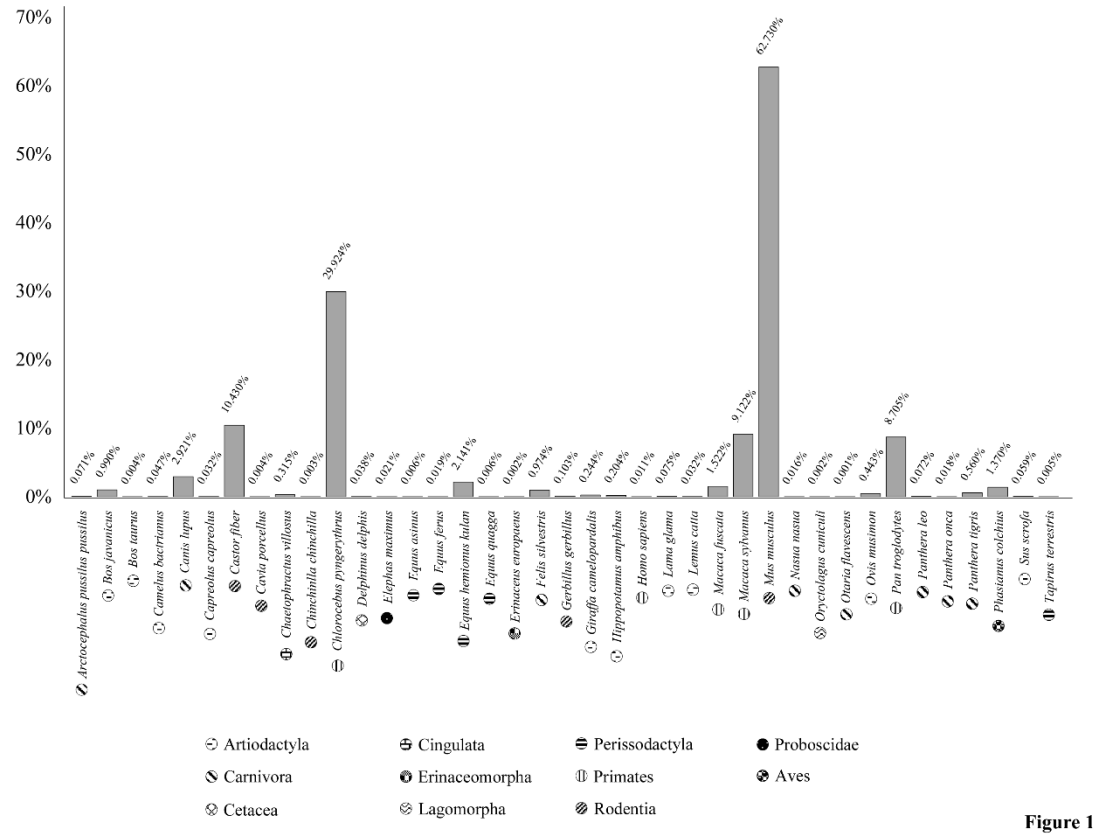


Figure 1

Figure 1. *Bifidobacterium animalis* profiling data obtained from faecal samples of different animals. In this bar-plot, the *x-axis* represents the animals tested for the presence of *B. animalis*, while the *y-axis* represents the percentage of *B. animalis* compared to other *Bifidobacterium* species present in the samples. Each color represents an animal order, as indicated in the legend.

Pan-genome and core-genome analysis of *B. animalis* species.

The reconstructed genomic datasets of the *B. animalis* species, encompassing a total of 56 chromosomal sequences, represents the genetic catalogue for this bifidobacterial species. The genetic

makeup of the whole taxon was employed to predict the pan-genome of the *B. animalis* species, i.e., the available collection of genes from strains of a given species (185). Moreover, these data were used to predict also the core-genome, i.e., the collection of gene families shared between organisms of a given species, i.e. the *B. animalis* taxon, as based on the clusters of orthologous groups (COGs) (186). The pan-genome size, consisting of 4,486 COGs, when plotted on a log-log scale as a function of the number of analyzed genomes, suggests that the power trend line has almost reached a plateau (Fig. 2). The average number of new genes discovered by sequential addition of genome sequences decreased from 130 COGs upon the addition of another genome, to 30 COGs in the final addition (Fig. 2). Thus, these findings indicate that genome sequencing of additional (novel) *B. animalis* strains are expected to increase the pan-genome size by less than 0.7 % (Fig. 2). Furthermore, the 56 *B. animalis* genomes were screened to identify shared orthologous genes as well as unique genes. *In silico* analyses reveals that 1,098 ORFs were shared between the assessed strains, representing the core-genome of this species. The functional examination of the core-genome, based on the eggNOG database (187), reveals that the 26.1 % of the identified core genes are predicted to encode housekeeping functions and enzymatic activities related to amino acid and carbohydrate metabolism, and their corresponding transport.

When we separately analyzed the core-genome of strains belonging to a specific *B. animalis* subspecies, subspecies-specific core genes could be identified (Fig. 2). In this context, 142 subspecies-specific genes were retrieved in the genomes of the *B. animalis* subsp. *animalis* subspecies, while just 82 were detected in the chromosome sequences of *B. animalis* subsp. *lactis* members. The existence of specific conserved genes between the two subspecies is suggestive of an evolutionary separation between these bifidobacterial taxa. Specifically, genes that have driven this differentiation are expected to be among the subspecies-specific core and include genes that are predicted to encode transporters and carbohydrate active proteins, i.e. 51 in the *B. animalis* subsp. *animalis*- and 31 in the *B. animalis* subsp. *lactis*-specific core genome, respectively (Table S2). Interestingly, the higher number of the above-mentioned genes in the *B. animalis* subsp. *animalis*-

specific core genome when compared to the corresponding number in the *B. animalis* subsp. *lactis*-specific core genome, suggests that *B. animalis* subsp. *animalis* strains are able to metabolize a larger number of glycan substrates compared to *B. animalis* subsp. *lactis* strains (Table S2). Furthermore, the subspecies-specific core genomes include various DNA binding proteins, with a distinctly higher abundance in *B. animalis* subsp. *animalis* (13 genes) as compared to *B. animalis* subsp. *lactis* (three genes), of which five belong to MarR family of transcriptional regulators (Table S2). Altogether, the observed differences in the number of subspecies-specific core genes between the *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* were shown to be statistically significant (P-value < 0.05).

An *in silico* approach was employed to calculate the Average Nucleotide Identity (ANI) values, defined as a measure of nucleotide-level genomic similarity between the coding regions of two genomes, between *B. animalis* genomes (188), showing a highly syntenic genome structure among members of this species, with associated ANI values ranging from 95.81 % to 99.99 %. Moreover, different ranges of ANI values were identified between strains belonging to *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis*. Interestingly, the lowest ANI value between *B. animalis* subsp. *lactis* genomes was 98.7 %, while for *B. animalis* subsp. *animalis* genomes this number was 96.1 %. These data reflect the differences between these two subspecies, highlighting a highly syntenic genome structure among members of the *B. animalis* subsp. *lactis* subspecies. This statement was further validated by the pan-genome analysis mentioned above that allowed us to highlight Truly Unique Genes (TUGs) of each *B. animalis* strain (Fig. 2). In this context, a variable number of TUGs, ranging from zero for 23 *B. animalis* subsp. *lactis* strains to 144 genes for *B. animalis* subsp. *animalis* 2022B, were detected (Fig. 2). Thus, the absence of TUGs within the majority of *B. animalis* subsp. *lactis* strains supports the previously noted high isogenic nature of members of this taxon (127). Furthermore, the ANI analysis highlights that genomes of two *B. animalis* subsp. *lactis* strains, i.e. ATCC 27674 and CNCM I-2494, displayed a genetic identity of 99.9 % when compared with that of the prototypical probiotic bifidobacterial strain, i.e. *B. animalis* subsp. *lactis* BB-12 (189). Thus, we can speculate that the latter strains exhibit similar probiotic characteristics (190). Nevertheless,

additional functional genomics analyses coupled with *in vivo* studies should be performed in order to confirm this notion.

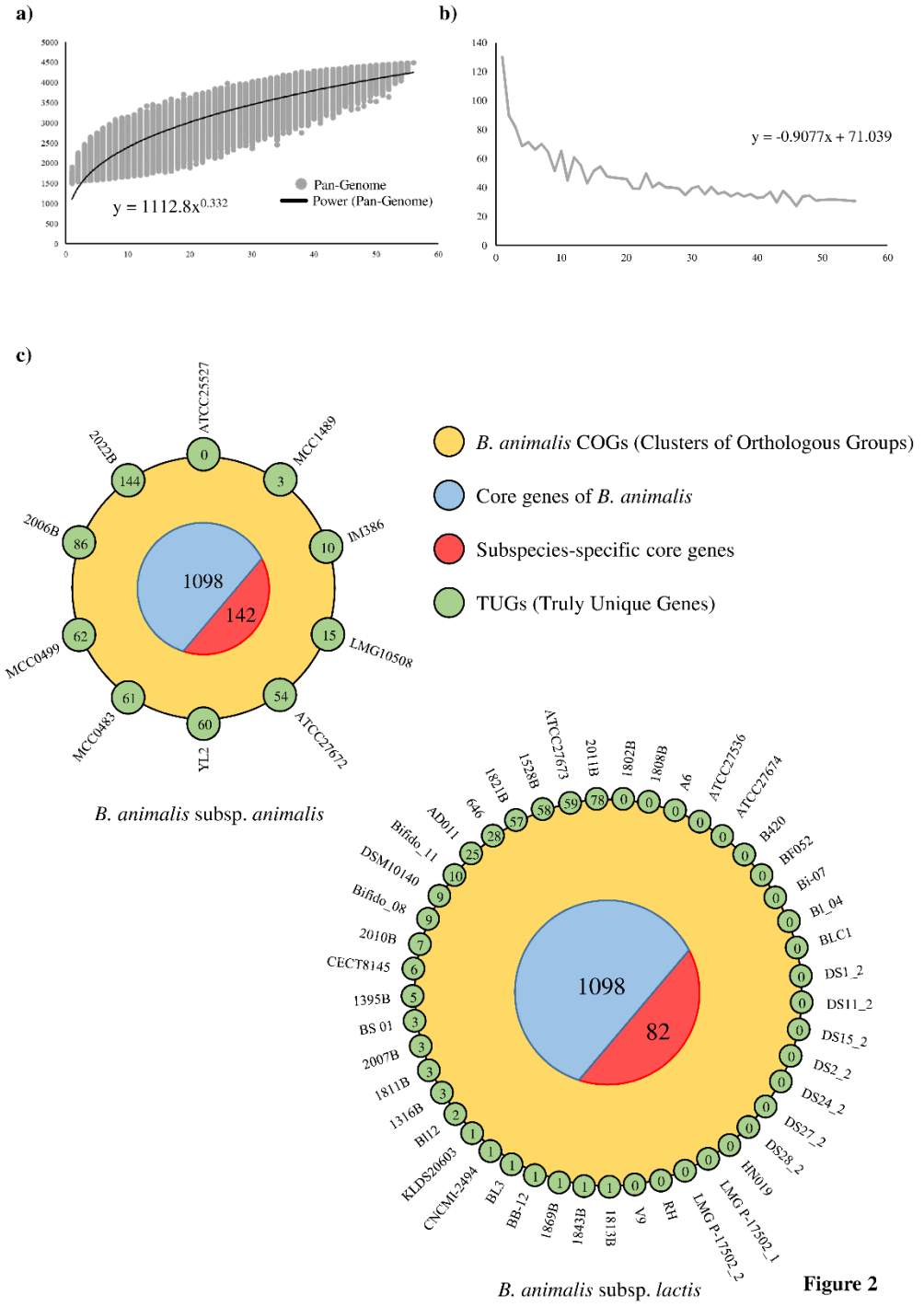


Figure 2

Figure 2. Pan-genome and core-genome of the *B. animalis* species. Panel a shows the pan-genome of the *B. animalis* species. Panel b exhibits the average of new genes upon sequential addition of the *B. animalis* genomes. Panel c displays two Venn diagrams representing shared orthologous as well as unique genes among the 56 *B. animalis* genomes. Numbers in blue circular segments represent the core genes of the *B. animalis* taxon, while numbers in red circular segments symbolize the subspecies-specific core genes. Moreover, numbers of unique genes are highlighted in small green circles.

Phylogenetic analyses of the *B. animalis* species.

Recently, a phylogenomic assessment of members of the genus *Bifidobacterium* allowed the identification of nine phylogenetic groups (67). Notably, *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* taxa are members of the *Bifidobacterium pseudolongum* group, which also includes *Bifidobacterium choerinum, Bifidobacterium cuniculi, Bifidobacterium gallicum, Bifidobacterium magnum*, *Bifidobacterium pseudolongum* subsp. *globosum* and *Bifidobacterium pseudolongum* subsp. *pseudolongum* (191). Accordingly, we re-evaluated the evolutionary development of the 56 *B. animalis* strains analyzed here using a phylogenomic approach, which also included the genome sequences of the bifidobacterial type-strains belonging to the *B. pseudolongum* phylogenetic group. *In silico* analyses identified 667 orthologous genes, which were shared among sequenced genomes of the *B. pseudolongum* group, which were then employed to build a so-called supertree (Fig. 3). This supertree showed that all 15 *B. animalis* strains isolated in this study, co-cluster with other publicly available *B. animalis* genomes. Furthermore, a clear division was identified between genomes belonging to the *B. animalis* subsp. *animalis* subspecies and those encompassing the *B. animalis* subsp. *lactis* subspecies (Fig. 3). As previously observed through molecular typing approaches, *B. animalis* subsp. *lactis* ATCC27672 clusters together with members of the *B. animalis* subsp. *animalis* group, suggesting a misclassification of this strain (181). Interestingly, *B. animalis* subsp. *lactis* 2011B clusters on a separate branch with respect to other *B. animalis* subsp. *lactis* strains, suggesting that this isolate may have followed a different evolutionary pathway compared to the other members of *B. animalis* subsp. *lactis* taxon. In order to assess the level of genetic differences between each *B. animalis* subspecies, we analyzed Single Nucleotide Polymorphisms (SNPs) among genomes of this taxon, using the software Mauve (192). The number of identified SNPs was higher in *B. animalis* subsp. *animalis* genomes (123,338 SNPs) as compared to those detected in the *B. animalis* subsp. *lactis* chromosomes (52,162 SNPs). In this context, 59.5 % of the *B. animalis* subsp. *animalis* SNPs were identified only in two strains, i.e., *B. animalis* subsp. *animalis* 2006B and *B. animalis* subsp. *animalis* 2022B, while 54.8 % of the *B. animalis* subsp. *lactis* SNPs were detected in only three

strains, i.e., *B. animalis* subsp. *lactis* 2010B, *B. animalis* subsp. *lactis* 2011B and *B. animalis* subsp. *lactis* 2007B. It should be noted that some of these differences may be correlated with the quality of the deposited genome sequences, which may have been affected by a low sequencing fold coverage. Nonetheless, strains that display the highest number of SNPs in their genomes also reflect their apparent phylogenetic distinctiveness in the supertree of the *B. pseudolongum* group (Fig. 3), perhaps reflecting divergent evolution when compared to other members of their subspecies. Furthermore, the performed phylogenetic analysis may assist in the selection of novel probiotic strains. In this context, 18 *B. animalis* subsp. *lactis* strains cluster in the BB-12 branch (Fig. 3). Their genomic relatedness was also highlighted in the pan-genome analysis, where half of the *B. animalis* subsp. *lactis* strains does not show any TUGs (Fig. 2).

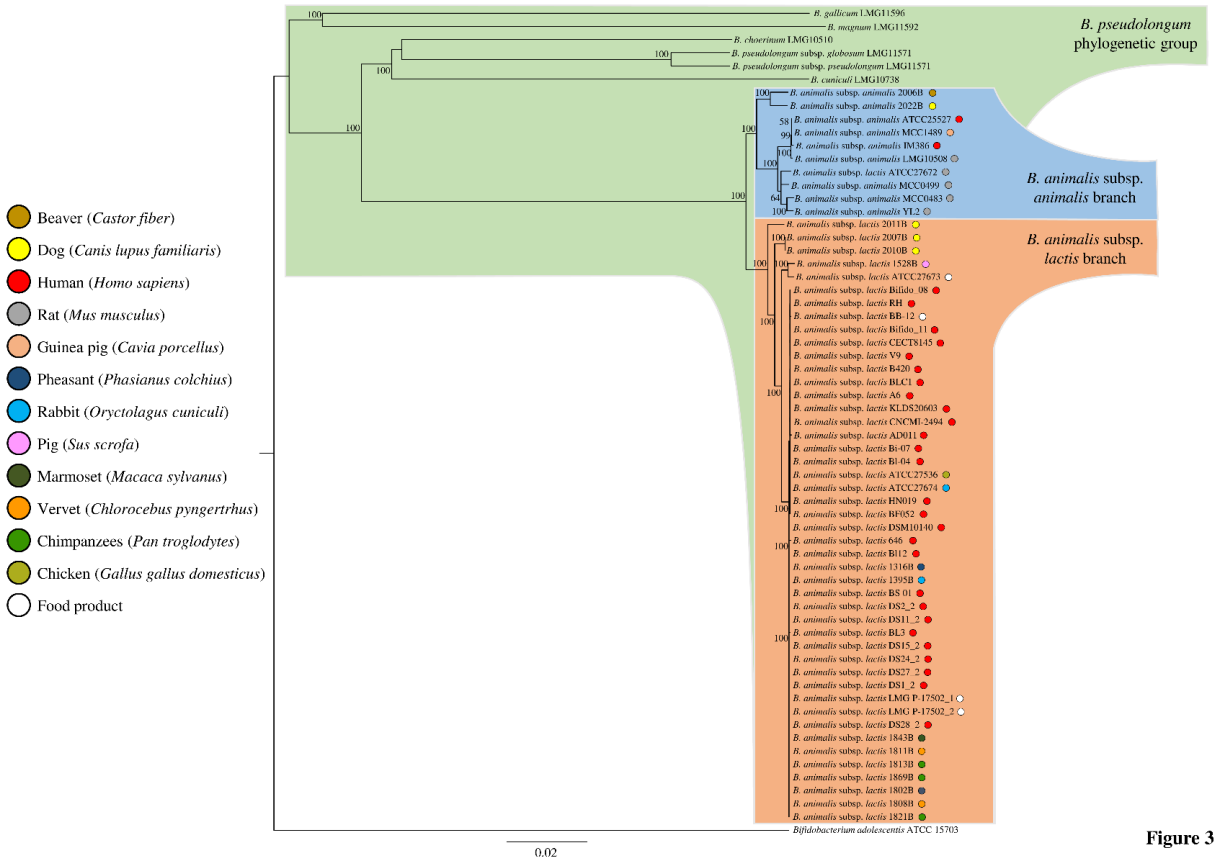


Figure 3

Figure 3. Phylogenomic tree of the *B. animalis* taxa. Proteomic tree based on the concatenation of 667 *B. animalis* core genes identified in the *B. pseudolongum* group phylogenomic analysis. This tree was constructed by the neighbor-joining method, and the genome sequence of *Bifidobacterium adolescentis* ATCC15703 was used as outgroup. Bootstrap percentages above 50 are shown at node points, based on 1,000 replicates. Colored small circles indicate the ecological origins of each bacteria.

The glycobiome of the *B. animalis* species.

Bifidobacteria are known to metabolize a wide range of carbohydrates, as a carbon and energy source, ranging from dietary- as well as host-derived glycans (86, 106, 117, 193). In order to assess carbohydrate fermentation capabilities of the two *B. animalis* subspecies, we performed growth experiments involving 19 *B. animalis* species cultivated on semi-synthetic medium with different carbohydrates as the sole carbon source. In order to obtain a complete overview of such carbohydrate metabolic abilities, we included both plant- and host- derived glycans (Fig. 4). As displayed in Figure 4, all *B. animalis* subsp. *lactis* strains were able to grow on a common set of sugars, such as lactose, maltose, raffinose and sucrose. In contrast, *B. animalis* subsp. *animalis* strains was shown to metabolize a broader array of sugars, with a high growth performance in media containing arabinose, galactose, glucose, maltose, melibiose, sucrose or xylose (194). Furthermore, *B. animalis* subsp. *lactis* 646, *B. animalis* subsp. *lactis* 1316B and *B. animalis* subsp. *lactis* 1395B, in contrast to other members of this subspecies exhibited appreciable growth on xylose (Fig. 4).

Statistical analyses were performed to corroborate the observed growth differences between *B. animalis* subsp. *lactis* and *B. animalis* subsp. *animalis* strains on different sugars. As shown in Figure 4, a significant growth difference (P-value < 0.05) for 14 carbohydrates was observed, with highest growth performances of *B. animalis* subsp. *animalis* strains (when compared to *B. animalis* subsp. *lactis* strains) in media containing arabinose, fructose, galactose, glucose, pullulan, trehalose or xylose (Table S1). On the other hand, *B. animalis* subsp. *lactis* strains were shown to grow significantly better (when compared to *B. animalis* subsp. *animalis* strains) in MRS medium supplemented with lactose (Fig. 4). Moreover, in five cases, the obtained growth performances were shown to be highly significantly different, with P-values < 0.001 (Fig. 4). Notably, none of *B. animalis* subsp. *lactis* strains was shown to be able to utilize mucin, N-acetyl-D-galactosamine and N-acetyl-D-glucosamine, which indicates that the tested strains possess limited metabolic capabilities concerning host-derived glycans (Fig. 4). In order to validate the observed metabolic differences of the *B. animalis* subspecies, we predicted the glycosyl hydrolase (GH) enzymes involved in

carbohydrate breakdown and belonging to the subspecies-specific core genes, as mentioned above. The *in silico* analyses were performed using the Carbohydrate-Active Enzymes (CAZy) database (109) involving the 56 *B. animalis* genomes mentioned above. Interestingly, 13 subspecies-specific core genes of *B. animalis* subsp. *animalis* genomes are predicted to be involved in sugar metabolism, while five genes indicated as carbohydrate-active enzymes are present in the subspecies-specific core genes of *B. animalis* subsp. *lactis* genomes. Among these subspecies-specific carbohydrate-active enzymes, we retrieved seven GH-encoding genes in *B. animalis* subsp. *animalis* genomes and four within *B. animalis* subsp. *lactis* strains. Interestingly, one of the seven *B. animalis* subsp. *animalis*-specific GH belongs to the GH2 family, which typically represent β-galactosidase (195) and exo-β-glucosaminidase (196) activities, confirming the observed high metabolic capabilities of this taxon towards galactose- and glucose-containing sugars (Figure 4). Furthermore, two *B. animalis* subsp. *animalis* GH-specific genes belong to the GH3 family, representing β-glucosidases and xylosidases (197) and the GH43 family, representing xylosidases (197) and arabinosidases (198), which are involved in the metabolism of xylose- and arabinose-containing glycans. Therefore, *in silico* analyses showed a larger number of GH-encoding genes among the *B. animalis* subsp. *animalis* subspecies-specific core genes when compared to the *B. animalis* subsp. *lactis* subspecies-specific core genes, confirming the observed broader carbohydrate-dependent growth performances displayed by this taxon.

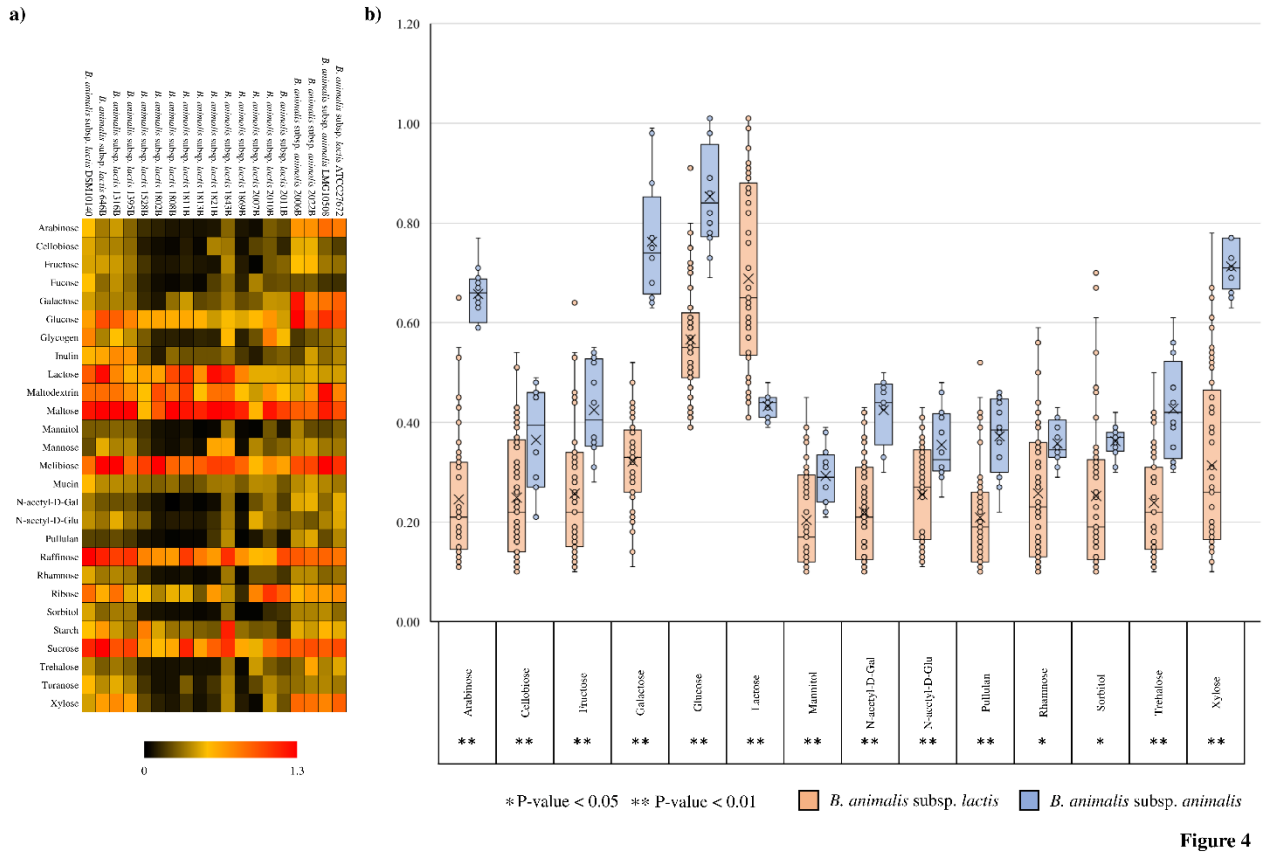


Figure 4

Figure 4. Evaluation of carbohydrate utilization by *B. animalis* strains. Panel a shows a heat map representing the growth performances of *B. animalis* strains on different sugars. Cultures were grown in biologically independent triplicates. Different shading represents the optical density reached by the assessed cultures. Panel b displays a whiskers plot based on optical density values of those sugars that results with a P-value < 0.05 between subspecies (Student's t-test). The *x-axis* represents the sole carbon source used for the growth experiments, while the *y-axis* shows the optical density values obtained for *B. animalis* subsp. *animalis* strains (blue color) and *B. animalis* subs. *lactis* strains (orange color). Points reflect the distribution of a data set, while the boxes represent 50 % of the data set, distributed between the 1st and 3rd quartiles. The Median divides the boxes into the interquartile range, while the X represents the Mean. The lines extending vertically outside of the boxes show the outlier range.

Evolutionary gain-gene and loss-gene analysis.

In order to identify genes that may have been acquired by Horizontal Genes Transfer (HGT), the genomes of the type strains of *B. animalis* subsp. *animalis*/*lactis* were analyzed with the software suite COLOMBO v 4.0 (176). Interestingly, 80 genes, representing 5.1 % of the *B. animalis* subsp. *lactis* genes, seem to have alien origins of which 42.5 % encode hypothetical proteins. Moreover, 7.5

% of the genes that may have been acquired by HGT are predicted to be enzymes involved in carbohydrate metabolism, while 12.6 % represent genes encoding transcriptional regulators, and genes involved in CRISPR-Cas systems (Fig. 5). In the case of *B. animalis* subsp. *animalis*, 4.6 % of the genes seem to have been acquired by HGT, of which 45.1 % represent hypothetical proteins. Moreover, 4.2 % of these genes encode transposase and 8.4 % are predicted to be involved in CRISPR-Cas and in transcriptional regulation. These data suggest that HGT events represent a minor force in the evolution of genomes of *B. animalis* species.

To evaluate the acquisition and loss of the subspecies-specific GH genes through the *B. pseudolongum* phylogenetic group, we analyzed the predicted subspecies-specific carbohydrate active enzymes using Count software (177), This evolutionary development analysis is based on the core-gene sequences retrieved from the type strains of the *B. pseudolongum* phylogenetic group. As indicated in Figure 5, the *B. animalis* subsp. *animalis* taxon seems to have acquired five carbohydrate active enzymes during evolution when compared to the common ancestor of the phylogenetic group. Furthermore, the *B. animalis* subsp. *lactis* taxon was shown to be the subspecies with the higher prevalence of subspecies-specific GH gene loss, encompassing five specific GHs (Fig. 5). These findings suggest that the *B. animalis* subspecies has followed a different evolutionary path, confirming our observed differences between these two taxa identified in the phylogenomic analyses.

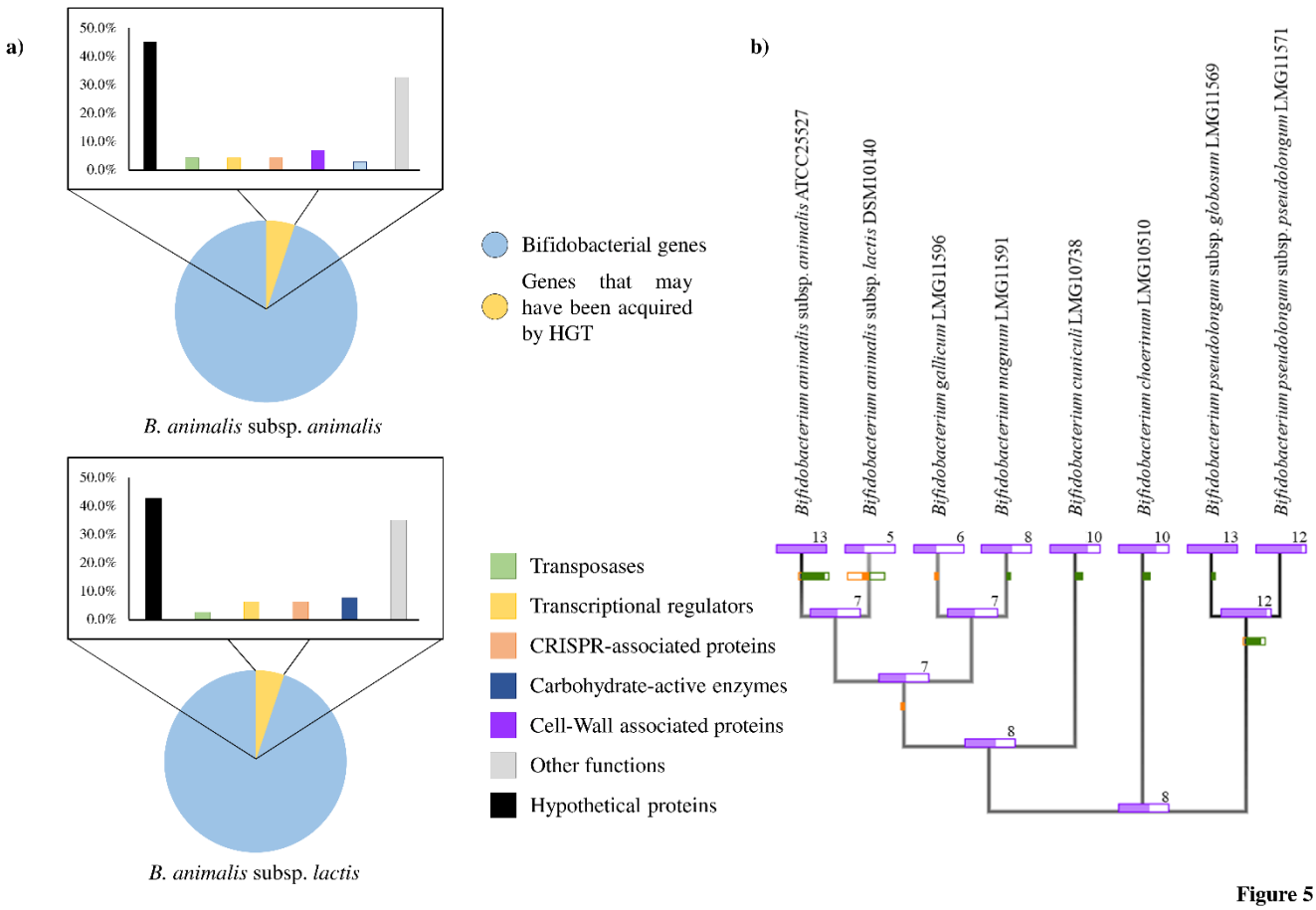


Figure 5

Figure 5. Evolutionary gene-gain and gene-loss analysis within the *B. pseudolongum* phylogenetic group as based on predicted subspecies-specific GHs. Panel a exhibits genes that are predicted to have been acquired by HGT events in the type strains of the *B. animalis* species. Bar-plots represent in different colors the functional annotations of the predicted genes. Panel b displays a tree based on the core genome of the *B. pseudolongum* phylogenetic group. The different sticks represent the predicted subspecies-specific GHs within *B. pseudolongum* phylogenetic group. Each node reports the number of predicted GHs identified in the type strains tested. Green and orange bars on the edge leading to each node indicate gain and losses.

CONCLUSIONS

Isolation of 15 *B. animalis* strains from the GIT of different animals, and representing the *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* taxa, revealed the cosmopolitan lifestyle of this species. Genome sequencing of the collected strains allowed us to reconstruct the genomic dataset of the *B. animalis* species, including 41 publicly available *B. animalis* genomes, unveiling that further genome sequencing of novel *B. animalis* strains will only slightly contribute to increase the pan-genome size. Nonetheless, phylogenetic analysis based on core genome sequences, among the 56 bifidobacterial genomes, showed a clear differentiation between the *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* branch. In fact, genome comparison of each strain showed the presence of a subspecies-specific core genome, representing the genetic differences between these two subspecies. Furthermore, the performed phylogenetic analysis highlights a cluster composed of 18 *B. animalis* subsp. *lactis* isolates that represent potential novel probiotic strains. Interestingly, a large proportion of the subspecies-specific genes of either *B. animalis* subspecies seems to be involved in sugar transport and metabolism. In this context, a larger number of such subspecies-specific transporter and GH activities was found in *B. animalis* subsp. *animalis* genomes. Growth performances on various sugars as their sole carbon source confirmed the ability of *B. animalis* subsp. *animalis* taxon to metabolize a broader set of sugars, e.g., arabinose, galactose, glucose, maltose, melibiose, sucrose and xylose, while *B. animalis* subsp. *lactis* strains seems to be more specialized using a smaller number of sugars, such as lactose, maltose, raffinose and sucrose. Altogether, these results seem to highlight a better ecological fitness of *B. animalis* subsp. *animalis* taxon compared to *B. animalis* subsp. *lactis* taxon. Moreover, a gene acquisition and loss analysis based on subspecies-specific glycosyl hydrolase genes revealed that *B. animalis* subsp. *animalis* taxon seems to have acquired several GHs through HGT, while *B. animalis* subsp. *lactis* species appears to have suffered loss of GH-encoding genes. Thus, these findings confirm the evolutionary differentiation between these two subspecies as highlighted in both phylogenetic and genomic analyses.

Table 1. *Bifidobacterium animalis* strain used in this study. The references were based on the decoding genomes project according to NCBI database.

Species	Strains	Ecological Origin	Genome Size (Mb)	No. of ORFs	GC content (%)	No of tRNA	rRNA locus	Coverage	Contigs	References
B. animalis subsp. *animalis*	2022B	*Castor fiber* feces	2.4	1935	61.08	65	2	142x	17	RSDC00000000
B. animalis subsp. *animalis*	2006B	*Canis lupus familiaris* (German shepherd) feces	2.16	1747	61.24	56	3	195x	47	RSDB00000000
B. animalis subsp. *animalis*	ATCC25527	Human feces	1.93	1622	61.35	52	2	-	-	(199)
B. animalis subsp. *animalis*	ATCC27672	Rat feaces	1.99	1611	60.97	52	1	-	-	NCBI database
B. animalis subsp. *animalis*	IM386	Human faeces	1.93	1623	61.35	52	1	-	-	NCBI database
B. animalis subsp. *animalis*	LMG10508	Rat feaces	1.92	1619	60.53	52	2	-	-	NCBI database
B. animalis subsp. *animalis*	MCC0483	Rat feaces	2.18	1922	60.97	53	1	-	-	NCBI database
B. animalis subsp. *animalis*	MCC0499	Rat feaces	2.13	1870	61.05	62	1	-	-	NCBI database
B. animalis subsp. *animalis*	MCC1489	Guinea pig feaces	1.91	1619	61.35	52	2	-	-	NCBI database
B. animalis subsp. *animalis*	YL2	Rat feces	2.02	1705	61.1	52	3	-	-	(200)
B. animalis subsp. *lactis*	646	Human feces	1.92	1673	61.4	52	4	-	-	NCBI database
B. animalis subsp. *lactis*	1316B	*Phasianus colchicus* feces	1.92	1556	60.47	52	3	195x	14	RSDA00000000
B. animalis subsp. *lactis*	1395B	*Oryctolagus cuniculus* feces	1.92	1557	60.47	52	2	99x	12	RSCZ00000000
B. animalis subsp. *lactis*	1528B	*Sus scrofa domesticus* feces	1.95	1600	61.47	55	2	97x	12	RSCY00000000
B. animalis subsp. *lactis*	1802B	*Macaca sylvanus* feces	1.92	1557	61.36	52	2	132x	15	RSCX00000000
B. animalis subsp. *lactis*	1808B	*Chlorocebus pygerytrhus* feces	1.92	1556	61.35	52	2	87x	15	RSCW00000000
B. animalis subsp. *lactis*	1811B	*Chlorocebus pygerytrhus* feces	1.68	1560	61.37	52	2	156x	16	RSCV00000000
B. animalis subsp. *lactis*	1813B	*Pan troglodytes* feces	1.68	1557	61.36	52	2	247x	12	RSCU00000000
B. animalis subsp. *lactis*	1821B	*Pan troglodytes* feces	1.75	1636	60.71	53	2	278x	40	RSCT00000000
B. animalis subsp. *lactis*	1843B	*Pan troglodytes* feces	1.68	1557	61.36	52	2	85x	14	RSCS00000000
B. animalis subsp. *lactis*	1869B	*Pan troglodytes* feces	1.92	1557	61.36	52	2	229x	14	RSCR00000000
B. animalis subsp. *lactis*	2007B	*Canis lupus familiaris* (Pomeranian) feces	1.97	1599	61.28	52	1	194x	25	RSCQ00000000

B. animalis subsp. *lactis*	2010B	*Canis lupus familiaris* (Alaskan malamute) feces	1.98	1607	61.21	52	2	121x	29	RSCP00000000
B. animalis subsp. *lactis*	2011B	*Canis lupus familiaris* (Flat coated retriever) feces	2.08	1700	61.32	54	1	181x	44	RSCO00000000
B. animalis subsp. *lactis*	A6	Human feces	1.96	1651	61.38	52	5	-	-	NCBI database
B. animalis subsp. *lactis*	AD011	Infant feaces	1.93	1642	61.38	52	2	-	-	(201)
B. animalis subsp. *lactis*	ATCC27536	Chicken feaces	1.91	1632	61.35	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	ATCC27673	Fermented milk sample	1.95	1685	61.52	52	3	-	-	(202)
B. animalis subsp. *lactis*	ATCC27674	Rabbit feaces	1.91	1629	61.35	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	B420	Human faeces	1.94	1633	61.37	52	3	-	-	(203)
B. animalis subsp. *lactis*	BB-12	Food matrices	1.97	1639	61.38	52	3	-	-	(204)
B. animalis subsp. *lactis*	BF052	Feces of breast-fed infant	1.94	1632	61.38	52	3	-	-	NCBI database
B. animalis subsp. *lactis*	Bi-07	Human faeces	1.94	1831	61.38	52	3	-	-	(203)
B. animalis subsp. *lactis*	Bifido_08	Human feces	1.95	1757	61.32	52	4	-	-	NCBI database
B. animalis subsp. *lactis*	Bifido_11	Human feces	1.94	1702	61.32	52	4	-	-	NCBI database
B. animalis subsp. *lactis*	Bl_04	Human feces	1.94	1633	61.38	52	3	-	-	(184)
B. animalis subsp. *lactis*	Bl12	Human colonoscopic sample	1.94	1633	61.37	52	3	-	-	NCBI database
B. animalis subsp. *lactis*	BL3	Human feces	1.94	1639	61.38	52	3	-	-	(205)
B. animalis subsp. *lactis*	BLC1	Human faeces	1.94	1630	61.37	52	3	-	-	(206)
B. animalis subsp. *lactis*	BS01	Human faeces	1.93	1632	61.37	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	CECT8145	Infant faeces	1.96	1766	61.38	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	CNCMI-2494	Human faeces	1.94	1635	61.38	52	3	-	-	(207)
B. animalis subsp. *lactis*	DS1_2	Human feces	1.92	1636	61.36	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	DS11_2	Human feces	1.92	1637	61.36	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	DS15_2	Human feces	1.92	1635	61.37	52	3	-	-	NCBI database
B. animalis subsp. *lactis*	DS2_2	Human feces	1.92	1634	61.35	52	2	-	-	NCBI database

B. animalis subsp. *lactis*	DS24_2	Human feces	1.92	1670	61.35	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	DS27_2	Human feces	1.92	1642	61.35	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	DS28_2	Human feces	1.92	1633	61.35	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	DSM10140	Human feces	1.94	1635	61.37	51	3	-	-	(184)
B. animalis subsp. *lactis*	HN019	Human faeces	1.92	1645	61.35	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	KLDS2.0603	Human feces	1.95	1646	61.37	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	LMG P-17502_1	Food sample	1.92	1628	61.36	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	LMG P-17502_2	Food sample	1.92	1628	61.36	52	1	-	-	NCBI database
B. animalis subsp. *lactis*	RH	Human feces	1.93	1629	61.37	52	2	-	-	NCBI database
B. animalis subsp. *lactis*	V9	Human feces	1.94	1633	61.38	52	3	-	-	NCBI database

Chapter 4

Bifidobacterial transfer from mother to child as examined by an animal model

Mancino W, Duranti S, Mancabelli L, Longhi G, Anzalone R, Milani C, Lugli GA, Carnevali L, Statello R, Sgoifo A, van Sinderen D, Ventura M, Turroni F.

The results of this chapter were published on Microorganisms, 2019 Aug 27;7(9):293. doi: 10.3390/microorganisms7090293.

ABSTRACT

Bifidobacteria commonly constitute the most abundant group of microorganisms in the healthy infant gut. Their intestinal establishment is believed to be maternally driven, and their acquisition has even been postulated to occur during pregnancy. In the current study, we evaluated bifidobacterial mother-to infant transmission events in a rat model by means of quantitative PCR (qPCR), as well as by Internally Transcribed Spacer (ITS) bifidobacterial profiling. The occurrence of strains supplied by mothers during pregnancy to their corresponding newborns was observed and identified by analysis immediately following C-section delivery. These findings provide intriguing support for the existence of an unknown route to facilitate bifidobacterial transfer during the very early stages of life.

For Supplemental Materials see the article published on *Microorganisms*.

MATERIALS AND METHODS

Experiment design and bifidobacterial treatment of rats.

Experiments were performed in accordance with the European Community Council Directive 2010/63/UE and approved by the Italian legislation on animal experimentation (D.L. 04/04/2014, n. 26, authorization no. 370/2018-PR). All efforts were made to reduce sample size and minimize animal suffering. For the purposes of the current study we employed adult (4-6 months) Wistar rats (*Rattus norvegicus*) which were bred in house under standard conditions (208). The timeline of animal treatment procedures is schematically depicted in supplemental Figure S1. The duration of the animal trial was four weeks in total. Mating was allowed during week 1 and for this purpose, 12 adult female rats (*Rattus norvegicus*) were coupled with 12 male rats. Pairs were kept in different cages in rooms with controlled temperature (22 ± 2 °C) and humidity (60 ± 10 %) and maintained in a 12/12 light/dark cycle (light on from 19:00 to 7:00 h), with food and water *ad libitum*. Following the one week mating period, all female rats were separated and kept in individual cages and were every day orally treated with bifidobacterial strains for three weeks, which represented the gestation period. One group of three female rats (which are here referred to as W1, W2 and W3) was orally inoculated with a mix of three different strains, i.e., *Bifidobacterium bifidum* PRL2010, *Bifidobacterium breve* 1895B and *Bifidobacterium longum* subsp. *longum* 1886B (Mix Colonization Group – MCG). The second group with the remaining nine female rats (W4, W5, W6, W7, W8, W9, W10, W11 and W12) was orally treated with *B. bifidum* PRL2010 only (PRL2010 Group – PG). All the strains used in this study were previously isolated from infant stool samples (80, 148). In order to evaluate bifidobacterial transfer, fecal samples were collected at five different time points. The first time point was before the oral administration of bifidobacteria (T0). Then, we collected fecal samples at four time points, i.e. at five, eight, 12 and 17 days (T1, T2, T3 and T4) (Fig. S1). On the 19th day of bifidobacterial treatment (i.e., on the day of birth, but

prior to labor) female animals were anesthetized with isoflurane (2 % in 100 % oxygen) and Caesarian delivery was performed under a laminar flow hood and all the surgical instruments used were previously sterilized in order to prevent any microbial contamination. Placentas were harvested and caecum and blood samples were collected from dams and pups.

***Bifidobacterium* strains growth conditions.**

All strains used in this study were cultivated in an anaerobic atmosphere (2.99 % H_2, 17.01 % CO_2 and 80 % N_2) in an anaerobic chamber (Concept 400, Ruskin) on De Man-Rogosa-Sharp (MRS) broth (Scharlau Chemie, Barcelona, Spain) supplemented with 0.05 % (w/v) L-cysteine hydrochloride (Sigma-Aldrich) and incubated at 37º C. Microbial cultures were harvested by centrifugation (3000 rpm for 8 minutes), washed and resuspended in 500 µL of 2 % (w/v) sucrose solution. The viable count of each inoculum was determined by retrospective plating on MRS.

DNA extraction and qPCR.

Bacterial DNA extraction from fecal samples was performed following the manufacturer's protocol of the QIAamp Fast DNA stool Mini Kit (Qiagen Ltd, Strasse, Germany). Furthermore, bifidobacterial DNA presence was evaluated for the mother's caecum, placentas, blood samples and from puppies' caecal samples. Specifically, the bacterial DNA from puppies and mothers' caecum and from placentas was extracted following the protocol of Power Viral Environmental RNA/DNA Isolation Kit (Mo Bio, Germany), whereas the bacterial DNA from mother's blood was extracted following the protocol of the DNeasy Blood and Tissue Kit (Qiagen Ltd, Strasse, Germany).

Quantitative PCR (qPCR) was performed as described previously (209). For *B. bifidum* PRL2010 were used primers Bbif_0282Fw (5'-GCGAACAATGATGGCACCTA-3') and

Bbif_0282Rv (5'-GTCGAACACCACGACGATGT-3') (128), for *B. breve* 1895B were used primers BBR7E_0534_fw (5'-AGCGACGATATGATGCAATG-3') and BBR7E_0534_rev (5'-CGTGAATACGCTGCACAGTC-3') and for *B. longum* subsp. *longum* 1886B were used primers B1886_0443_fw (5'-AAGCCAAGGACATGTTCGAC-3') and B1886_0443_rev (5'-TGGTGTATCTGGCGTTCTTG-3') (148). For species specific qPCR were used following primer pairs: Bbif1 (5'-CCACATGATCGCATGTGATTG-3') and Bbif2 (5'-CCGAAGGCTTGCTCCCAAA-3') for *B. bifidum* species (193), Bbre1 (5'-CCGGATGCTCCATCACAC-3') and Bbre2 (5'-ACAAAGTGCCTTGCTCCCT-3') for *B.breve* species (210), Blon1 (5'-TTCCAGTTGATCGCATGGTC-3') and Blon2 (5'-GGGAAGCCGTATCTCTACGA-3') for *B. longum*.

Bifidobacterium strain isolation from mothers' caecum.

The collected caecum of the mothers was homogenized and serial dilutions (1:10) were performed. All dilutions were cultivated on MRS agar (Scharlau Chemie, Barcelona, Spain) supplemented with 0.05 % (w/v) L-cysteine hydrochloride (Sigma-Aldrich) and 50 μg/ml mupirocin (Delchimica, Italy). After 48h of incubation in an anaerobic atmosphere (2.99 % H_2, 17.01 % CO_2 and 80 % N_2) in a chamber (Concept 400, Ruskin), morphologically distinct colonies were selected and cultivated in MRS broth (Scharlau Chemie, Barcelona, Spain) supplemented with 0.05 % (w/v) L-cysteine hydrochloride (Sigma-Aldrich). Subsequently, after overnight growth, bacterial DNA was extracted as described previously (211). The identification of specific strains was obtained using a PCR approach based on strain-specific primers (see above).

Bifidobacterial ITS PCR amplification and sequencing.

Following bacterial DNA extraction from caecal samples of mothers and puppies, partial ITS sequences were amplified using primer pair Probio-bif_Uni/Probio-bif_Rev, which targets the spacer region between the 16S rRNA and 23S rRNA genes within the ribosomal RNA (rRNA) locus (165). At the same time, we prepared a mock community (Mock Bifidobacterial Community), consisting of a pool of known concentrations of 11 different *Bifidobacterium* strains prepared by combining equal concentration of bacterial DNA. The DNA from the mix was diluted to produce a final DNA concentration of 2 ng/µL, and 4 µL of these dilutions were used in each PCR reaction using primer pair Probio-bif_Uni/Probio-bif_Rev. Illumina adapter overhang nucleotide sequences were added to the partial ITS amplicons, which were further processed employing the 16S Metagenomic Sequencing Library Preparation Protocol (Part #15044223 Rev. B – Illumina). The library preparation was performed as described above for the 16S rRNA microbial profiling analyses. Following sequencing, the .fastq files were processed using a custom script based on the QIIME software suite (212). In order to reconstruct the complete Probio-bif_Uni / Probio-bif_Rev amplicons, the paired-end read pairs were assembled. Quality control retained sequences with a length between 100 and 400 bp and mean sequence quality score of >20, while sequences with homopolymers >7 bp in length and mismatched primers were removed. ITS Operational Taxonomic Units (OTUs) were defined at 100 % sequence homology using uclust software (213). All reads were classified to the lowest possible taxonomic rank using QIIME2 (212, 214) and a reference dataset, consisting of an updated version of the bifidobacterial ITS database (165).

RESULTS AND DISCUSSION

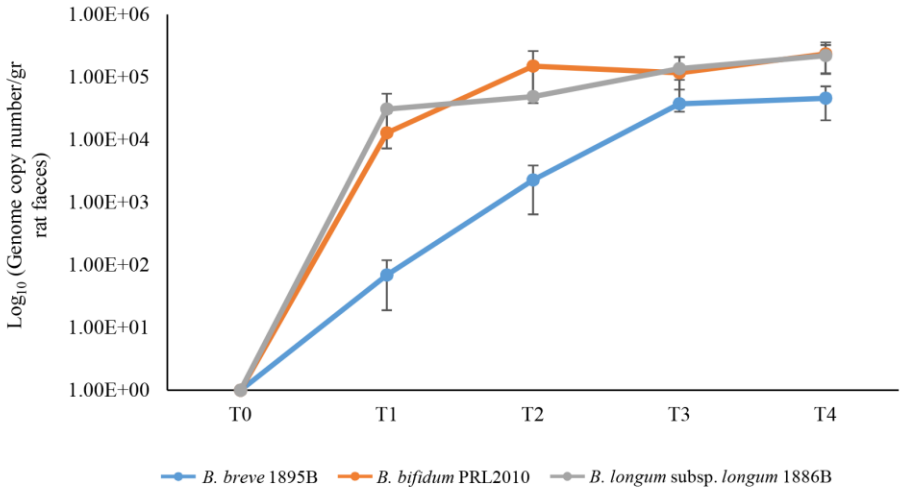
Evaluation of vertical transmission of bifidobacteria under in vivo conditions.

As previously shown by metagenomic attempts, identical bifidobacterial strains have been found in the fecal samples of mother-newborn dyads (151, 215). In order to evaluate possible microbial transfer by pregnant rats treated with a bifidobacterial strain mix consisting of *B. bifidum* PRL2010, *B. breve* 1895B and *B. longum* subsp. *longum* 1886 strains, or rats treated with *B. bifidum* strain PRL2010 only, animals were divided in two different groups. The first group consisted of three female rats, which were treated with the bifidobacterial mix (Mix Colonization Group – MCG), whereas the second group encompassed nine female rats, to which only PRL2010 was administered (PRL2010 Group - PG). Notably, the bifidobacterial strains used were, previously, isolated from fecal samples from healthy breast-fed infants (75, 80, 148). However, even if these bacterial strains were not belonging to the indigenous microbiota of rat, when administered to these animals during pregnancy, they are able to colonize the gut of rat (see below). Jimenez et al. have described a similar result previously, where a human infant gut commensal, *Enterococcus faecium*, could be vertically transmitted in pregnant mice.

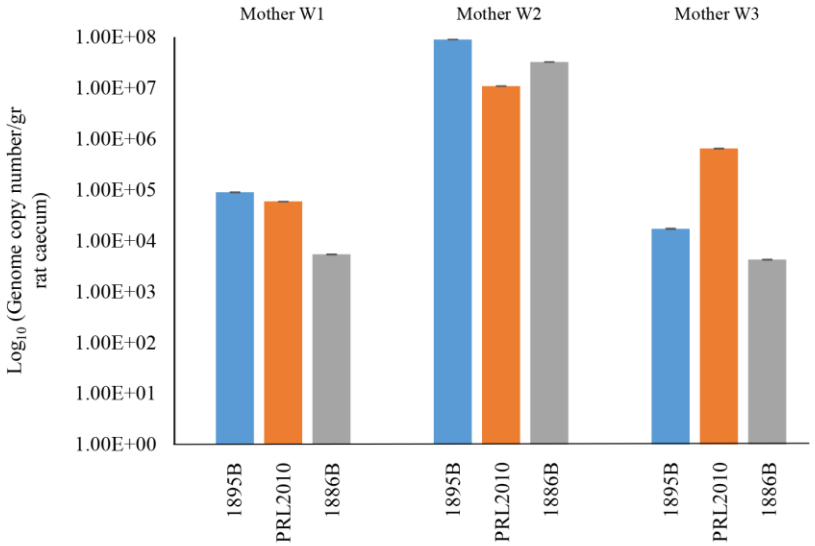
At the start of the experiment, rats were checked for the presence of strains PRL2010, 1895B and/or 1886B in fecal samples by means of PCR using strain-specific primer pairs, revealing that, as expected, these bifidobacterial strains were absent in the animals enrolled in this study and prior to them being fed any of the strains. Dams were administered a daily dose of 10^9 colony forming units (CFU) of *B. bifidum* PRL2010 or a mix of *B. bifidum* PRL2010, *B. longum* subsp. *longum* 1886B and *B. breve* 1895B for 21 days. Presence and clearance of *B. bifidum* PRL2010, *B. longum* subsp. *longum* 1886B and *B. breve* 1895B were monitored during the gestation period using a qPCR approach based on strain-specific primers (Fig. 1). Following Caesarian delivery of the pups, rats were sacrificed and their caecum was removed and assayed

for the presence of PRL2010, 1886B and 1895B by means of qPCR. Interestingly, the microbial density estimated by qPCR of *B. bifidum* PRL2010 in the caecal samples of mothers ranged from 10^4 to 10^7 CFU, whereas in caecal samples of pups the abundance ranged from 10^3 to 10^6 CFU (Fig. 1). Similarly, the qPCR estimated abundance of *B. breve* 1895B in the caecum of mothers and of newborns ranged from 10^4 to 10^7 and 10^2 to 10^6, respectively (Fig. 1). The cell load determined by qPCR of *B. longum* subsp. *longum* 1886B in the mother's gut and in pup's caecum varies from 10^3 to 10^7 and 10^2 to 10^4, respectively (Fig. 1).

a)


1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Log10 (Genome copy number/gr rat faeces)
T0
T1
T2
T3
T4
B. breve 1895B
B. bifidum PRL2010
B. longum subsp. longum 1886B

b)


Mother W1
Mother W2
Mother W3
1.00E+08
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Log10 (Genome copy number/gr rat caecum)
1895B
PRL2010
1886B

c)

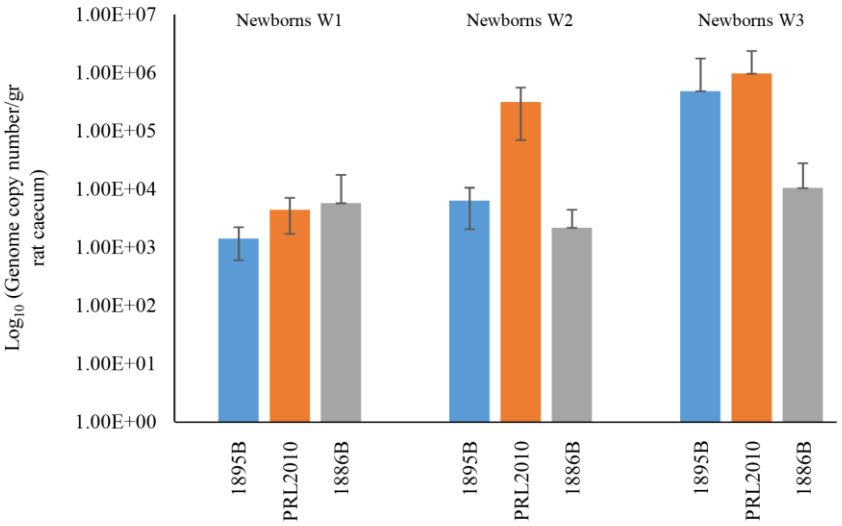

Newborns W1
Newborns W2
Newborns W3
1.00E+07
1.00E+06
1.00E+05
1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Log10 (Genome copy number/gr rat caecum)
1895B
PRL2010
1886B

Figure 1

Figure 1. Schematic representation of the vertical transmission of the bifidobacterial mix in rat model treated with B. breve 1895B, B. bifidum PRL2010 and B. longum subsp. longum 1886B strains (Mix Colonization Group – MCG). Panel a shows the average of DNA presence of the three strains observed during the bifidobacterial administration. Each point represents the average of the log-population size ± standard deviation for three rats. Panel b displays the presence of *B. breve* 1895B, *B. bifidum* PRL2010 and *B. longum* subsp. *longum* 1886B in the caecum of female rats. Panel c exhibits bifidobacterial retrieval from puppies' caecum. Each pillar represents the average presence for each *Bifidobacterium* strain ± standard deviation.

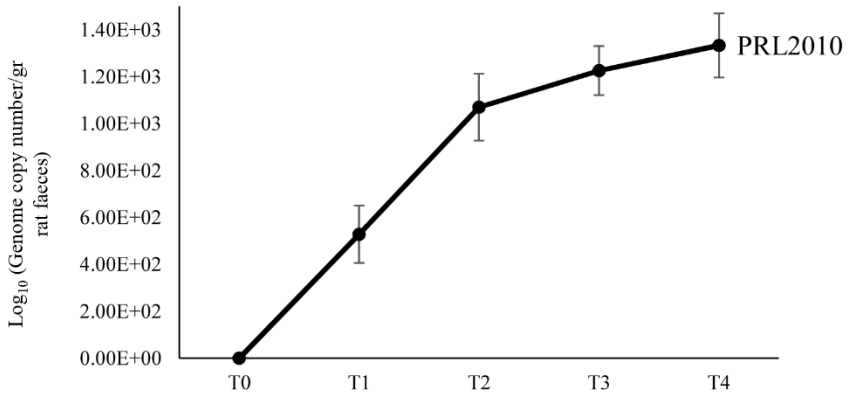
In addition, we isolated viable PRL2010/1895B/1897B cells from the caecal samples of the mothers by means of direct cultivation of caecal contents of animals on mMRS followed by the precise strain-assignment of the isolated cells using a PCR approach based on strain-specific primers.

Maternal inheritance of *B. bifidum* PRL2010 strain.

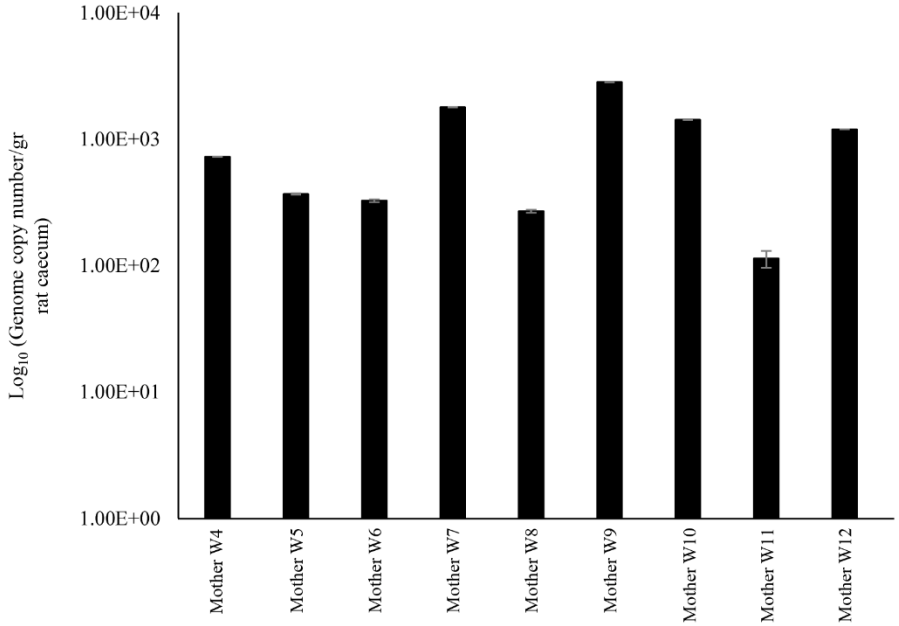
Since *B. bifidum* strain PRL2010 has been shown to represent a model infant gut commensal (80-82, 128, 178, 193, 209, 215-224), and as described above for the Mix Colonization Group – MCG , this strain displayed the highest level of vertical transmission from mother to newborns, we decided to further examine the maternal inheritance of strain PRL2010. Bifidobacterial inheritance of nine pregnant rats receiving a daily dose of 10^9 CFU of *B. bifidum* PRL2010 for three weeks was monitored by qPCR (Fig. 2). These rats delivered their puppies by Caesarian section and PRL2010 levels were evaluated by qPCR using strain-specific primers targeting both the caecal sample of the mother and the corresponding pups (Fig. 2). The estimated abundance of PRL2010 cells by qPCR ranged from 10^3 to 10^4 CFU and from 10^2 to 10^4 CFU in mother and newborns samples, respectively (Fig. 2). Notably, the PRL2010 colonization was lower in PG mothers respect to MCG mothers (Pvalue < 0.001). Concurrently,

the genome copy number of PRL2010 in the newborns from PG group was lower respect those observed in MCG newborns group (Pvalue < 0.001). These findings suggest that microbe-microbe interactions provide an advantage in the vertical transmission efficiency of these species in the infant gut. In addition, such results corroborate previous data regarding the existence of syntrophic interactions between bifidobacteria in the infant gut (75, 178, 215). However, the bacterial load of PRL2010 in the MCG newborns group resulted lower respect to their mothers. In contrast, the PG newborns group showed a higher abundance of PRL2010 cells when compared to their mothers (Fig. S2). Probably, the load of *B. bifidum* PRL2010 cells in the newborns of PG group is higher respect to their mothers because it has been administered as a mono-strain supplement. Conversely, the simultaneously supplementation of bifidobacterial mix encompassing different strains in MCG group contributed to the decreased level of PRL2010 cells in the newborns.

a)


1.40E+03
1.20E+03
1.00E+03
8.00E+02
6.00E+02
4.00E+02
2.00E+02
0.00E+00
Log10 (Genome copy number/gr rat faeces)
PRL2010
T0
T1
T2
T3
T4

b)


1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Log10 (Genome copy number/gr rat caecum)
Mother W4
Mother W5
Mother W6
Mother W7
Mother W8
Mother W9
Mother W10
Mother W11
Mother W12

c)

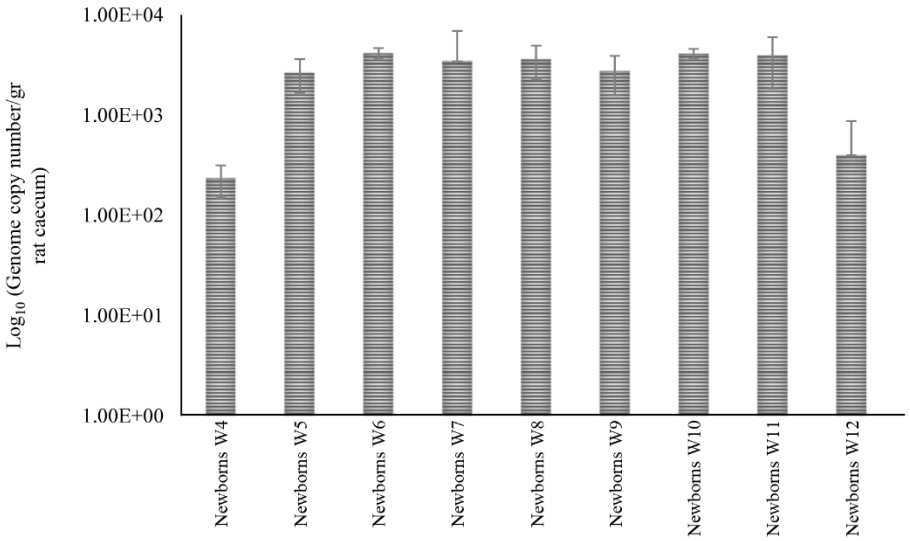

1.00E+04
1.00E+03
1.00E+02
1.00E+01
1.00E+00
Log10 (Genome copy number/gr rat caecum)
Newborns W4
Newborns W5
Newborns W6
Newborns W7
Newborns W8
Newborns W9
Newborns W10
Newborns W11
Newborns W12

Figure 2

Figure 2. Schematic representation of vertical inheritance of *B. bifidum* strain PRL2010, when administered solely in a rat model (PRL2010 Group – PG). Panel a shows the population sizes of *B. bifidum* PRL2010 being present in the intestine of female rats. Each point represents the average of the log-population size ± standard deviation for nine rats. Panel b displays the *B. bifidum* PRL2010 retrieval ± standard deviation for each female rat used. Panel c exhibits the presence of *B. bifidum* PRL2010 in the caecum from puppies. Each pillar represents the average colonization ± standard deviation of puppies for each female rats.

Identification of DNA belonging to PRL2010 in different rat body sites.

Bifidobacterial DNA occurrence in other body sites of rats such as placenta and blood samples, which was collected from mothers immediately after Caesarian delivery, was investigated by qPCR using PRL2010-specific primers. Remarkably, these experiments revealed the presence of DNA belonging to PRL2010 in the placental tissue but not in blood (Fig. 3). Similar results were described previously by Jimenez et al. in a trial examining the vertical transmission of *Enterococcus faecium* HA1 in pregnant mice (225). Any attempts directed to cultivate PRL2010 cells from placenta samples under mMRS did not result in the isolation of viable cells. These findings suggest that either the rat placenta can only be reached by DNA from lysed PRL2010 cells or that this body compartment only contains dormant PRL2010 cells. These data may also open a novel and intriguing scenario of prenatal colonization of rats by *B. bifidum* PRL2010.

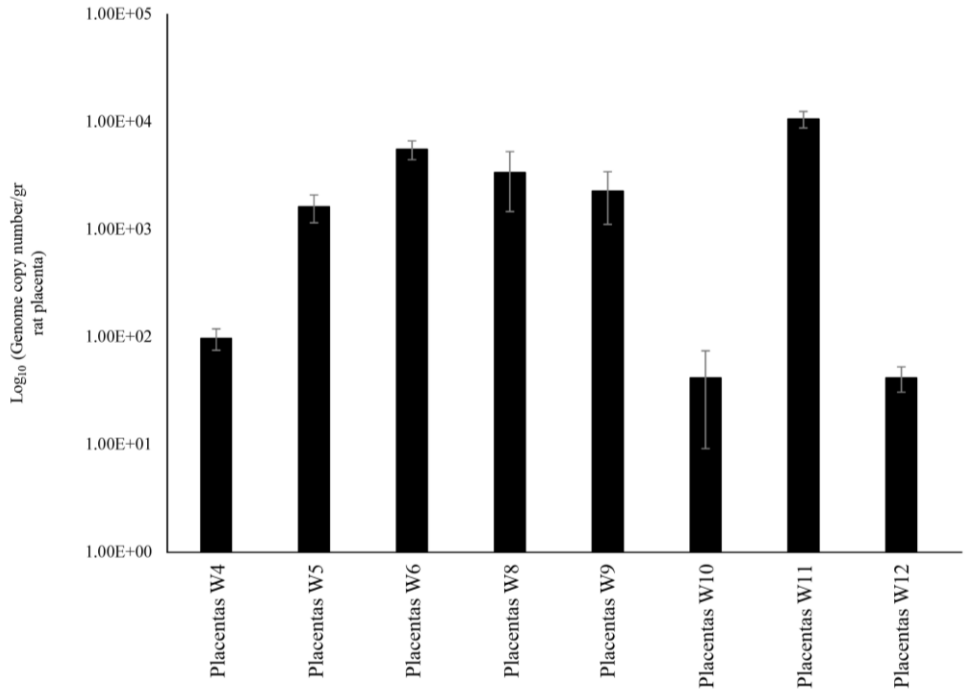


Figure 3. Schematic representation of the *B. bifidum* PRL2010 DNA load observed in placentas' samples. This graphic displays the level of DNA belonging to *B. bifidum* PRL2010 present in the placenta of rats. Each pillar represents the average retrieval ± standard deviation of *B. bifidum* PRL2010 from the placenta of rats.

ITS bifidobacterial profiling of the caecum of mothers and newborns.

In order to further characterize the bifidobacterial composition of caecal samples from mothers and puppies, ITS bifidobacterial profiling analyses were performed (165) on these 30 samples, producing a total of 212,034 reads with an average of 7,068 ± 6457 reads per sample (Table 1). These raw data were processed to identify and classify reads into clusters of identical sequences (OTUs). In detail, we focused our interest on OTUs belonging to the species *B. bifidum*, *B. breve* and *B. longum* subsp. *longum*. Interestingly, all puppies of the MCG group shared these three OTUs with the corresponding mothers' samples (Fig. 4). Moreover, comparison between PG mothers with their corresponding puppies showed that in 70 % of

cases the OTUs belonging to the species *B. bifidum*, *B. breve* and *B*. *longum* subsp. *longum* are shared (Fig. 4). In order to confirm the presence of *B. bifidum*, *B. breve* and *B*. *longum* subsp. *longum* in PG group, we performed a qPCR using species-specific primers using mothers' caeca (Fig. 4). These data revealed that *B. breve* and *B*. *longum* subsp. *longum* species were present in the gut microbiota of pregnant rats, even if they were not supplemented to the animals (Fig.4).

Table 1. Filtering table of the analyzed caecal samples.

		Input reads	Final reads
PG	**W4-mother**	3423	3126
PG	**F4-1**	2589	2386
PG	**F4-2**	8551	8012
PG	**F4-3**	4126	3776
PG	**W6-mother**	2078	1970
PG	**F6-1**	2164	2036
PG	**F6-2**	5050	4852
PG	**W8-mother**	1239	1201
PG	**F8-1**	6365	5797
PG	**F8-2**	4527	4050
PG	**W11-mother**	1009	942
PG	**F11-1**	4106	3908
PG	**F11-2**	2133	2036
PG	**F11-3**	1692	1642
MCG	**W2-mother**	22528	22026
MCG	**F2-1**	12703	11906
MCG	**F2-2**	32065	29485
MCG	**F2-3**	11715	10850
MCG	**F2-4**	17106	15929
MCG	**F2-5**	13992	12724
MCG	**F2-6**	10339	9739
MCG	**W3-mother**	2686	2640
MCG	**F3-1**	4818	4620

MCG	**F3-2**	8970	8489
MCG	**F3-3**	5346	4987
MCG	**F3-4**	5664	5373
MCG	**F3-5**	2765	2584
MCG	**F3-6**	12523	11182
MCG	**F3-7**	6129	5701
MCG	**F3-8**	8524	8065

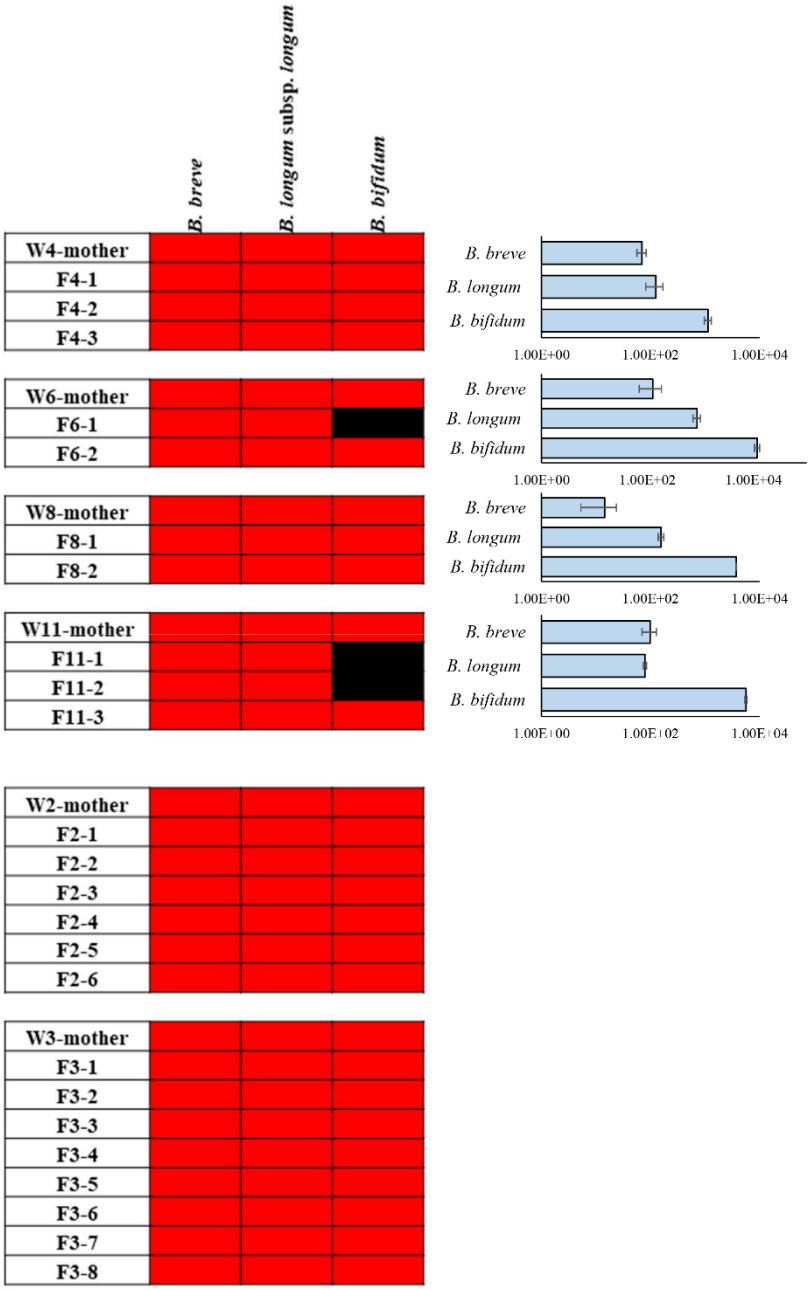


Figure 4. Heat map of bifidobacterial OTUs belonging to *B. bifidum*, *B. breve* and *B. longum* subsp. *longum*. Red shading represents presence, and black shading indicates absence. The pillars next to each heat map display the DNA load of these bifidobacterial species in mothers' caeca of PG group. The x axes represent the genome copy number /gr of cecum samples

These results further support the notion of transfer of (bifido)bacterial DNA and/or cells between mother and puppies (75, 163) and reinforced the earlier finding that co-existence of multiple bifidobacterial species allows a more efficient transmission of (bifido)bacterial DNA and/or cells from mothers to their newborns (75, 178, 215).

CONCLUSIONS

This study was aimed to investigate the DNA bifidobacterial transmission from female rats to their corresponding pups through a vertical route. Notably, the efficiency/yield of these DNA transmission events may be influenced by the co-presence of different bifidobacterial strains. The above results confirm previous genus-level overviews (163) and underpin the notion that bifidobacterial transmission is characterized by the development of extensive microbe-microbe co-operation (75, 151, 178). In fact, our data suggest that bifidobacterial communities have established co-operative behavior between co-colonizers, acting as evolutionary drivers in the mammalian gut microbiota. This assumption is further reinforced by previous studies that have described the occurrence of cross-feeding between specific bifidobacterial species (75, 151). In addition, our findings have provided preliminary insights about the existence of possible bifidobacterial colonization at preterm level. Nevertheless, the lack of any data about the viability of PRL2010 cells cannot fully support the existence of fetal colonization of bifidobacteria and further experiments will be needed in order to investigate this intriguing route of bacterial colonization.

Chapter 5

Mobilome and resistome reconstruction from genomes belonging to members of the *Bifidobacterium* genus

Mancino W, Lugli GA, van Sinderen D, Ventura M, Turroni F.

The results of this chapter were published on Microorganisms, 2019 Dec 2;7(12):638. doi: 10.3390/microorganisms7120638.

ABSTRACT

Specific members of the genus *Bifidobacterium* are among the first colonizers of the human/animal gut, where they act as important intestinal commensals associated with host health. As part of the gut microbiota, bifidobacteria may be exposed to antibiotics, used in particular for intrapartum prophylaxis, especially to prevent *Streptococcus* infections, or in the very early stages of life after the birth. In the current study, we reconstructed the *in silico* resistome of the *Bifidobacterium* genus, analyzing a database composed of 625 bifidobacterial genomes, including partial assembled strains with less than 100 genomic sequences. Furthermore, we screened bifidobacterial genomes for mobile genetic elements such as transposases and prophage-like elements, in order to investigate the correlation between the bifido-mobilome and the bifido-resistome, also identifying genetic insertion hotspots that appear to be prone to Horizontal Gene Transfer (HGT) events. These insertion hotspots were shown to be widely distributed among analyzed bifidobacterial genomes, and suggest acquisition of antibiotic resistance genes through HGT events. These data were further corroborated by growth experiments directed to evaluate bacitracin A resistance in *Bifidobacterium* spp., a property that was predicted by *in silico* analyses to be part of the HGT-acquired resistome.

For Supplemental Materials see the article published on *Microorganisms*.

MATERIALS AND METHODS

Bacterial Strains.

We retrieved the genome sequence of 625 public available *Bifidobacterium* genomes from the National Center for Biotechnology Information (NCBI) public database (Table S1). Collected genomes with more than 100 genomic sequences were discarded to analyze high quality genome sequences only. As reported in Table S1, strains, which were not classified at species level, were validated using the average nucleotide identity (ANI) approach. Strains used for this analysis were compared with the 67 type strains of the *Bifidobacterium* genus. Notably, two bifidobacterial strains displaying an ANI value of < 94 % may be considered to belong to two different species (29222102 25107967 28764658 19855009).

IS elements identification.

Predicted genes of 625 bifidobacterial strains used for this study were screened for the presence of IS elements. We used a custom database composed by 329,372 RefSeq sequences belonging to the Actinobacteria phylum, retrieved from the NCBI database. The alignment was performed through BLASTP analysis with an E-value cutoff of $1e^{-5}$ (171). After the manual control of the sequences with an amino acid length less than 100 amino acid of the type strains of the species studied, we decided to discard these sequences because they were considered non-functional or truncated. Finally the selected IS element sequences were validated and classified into IS families using the IS finder database (226).

Bifidoprophages identification.

The 625 bifidobacterial genomes were screened for prophage-like elements using a custom database based on already identified sequences through BLASTP analysis (171) (E-value cutoff of $1e^{-5}$). The custom database was constructed through previously bifidoprophage validate sequences retrieved from 60 bifidoprophages identified by Lugli *et al.* (99), considering genetic islands presenting different genes encoding for phage functions. Following this, a manual examination of the DNA region surrounding a putative phage-encoding gene was performed. These manual screening allowed us to identify complete prophage like-sequences, while discarding incomplete or remnant phage sequences, as previously performed by Lugli *et al.* (99).

Prediction of the antibiotic resistance genes.

The *in silico* proteome of 625 *Bifidobacterium* genomes used in this study was screened for proteins that can act as antibiotic resistance proteins through inactivation and/or removal of antibiotic molecules. The screening was carried out using the MEGAREs database through BLASTP analysis (E-value cutoff of $1e^{-18}$) (171, 227). The E-value cutoff was chosen based on a manual editing performed to identify false positive sequences. The core database was obtained by non-redundant compilation of sequences contained in Resfinder, ARG-ANNOT, the Comprehensive Antibiotic Resistance Database (CARD) and the NCBI Lahey Clinic beta-lactamase archive (228-231). Following this, a manual examination of the sequence with an E-value less than $1e^{-18}$ was performed in order to explore all the biodiversity of the AR genes of the *Bifidobacterium* species. We excluded the putative AR genes encoding for transporters for low accuracy in their prediction (232). The predicted AR genes were

classified according to the presumed mechanism of action and the antibiotic molecules they counteract.

Moreover, for the 625 bifidobacterial genomes analyzed, we manually evaluated the genes flanked by the predicted AR genes, forming the *Bifidobacterium* resistome, in order to identify mobile genetic hotspots that may promote HGT events.

Phylogenomic analyses.

The nucleotide similarity of each obtained bifidoprophage sequence was calculated using the software package LAST (233). Results were employed to build a matrix representing the genome similarity among different prophage and to generate a clustering tree. The bifidophage sequences were aligned using Mafft software (173) and the clustering tree was constructed using ClustalW (174). The constructed clustering tree was visualized using the FIGTREE software (http://tree.bio.ed.ac.uk/software/figtree/).

Bacitracin A antibiotic susceptibility tests.

The Minimal Inhibitory Concentration (MIC) breakpoints (micrograms per milliliter) of bacitracin A were determined using the broth microdilution method (MDIL) according to the ISO standard guidelines (234). Bacitracin A antibiotic was purchased from Merck (Germany). Microplates were incubated under anaerobic conditions for 48 h at 37°C. Cell density was monitored by optical density measurements at 600 nm (OD600) using a plate reader (BioTek, VT, USA). The MIC breakpoint represents the highest concentration of a given antibiotic to which a particular bacterial strain is resistant.

Statistical analyses.

SPSS software (IBM, Italy) was used to perform statistical analysis between BacA strains group and control group by T-student test.

RESULTS AND DISCUSSION

The putative resistome of the genus *Bifidobacterium*.

In order to investigate the genetic AR arsenal carried by members of the *Bifidobacterium* genus, we investigated the resistome of 625 bifidobacterial genomes. We enlarged the previously published database on the resistome of the *Bifidobacterium* genus, which were based on 91 different genomes (232). Putative AR genes encoding transporters were excluded from this analysis due to the inaccuracy of their bioinformatic prediction (232). The overall number of putative antibiotic resistance genes identified among these 625 genomes was 13,870, representing less than 1 % of the total *Bifidobacterium* genes analyzed (Table S1). According to the predicted mechanism of action and the antibiotic molecules that could be counteracted, seven different AR gene classes were identified (Fig. 1). The AR class with the highest number of representatives was the one conferring glycopeptide resistance, which corresponds to 5,999 putative enzymes acting against glycopeptide antibiotics, such as vancomycin, teicoplanin and telavancin (Fig. 1) (235-237). Notably, *Bifidobacterium bifidum* 791, *Bifidobacterium longum* subsp. *infantis* 1888B and *B. bifidum* AM42-15AC were strains containing the highest number of genes predicted to belong to this glycopeptide resistance class, each encoding 29 distinct enzymes predicted to confer such resistance. Moreover, we identified 2,178 genes putatively belonging to the tetracycline resistance class (Fig. 1) (238-241). Members of the *B. bifidum* species, isolated from fecal samples of healthy Chinese individuals (242), i.e., strains TM05-15, TF07-22, TM02-15, TM02-17, TM06-10 and TM07-4AC, were shown to contain the highest number of genes encoding proteins predicted to counteract tetracycline antibiotics, ranging from 29 genes of *B. bifidum* TM05-15 to 28 genes for the other *B. bifidum* strains. Notably, 484 analyzed strains did not appear to contain genes encoding tetracycline resistance proteins, representing the 77.5 % of the total *Bifidobacterium* strains analyzed. Furthermore, 2,437 genes were found to belong to beta-

lactamase class and *Bifidobacterium animalis* subsp. *animalis* ATCC 25527 was shown to be the strain with the highest number (i.e. 32) of predicted beta-lactamase-encoding genes, while 469 of the 625 analyzed genomes did not appear to encompass genes belonging to this AR class (Fig. 1).

Moreover, 2,618 genes were predicted to belong to the methyltransferase AR class, including 23S rRNA methyltransferase which may confer resistance toward erythromycin and clindamycin, as demonstrated in a previous study (243) (Fig. 1). In addition, we identified 500 genes predicted to belong to the sulfonamide-resistance class, which includes genes encoding enzymes counteracting sulfonamide antimicrobial agents, also known as *sul* genes (244) (Fig. 1). The *sul* gene appears to be present as three variants in the investigated genomes, i.e., s*ul1*, *sul2* and *sul3*, all encoding a dihydropteroate synthase (244, 245). Interestingly, in the assessed genomes of the *Bifidobacterium* genus the most prevalent gene variant was *sul3*, found in 90.4 % of all identified sulfonamide-resistance genes.

Finally, the aminoglycoside class and the metronidazole class were the two least represented classes of AR genes in bifidobacteria, with just 73 and 64 identified genes predicted to be members of these two respective classes (Fig. 1). Notably, *B. bifidum* AF11-25B was predicted to contain the highest number of genes encoding enzymes that counteract aminoglycoside antibiotics, such as streptomycin, kanamycin and gentamicin.

Moreover, *B. bifidum* TF05-1 was the only strain whose chromosome contains a gene encoding a putative quinolone resistance protein. This gene encodes a pentapeptide repeat protein, which is predicted to be involved in the fluoroquinolone resistance (246-248). Interestingly, comparing the identified *Bifidobacterium* resistome with antibiotic resistance determinants of other gut commensal, such as members of *Lactobacillus* genus, we observed a less complexity of the resistome (249). In fact, the *Lactobacillus* genus included different genes that could confer resistance toward a wide range of antibiotic molecules, such as

vancomycin, erythromycin, penicillin, but also tetracycline, chloramphenicol and aminoglycoside antibiotics (250-254). Furthermore, different *Escherichia coli* strains presented in their genomes AR genes which counteracted carbapenem antibiotics (255, 256), whereas bifidobacteria seemed to be very sensitive to this antibiotic class, and their genomes do not encompass any genes that could confer resistance toward this antibiotic. Notably, a recent study based on metagenomics analyses of the human gut microbiota revealed that *Entrococcus* and *Enterobacter* genera possessed very high antibiotic resistance load (257). Moreover, different studies have been demonstrated the presence of AR genes in the genomes of the members of *Bacillus* genus, used as probiotic bacteria in functional food and for animal feed (249, 258). In the latter genus have been identified macrolide-resistance genes present on extra-chromosomal elements, tetracycline resistance genes, but also *cfr*-like genes (i.e., conferring resistance toward several classes of antibiotics including phenicols, oxalozidinone, lincosamides, pleuromutilinis and streptogramin) that have not been identified in the genomes of the members of the *Bifidobacterium* genus (259-261). Our resistome analyses revealed the lacking of specific *Bifidobacterium* AR genes, corroborating the safer behavior of the *Bifidobacterium* genus compared to other human gut commensals. Although we do acknowledge the limitations of the *in silico* analysis in assigning antibiotic resistance functions to these identified genes, they are nonetheless considered to represent the potential arsenal to counter antimicrobial molecules.

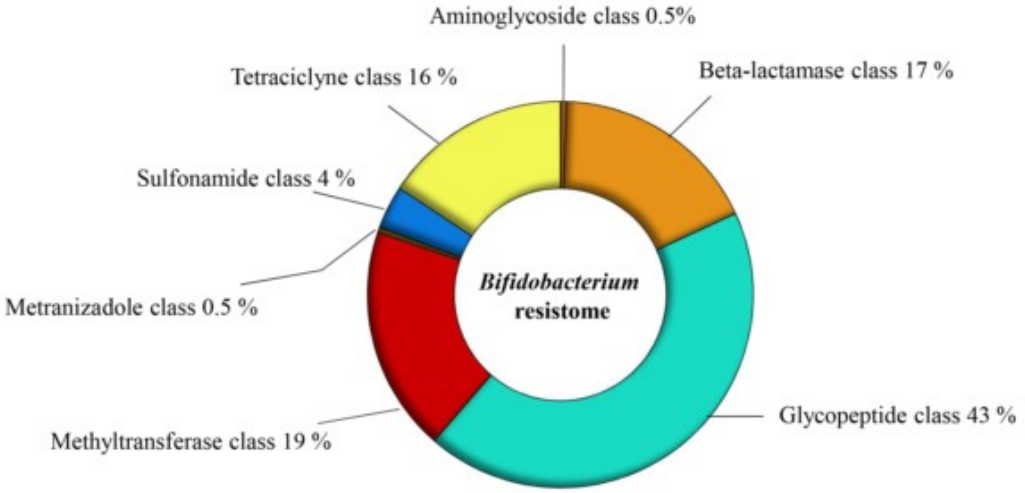


Figure 1

Figure 1. Predicted resistome of the *Bifidobacterium* genus. Abundance of different predicted antibiotic resistance gene classes identified among the 625 analyzed *Bifidobacterium* genomes.

The predicted mobilome of the *Bifidobacterium* genus.

The mobilome is defined as genetic elements that can move within a genome and between different genomes, including transposable elements, bacteriophages and plasmids (262-264). Similar to other members of the gut microbiota, it has been demonstrated that bifidobacteria possess genetic elements whose action is responsible for shaping their genomes (58, 99, 264, 265). In order to explore the mobile element repertoire of the *Bifidobacterium* genus, we analyzed the same 625 *Bifidobacterium* genomes as indicated above (Table S1). Our analyses updated previously published data based on the reconstruction of 60 different bifidoprophage-like elements of 48 species of the *Bifidobacterium* genus.

A screening among analyzed bifidobacterial genomes revealed 16,065 different genes encoding transposases, excluding genes that were truncated at the start or end codon (Table S1). The sequence of each IS element has been classified according to the ISFinder database (266), showing that members of the IS3 family are the most widespread among the

Bifidobacterium genus (Table S2). Notably, members of the *Bifidobacterium breve* species showed the highest number of IS elements, i.e., strains BR-06, BR-H29, BR-21, BR-L29 and BR-C29, ranging from 174 to 102 (Table S1). Moreover, 16.5 % of the analyzed genomes were predicted to contain less than 10 genes encoding transposases in their chromosomes, while *Bifidobacterium commune* LMG 28292 does not appear to encompass any IS element at all (Table S1).

Recently, Lugli *et. al.* recognized and classified all prophage-like elements (referred to as bifidoprophages) present in 48 genomes of type strains belonging to different bifidobacterial species (99) and Mavrich *et al.* characterized three of these identified groups of prophages integrated in members of *B. breve* and *B. longum* species by means of induction experiments (267). In the current study, the screening for bifidoprophages was further extended to 625 different bifidobacterial genomes, resulting in the identification of 598 putative and apparently complete prophage sequences (Table S1). Notably, the genomes of *Bifidobacterium biavatii* DSM 23969, *Bifidobacterium imperatoris* LMG 30297 and *Bifidobacterium cuniculi* LMG 10738 were predicted to contain the highest number of prophages in their genomes, i.e., seven, six and five prophage-like elements, respectively (Table S1). In order to evaluate the homology among identified prophage-like elements a genomic based-alignment clustering was performed. We observed the presence of four main homology clusters, in which the taxonomic origin of the corresponding *Bifidobacterium* hosts was highly heterogeneous. Each identified cluster showed several sub-clusters consisting of different prophage-like elements belonging to bifidobacterial strains of the same species, highlighting a sub-cluster phage specificity that appears to be host-related (Fig. 2). As reported in previous studies, prophages contribute to the genetic individuality of bacterial strains, containing many unique genes that in some instances may confer fitness advantage to the host such as gene related to antibiotic resistance (268-270). The mobile nature of

phages may then allow transfer of such advantageous genes to human/animal pathogens or to other non-resident microorganisms of the gut microbiota.

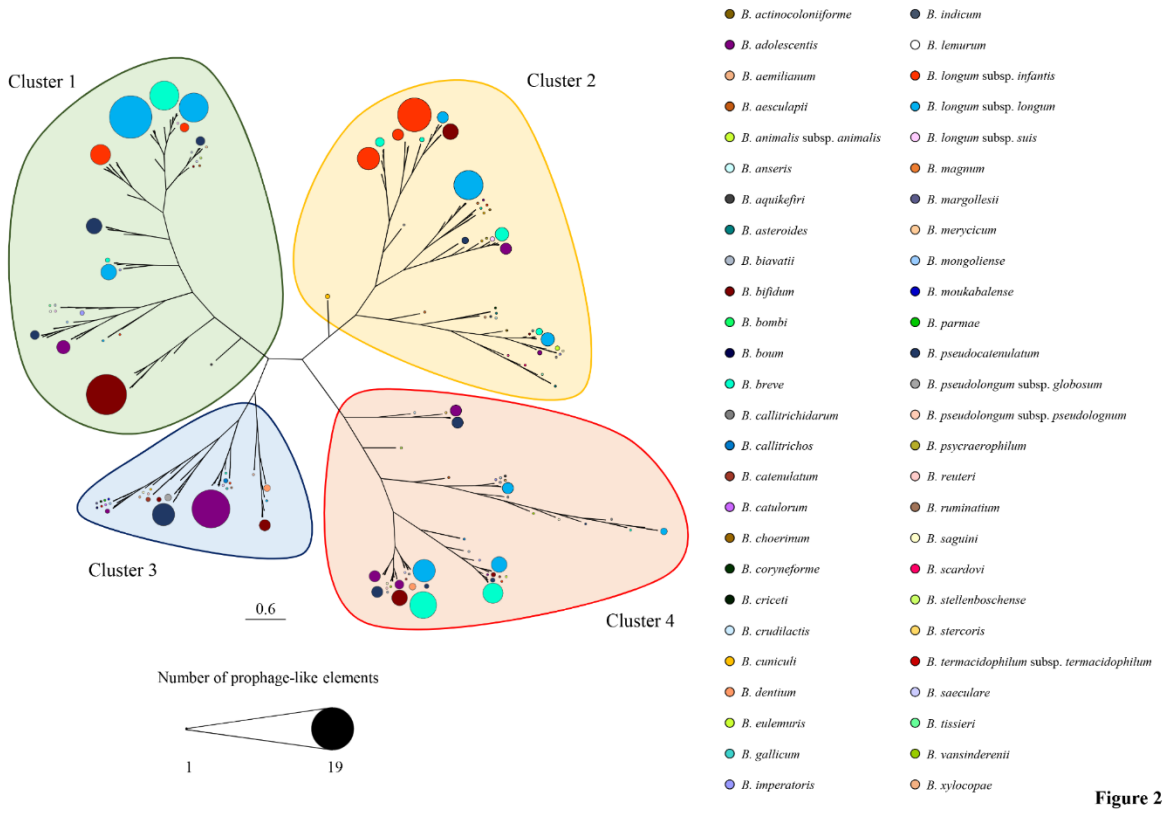


Figure 2. Phylogenetic tree of identified bifidoprophages. Genomic alignment-based clustering of 598 prophages identified within bifidobacterial strain genomes. Each colored dot represents the *Bifidobacterium* host species origin of a given bifidoprophage. The dot size refers to the number of prophage-like sequence identified within the same branch-tree. The four different clusters are highlighted with different color.

Identification of the putative mobile resistome of *Bifidobacterium* genus.

In order to evaluate the insurgence of AR genes located on or close to mobile elements, such as transposases and bifidoprophages, we investigated the flanking genes of the predicted resistome of the 625 bifidobacterial genomes. These regions may represent Mobile Genetic Hotspots (MGHs) that promote HGT events, thereby transferring antibiotic resistance to

other bacteria. We identified 201 putative MGHs distributed in 120 of the 625 *Bifidobacterium* strains studied. The number of AR genes involved in MGHs were very small compared to the total number of resistome genes (i.e. 13,870) representing less than 1.5 % of the total *Bifidobacterium* resistome. Interestingly, we could not observe a correlation between a specific type of IS element and a class of AR genes.

As already noted in previous studies, 37 of the 41 strains of *Bifidobacterium animalis* subsp. *lactis* species contain a *tet*W gene flanked by a putative conjugative transposon (Fig. 3) (127, 241, 271, 272). The *tet*W gene encodes a protein belonging to the GTP-binding elongation factor family that protects ribosomes from translation inhibition activity of tetracycline (273). Notably, this MGH is also present in 15 other genomes belonging to members of the *Bifidobacterium adolescentis*, *B. animalis* subsp. *animalis*, *B. breve*, *Bifidobacterium longum* spp., *Bifidobacterium pseudolongum* subsp. *pseudolongum* and *Bifidobacterium pullorum* species. Remarkably, *tet*W appears to be well conserved among different species (Fig. 3), suggesting the involvement of HGT events that could have transferred this tetracycline resistance gene to different bifidobacterial strains.

Interestingly, 67 MGHs involved prophage-like elements, which harbor a gene encoding for an UDP pyrophosphate phosphatase within their sequence (Fig. 3), revealing a domain in the amino acid sequence that resembles a bacitracin resistance protein (BacA) (274, 275). These 67 MGHs were present in members of three different *Bifidobacterium* species, i.e., *B. breve*, *B. longum* spp. and *Bifidobacterium pseudocatenulatum*, putatively conferring resistance to bacitracin through the phosphorylation of undecaprenol (274, 275). Prophages influence the biodiversity and abundance of bacteria in the human/animals intestinal tract, conferring new capabilities to their host (269). The acquisition of a prophage-like element may thus confer a fitness advantage (269, 270), in this particular case by conferring bacitracin resistance to these *Bifidobacterium* strains.

Moreover, a gene encoding a 23S rRNA methyltransferase flanked by a transposase was identified in 53 putative MGHs (Fig. 3). In a recent study, Martinez *et al.* demonstrated the existence of a 23S rRNA methylase that confers erythromycin and clindamycin resistance to *B. breve* CECT7263 (243). We found these MGHs in ten different *Bifidobacterium* species, including *B. adolescentis*, *B. bifidum*, *B. breve*, *Bifidobacterium choerinum*, *Bifidobacterium kashiwanohense*, *B. longum* spp., *B. pseudocatenulatum*, B. *pseudolongum* subsp. *pseudolongum* and *B. pullorum*. The highest occurrence of this genetic hotspot was in *B. breve* strains, where this hotspot was present in 15 out of 88 *B. breve* genomes analyzed. This methyltransferase is responsible for the enzymatic modification of the nucleotide sequence of the 23S rRNA gene, adding a methyl group, preventing the linking of macrolide molecules (243). Notably, the transposases that encompass these MGHs are predicted to be replicative transposons that may cause a rearrangement within bifidobacterial genomes, indicating that these MGHs rarely transfer to other genomes.

Remarkably, *B. longum* subsp. *longum* E18, isolated from healthy-adult feces sample (276), is the only strain whose chromosome contains a prophage-like element including a gene predicted to encode a protein with a complete beta-lactamase domain (Fig. 3). Furthermore, the genome of strain *Bifidobacterium parmae* LMG 30295 contains a *van*Z homolog flanked by a predicted transposase-encoding gene, belonging to the transposon family IS256. The *van*Z gene is predicted to confer low-level resistance to the glycopeptide antibiotic teicoplanin (Te), which prevents incorporation of D-alanine into peptidoglycan precursors. This hotspot did not include a conjugative transposon, decreasing possible transfer events and bringing to possible genomic rearrangements (277, 278). Therefore, more than 50 % of putative MGHs identified encompassed transposons that can not be classified as conjugative transposons, reducing possible HGT events involving AR genes, corroborating previously published data (232).

The distribution of putative AR genes among analyzed bifidobacteria could be due to selective pressure imposed by intensive antibiotic use in their animal/human hosts, similar to what has been observed for lactic acid bacteria (LAB) (232, 249). These findings underline the safety of this genus and the very low frequency by which these AR genes may transfer to other members of the gut microbiota.

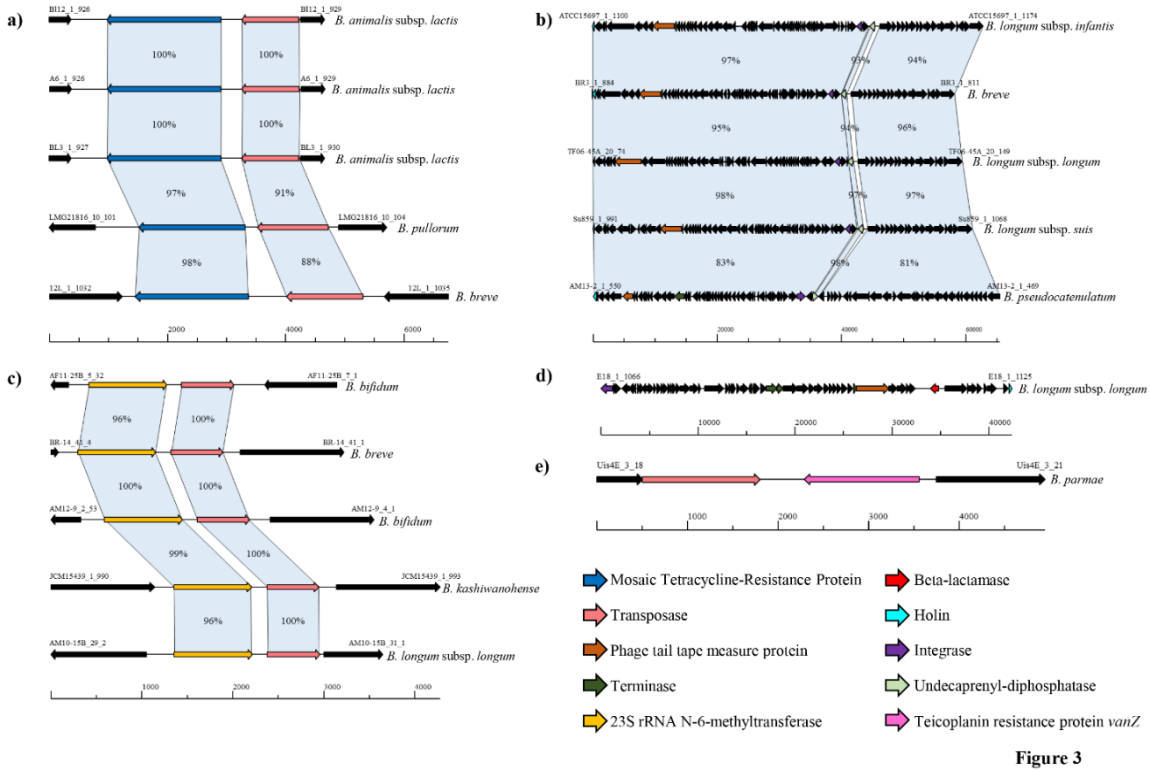


Figure 3

Figure 3. Mobile genetic hotspots identified in *Bifidobacterium* genus. Bifidobacterial genomic regions containing putative MGHs. Different species and gene names are reported next to each genomic region. Panels a to c show genomic regions conserved among different *Bifidobacterium* species. Panels d and e display unique mobile genetic hotspots identified in *B. longum* subsp. *longum* E18 and *B. parmae* LMG 30295 strains. Each arrow indicates a gene and the different colors indicate the function of the gene product.

Assessment of bacitracin A resistance of *Bifidobacterium* spp.

In order to validate our *in silico* predictions, we further investigated the antibiotic resistance of bifidobacterial strains whose genomes were shown to contain a *bac*A gene located in the sequence of a prophage-like element. Thus, *in vitro* measurements of MIC breakpoints for the bacitracin A antibiotic were monitored including three *Bifidobacterium* strains, i.e., *B. breve* 1891B, *B. longum* subsp. *longum* 35B and *B. longum* subsp. *infantis* ATCC 15697, whose genomes encompasses a predicted *bac*A gene and three additional strains as control, i.e., *B. breve* LMG 13208, *B. longum* subsp. *longum* LMG 13197 and *B. longum* subsp. *infantis* 1888B, whose chromosomes do not include a predicted *bac*A gene.

As indicated by this *in silico* analysis, those strains containing the *bac*A gene in their genomes exhibit a higher resistance level to bacitracin A (ranging from 16 fold to 32 fold) when compared to control strains (Fig. S1). In this context, the bacitracin A breakpoints MIC values of *B. breve* 1891B, *B. longum* subsp. *longum* 35B and *B. longum* subsp. *infantis* ATCC 15697 were, respectively, 16 µg/mL, 32 µg/mL and 64 µg/mL, whereas the MIC values of the members of the control group were 2 µg/mL for *B. breve* LMG 13208 and 1 µg/mL for *B. longum* subsp. *longum* LMG 13197 and *B. longum* subsp. *infantis* 1888B (Fig. S1). Statistical analyses were performed to corroborate the observed MIC differences, resulting in a significant growth difference between the two groups analyzed (P-value <0.001) (Fig. S1). These results confirmed the *in silico* predicted resistance to bacitracin of those bifidobacterial strains possessing the MGHs related to the *bac*A gene. The fact that the *in silico* analyses matched with the *in vitro* data highlighted the validity of an *in silico* resistome prediction (279).

CONCLUSIONS

Bifidobacteria are dominant members of the human/animals GIT, especially during the early stage of life. It has previously been demonstrated that the presence of this genus in the microbiota is associated with health-promoting effects (280). In the current study, we reconstructed the resistome and the mobilome of members of the *Bifidobacterium* genus, evaluating genetic hotspots that could be involved in HGT events. The reconstructed putative resistome revealed that only a limited number of bifidobacterial genes are likely to be involved in putative AR spread. Moreover, the AR genetic arsenal of the *Bifidobacterium* genus seems to be less complex compared to the resistome of other Gram-positive bacteria, such as members of the *Lactobacillus* genus or other species included in food supplements and used as probiotics, such as members of the *Bacillus* genus (249, 251, 260, 281, 282). Identified MGHs were restricted to less than 20 % of the analyzed strains, of which most isolated from the human GIT, suggesting the occurrence of AR in members of the human microbiota because of intense antibiotic therapies. Remarkably, the acquisition of phages encompassing AR genes in their sequence could confer ecological advantages, increasing the biological fitness of their host (269, 270). Nevertheless, the vast majority of identified MGHs in the *Bifidobacterium* genus are unlikely to be transferred to other microorganisms, due to the transposition mechanisms of the identified IS elements flanking putative AR genes. Moreover, *in vitro* bacitracin A antibiotic resistance tests based on bifidobacterial strains containing *bac*A located in an MGH confirmed our *in silico* prediction. Finally, these findings underpin the safety of the *Bifidobacterium* genus compared to other taxa such as *Escherichia coli* and members of the Gammaproteobacteria class, which were shown to contribute to a high antibiotic resistance load in the human microbiota (283).

Chapter 6

Investigating bifidobacterial resistance against amoxicillin-clavulanic acid: genetic and ecological insights

Mancino W, Mancabelli L, Lugli GA, Viappiani A, Anzalone R, Longhi G, van Sinderen D, Ventura M, Turroni F.

The results of this chapter were submitted to Applied and Environmental Microbiology.

ABSTRACT

Amoxicillin-Clavulanic acid (AMC) is one of the most frequently prescribed antibiotic formulations in the Western world. Extensive oral use of this antimicrobial combination influences the gut microbiota. One of the most abundant early colonizers of the human gut microbiota is represented by different taxa of the *Bifidobacterium* genus, which include many members that are considered to bestow beneficial effects upon their host. In the current study, we investigated the impact of AMC administration on the gut microbiota composition, comparing the gut microbiota of 23 children that had undergone AMC antibiotic therapy to that of 19 children that had not been treated with antibiotics during the preceding six months. Moreover, we evaluated the sensitivity to AMC by Minimal Inhibitory Concentration (MIC) test of 261 bifidobacterial strains, including reference strains for the currently recognized 64 bifidobacterial (sub)species, as well as 197 bifidobacterial isolates of human origin.. These assessments allowed the identification of four bifidobacterial strains, which exhibit a high level of AMC insensitivity, and which were subjected to genomic and transcriptomic analyses to identify the genetic determinants responsible for this AMC insensitivity. Furthermore, we investigated the ecological role of AMC-resistant bifidobacterial strains by in vitro batch-cultures.

IMPORTANCE

Based on our results, we observed a drastic reduction in gut microbiota diversity of children treated with antibiotics, especially affecting bifidobacterial abundance. Consequently, the MIC experiments revealed that more than the 98% of the strains tested resulted to be inhibited from AMC antibiotic. The isolation of four insensitive strains and the sequencing of their genome allowed understanding the possible genes involved in the resistance mechanisms. Moreover, gut-simulating *in-vitro* experiments revealed that one strain, i.e. *B. breve* 1891B took over other bacteria also in presence of AMC.

MATERIALS AND METHODS

Sample collection.

For the purpose of this study a total of 42 human fecal samples were collected, divided in two groups, the first one was represented by fecal samples obtained from 23 children, who were undergoing AMC (Amoxicillin:Clavulanic acid ratio of 7:1) treatment from seven to 10 days (Average of 8.9 days ± 1.3), while the second group was represented by fecal samples from 19 healthy children (Table S1). Collected samples, which consisted of approximately 10 g of fresh fecal material, were kept on ice, shipped under sub-zero conditions to the laboratory and stored at -80 °C until further processing.

Bacterial DNA Extraction, 16S rRNA Gene PCR Amplification and Sequencing.

Stool samples were subjected to DNA extraction using the QIAamp DNA Stool Mini kit following the manufacturer's instructions (Qiagen). Partial 16S rRNA gene sequences were amplified from extracted DNA using primer pair Probio_Uni / Probio_Rev, targeting the V3 region of the 16S rRNA gene sequence (284). Illumina adapter overhang nucleotide sequences were added to these partial 16S rRNA gene-specific amplicons, which were further processed employing the 16S Metagenomic Sequencing Library Preparation Protocol (Part #15044223 Rev. B – Illumina). Amplifications were carried out using a Verity Thermocycler (Applied Biosystems). The integrity of the PCR amplicons was analyzed by electrophoresis on a 2200 TapeStation Instrument (Agilent Technologies, USA). DNA products obtained following PCR-mediated amplification of the 16S rRNA gene sequences were purified by a magnetic purification step involving the Agencourt AMPure XP DNA purification beads (Beckman Coulter Genomics GmbH, Bernried, Germany) in order to remove primer dimers. DNA concentration of the amplified sequence library was determined by a fluorimetric Qubit quantification system (Life Technologies, USA). Amplicons were diluted to a concentration of 4 nM, and 5 µL quantities of each diluted DNA amplicon sample were mixed to prepare the pooled final library. Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3

chemicals. Following sequencing, the .fastq files were processed using a custom script based on the QIIME software suite (212). Paired-end read pairs were assembled to reconstruct the complete Probio_Uni / Probio_Rev amplicons. Quality control retained sequences with a length between 140 and 400 bp and mean sequence quality score > 20, while sequences with homopolymers > 7 bp and mismatched primers were omitted. In order to calculate downstream diversity measures (alpha and beta diversity indices, Unifrac analysis), 16S rRNA Operational Taxonomic Units (OTUs) were defined at 100 % sequence homology using DADA2 (285); OTUs not encompassing at least 2 sequences derived from the same sample were removed. Notably, this approach allows highly distinctive taxonomic classification at single nucleotide accuracy (285). All reads were classified to the lowest possible taxonomic rank using QIIME2 (212, 214) and a reference dataset from the SILVA database. Biodiversity within a given sample (alpha-diversity) was calculated based on the observed OTU index. Similarities between samples (beta-diversity) were calculated by weighted uniFrac (286). The range of similarities is calculated between values 0 and 1. PCoA representations of beta-diversity were performed using QIIME2 (212, 214).

Evaluation of cell density by flow cytometry assay.

For bacterial cell counting, 0.2 g of fecal sample was diluted in physiological solution (Phosphate-buffered saline, PBS). Subsequently, bacterial cells were stained with one µL SYBR®Green I and incubated in the dark for at least 15 min before measurement. All count experiments were performed in triplicate using an Attune NxT Flow flow cytometer (Invitrogen, ThermoFisher Scientific) equipped with a Blue Laser set at 50 mWatt and tuned to an excitation wavelength of 488 nm. Multiparametric analyses were performed on both scattering signals (FSC, SSC) and SYBR Green I fluorescence was detected on FL1 channel. Cell debris and eukaryotic cells were excluded from acquisition analysis by a sample-specific FL1 threshold. All data sets were statistically analyzed with Attune NxT Flow Cytometer Software. Utilizing these cell counts to normalize the sequencing data

into absolute abundance of each profiled taxa, we were capable to perform quantitative microbiome profiling using a previously described method (287).

Isolation of novel *Bifidobacterium* strains.

In order to explore the AMC resistance of *Bifidobacterium* genus, 24 novel strains were isolated from fecal samples of individuals that had been treated with AMC for a varying number of days (Table 1). One gram of feces from each collected fecal sample was mixed with nine mL of PBS (pH 6.5). Serial dilutions and subsequent plating were performed using MRS agar, supplemented with 50 μg/mL mupirocin (Delchimica, Italy), 0.05 % (wt/vol) L-cysteine hydrochloride and 8 μg/mL of AMC (Merck, Germany). Morphologically distinct colonies that developed on MRS plates were randomly picked and re-streaked in order to isolate purified bacterial strains. All novel isolates were subjected to DNA isolation and characterized as previously described (164). The *Bifidobacterium* strains isolated in this study are listed in Table 1.

Amoxicillin-clavulanic acid susceptibility tests.

The MIC breakpoints (micrograms per milliliter) of AMC were determined using the broth microdilution method (MDIL) (234). Microplates were incubated under anaerobic conditions for 48 h at 37°C. Cell density was monitored by optical density measurements at 600 nm (OD600) using a plate reader (BioTek, VT, USA). Furthermore, the same MIC analysis was performed for the antibiotic amoxicillin alone (European Standard, Merck, Germany). The MIC breakpoint represents the highest concentration of a given antibiotic to which a particular bacterial strain was shown to be resistant.

Genome sequencing and assemblies.

DNA samples extracted from AMC-resistant bifidobacterial isolates were subjected to genome sequencing using MiSeq (Illumina, UK) at GenProbio srl (Parma, Italy) according to the supplier's protocol (Illumina, UK). 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 (166). SPAdes software was used for *de novo* assembly of each *Bifidobacterium* genome sequence (167, 168), while protein-encoding ORFs were predicted using Prodigal (169).

Comparative Genomics.

In order to identify unique protein families encoded by new isolated AMC resistant *Bifidobacterium* strains a PGAP analyses was performed (170). *B. breve* 1891B and *B. breve* M1D genomes were analyzed with six other *B. breve* genomes, tested for AMC resistance (Table S3). Simultaneously, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B genomes were compared with five *B. longum* subsp. *longum* genomes used in the MIC_{AMC} analyses (Table S3). Each predicted proteome of a given *Bifidobacterium* genome was screened for orthologues against the proteome of every *Bifidobacterium* strain belonging to the same species by means of BLAST analyses (171) (cutoff, E value of $<1 \times 10^{-4}$ and 50 % identity across at least 80 % of both protein sequences). Protein families shared between analyzed genomes allowed us to identify the core genome of the *B. breve* and *B. longum* subsp. *longum* (sub)species.

Functional annotation of each protein of *B. breve* 1891B was performed employing the eggNOG database (187).

Prediction of antibiotic resistance genes.

The *in silico* proteome of four *Bifidobacterium* genomes isolated in this study was screened for proteins with similarity to antibiotic resistance proteins acting through inactivation and/or removal of antibiotic molecules. The screening was carried out using the MEGAREs database through BLASTP analysis (E-value cutoff of $1e^{-5}$) (171, 227, 288). The core database was obtained by non-redundant compilation of sequences contained in Resfinder, ARG-ANNOT, the Comprehensive Antibiotic Resistance Database (CARD) and the NCBI Lahey Clinic beta-lactamase archive (228-231). Following this, a manual examination of sequences with an E-value below $1e^{-5}$ was performed in order to detect distant homologs.

RNA extraction.

Aliquots of *B. breve* 1891B cells were grown to an optical density at 600 nm ranging from 0.6 to 0.8. The experiment was conducted in three biological replicates. Total RNA was isolated from bifidobacterial cultures grown in MRS or MRS with 32 μg/mL of AMC. RNA extraction was performed as previously described (289). Briefly, cultures were centrifuged at 4000 rpm for 10 min at 4ºC. Cells pellets were treated with TES buffer (Tris - Ethylenediaminetetraacetic acid - Sodium dodecyl sulfate buffer) and lysozyme followed by resuspension in 1 mL of QIAzol Lysis Reagent (Qiagen, UK) and placed in a sterile tube containing glass beads (Merck, Germany). The cells were lysed by shaking the mix on MINI-Bead Beater 24 (BioSpec Products, USA) for 2 min followed by 2 min of static cooling; this step was repeated for three times. The lysate was centrifuged at 12,000 x *g* for 15 min, and the upper phase was recovered. RNA samples were washed from proteins by means of chloroform and finally samples were purified by means of the RNeasy mini kit (Qiagen, UK), following the manufacturer's instructions. Quality and integrity of the RNA were checked by the Tape station 2200 (AgilentTechnologies, USA) analysis. RNA concentration was then determined by a fluorimetric Qubit quantification system (Life Technologies, USA).

RNAseq analysis performed by the MiSeq Illumina.

One μg of total RNA was treated by the TruSeq Stranded Total RNA protocol (Illumina, USA), following manufacturer's instructions. Ribosomal depletion is included in the protocol. Quality and quantity of libraries were checked by the Tape station 2200 (AgilentTechnologies, USA) analysis and fluorimetric Qubit quantification system (Life Technologies, USA), respectively. Samples were loaded into a Flow Cell V3 600 cycles (Illumina) as reported by the technical support guide. The reads were depleted of adapters, quality filtered (with overall quality, quality window and length filters) and aligned to the *Bifidobacterium* reference genomes through Bowtie2 software (290). RPKM values were evaluated by means of Artemis software (291).

***In vitro* simulation of the effect of AMC and *Bifidobacterium* strains on the human gut microbiota.**

The effect of AMC antibiotic and four selected AMC-insensitive *Bifidobacterium* strains on the human gut microbiota was evaluated *in vitro* through anaerobic, pH- and temperature-controlled batch cultures. We evaluated four different conditions and the growth medium used was based on the fecal medium described by Macfarlane et al. (292). For each batch culture experiment, a fresh fecal sample from the same donor was used. Fecal sample was collected from a child, aged 4 years, who had not undergone antibiotic treatment for at least three months prior to sample collection, and who had not consumed probiotic bacteria. The fresh fecal sample was previously analyzed in order to confirm the absence of *B. breve* and *B. longum* species DNA. We tested four different growth conditions for each strain. Briefly, the first condition consisted of 40 mL of growth media supplemented with AMC (20 μM) in which an overnight culture of a particular *Bifidobacterium* strain and the fresh fecal sample were inoculated, each at 1% (v/v and w/v, respectively). The concentration of AMC used was based on another study in which different antibiotic and non-antibiotic compounds were tested for their impact on the gut microbiota (293). The second condition consisted of 40 mL of

fecal media inoculated with 1 % (v/v) of an overnight culture of a *Bifidobacterium* strain and 1 % (w/v) of the fresh fecal sample. The third batch culture was made of a 40 mL of fecal media supplemented with AMC 20 µM and in which the fecal sample was inoculated at 1 % (w/v). Finally, the fourth condition was composed of a 40 mL of growth media in which only the fecal sample was inoculated at 1 % (w/v). For each experiment, an aliquot of culture was taken at four different time points: 12h, 18h, 24h and 36h after inoculum. Each aliquot was subjected to DNA extraction using the QIAamp DNA Stool Mini kit following the manufacturer's instructions (Qiagen, UK) for sequencing library preparation.

Shallow shotgun metagenomics and evaluation of cell density by flow cytometry assay of co-culture experiments.

Extracted DNA was prepared for sequencing purposes following the Illumina Nextera XT protocol. Briefly, DNA samples were enzymatically fragmented, barcoded and purified involving magnetic beads. Then, samples were quantified using fluorometric Qubit quantification system (Life Technologies, USA), loaded on a 2200 Tape Station Instrument (Agilent Technologies, USA) and normalized to 4 nM. Sequencing was performed single-end using an Illumina MiSeq sequencer with flow cell v3 600 cycles (Illumina Inc., San Diego, USA). For bacterial cell counting, the batch cultures were diluted in PBS physiological solution. Subsequently, bacterial cells were stained with one µL SYBR®Green I and incubated in the dark for at least 15 min before measurement. All count experiments were performed in triplicate as described previously (see above). All data sets were statistically analyzed with Attune NxT Flow Cytometer Software. Utilizing these cell counts to normalize the sequencing data into absolute abundance of each profiled taxa, we were able to perform quantitative microbiome profiling using a previously described method (287).

Statistical analysis.

SPSS software (IBM, Italy) was used to perform statistical analysis between AMC group and Healthy group, and for RNAseq data by Student t-test. PCoA statistical analysis were performed by PERMANOVA test.

Data Deposition.

Raw sequences of the 16S profiling experiments, shallow shotgun metagenomics experiments and RNAseq experiments are accessible through SRA study accession numbers PRJNA663786. Newly isolated *Bifidobacterium* genomes were sequenced and deposited at DDBJ/ENA/GenBank under accession numbers reported in Table 2.

RESULTS & DISCUSSION

Microbial composition of gut microbiota of children taking AMC. In order to assess the differences between the gut microbiota composition of 23 children undergoing AMC therapy (AMC group) compared to that of 19 healthy, age-matched individuals (CTRL group) that had not been taking any antibiotic during the preceding six months, 16S rRNA gene-based microbial profiling analyses were carried out on fecal samples, as described previously (284). The analyses resulted in a total of 2,274,698 reads with an average of 54,159 ± 11,080 reads per sample. Statistical Whisker plot analyses revealed a notable difference between the two groups analyzed (Fig. 1). Specifically, the 16S rRNA gene-based analysis showed a higher diversity of the CTRL group microbiota when compared to the AMC group (p-value < 0.05), demonstrating that the microbiota composition is significantly influenced by the antibiotic therapy to which individuals were subjected (Fig. 1). In order to explore the absolute bacterial abundance of the analyzed fecal samples, we used a quantitative microbiome profiling approach based on flow cytometric analyses for the enumeration of microbial cells present in each fecal sample assayed. This analysis allowed the identification of the microbial load of the 42 fecal samples, which was then used to normalize the 16S rRNA sequencing data. This procedure therefore generated absolute abundance data for each profiled taxa as previously described (287). Notably, the absolute bacterial abundance was shown to be significantly different between the two groups (average absolute abundance of 2,448,888 ± 2,875,900 and 5,438,290 ± 3,345,390 in AMC and CTRL groups, respectively, p-value < 0.001), revealing that the microbial abundance of the AMC samples was significantly reduced when compared to that of the CTRL group (Fig. 1). Interestingly, these analyses revealed differences in the composition, at genus level of the AMC group compared to the CTRL group microbiota, probably due to the intensive selective pressure imposed by the antibiotic therapy to which the AMC children had been subjected. In particular, comparison of the gut microbiota composition of AMC and CTRL groups revealed significantly (p-value < 0.05) higher absolute abundance of the *Barnesiella*, *Bacteroides*, *Alistipes*, *Parabacteroides*, *Blautia*, *Faecalibacterium*, *Dialister* and *Odoribacter* genera in CTRL samples (ranging from 2.1-fold for the

Bacteroides genus to 13.7-fold in the case of the *Dialister* genus) (Fig. 1). This finding is consistent with other studies that had previously assessed the microbiota of individuals treated with antibiotics (294-297). Interestingly, the comparison of the gut microbiota between the AMC and the CTRL groups also showed a 1.8-fold decrease in absolute abundance of the genus *Bifidobacterium*. Bifidobacteria are widely considered to represent a positive biomarker for a healthy gut status (217, 280, 298, 299). The decrease in members of this genus is consistent with previous data that had also observed a decrease in bifidobacterial abundance following antibiotic exposure (300, 301). In order to evaluate the inter-individual differences between AMC and CTRL samples, we analyzed the beta diversity based on weighted UniFrac distance metric and represented the obtained results through Principal Coordinate Analysis (PCoA) (Fig. 1). Interestingly, the PCoA analysis showed that most of the samples were grouped as two different clusters corresponding to AMC or CTRL samples, respectively, thus underlining the distinct microbiota composition of individuals that had or had not been treated with AMC (Fig. 1). Moreover, a PERMANOVA statistical analysis on the PCoA data revealed a significant division between the two analyzed groups (p-value < 0.05). These findings corroborate the fact that microbiota profiles differ considerably between samples, yet in the case of children that had been treated with antibiotics generally it was clear that such samples had suffered from bacterial depletion when compared to control samples, underlining the expected detrimental effect of AMC on the microbiota composition.

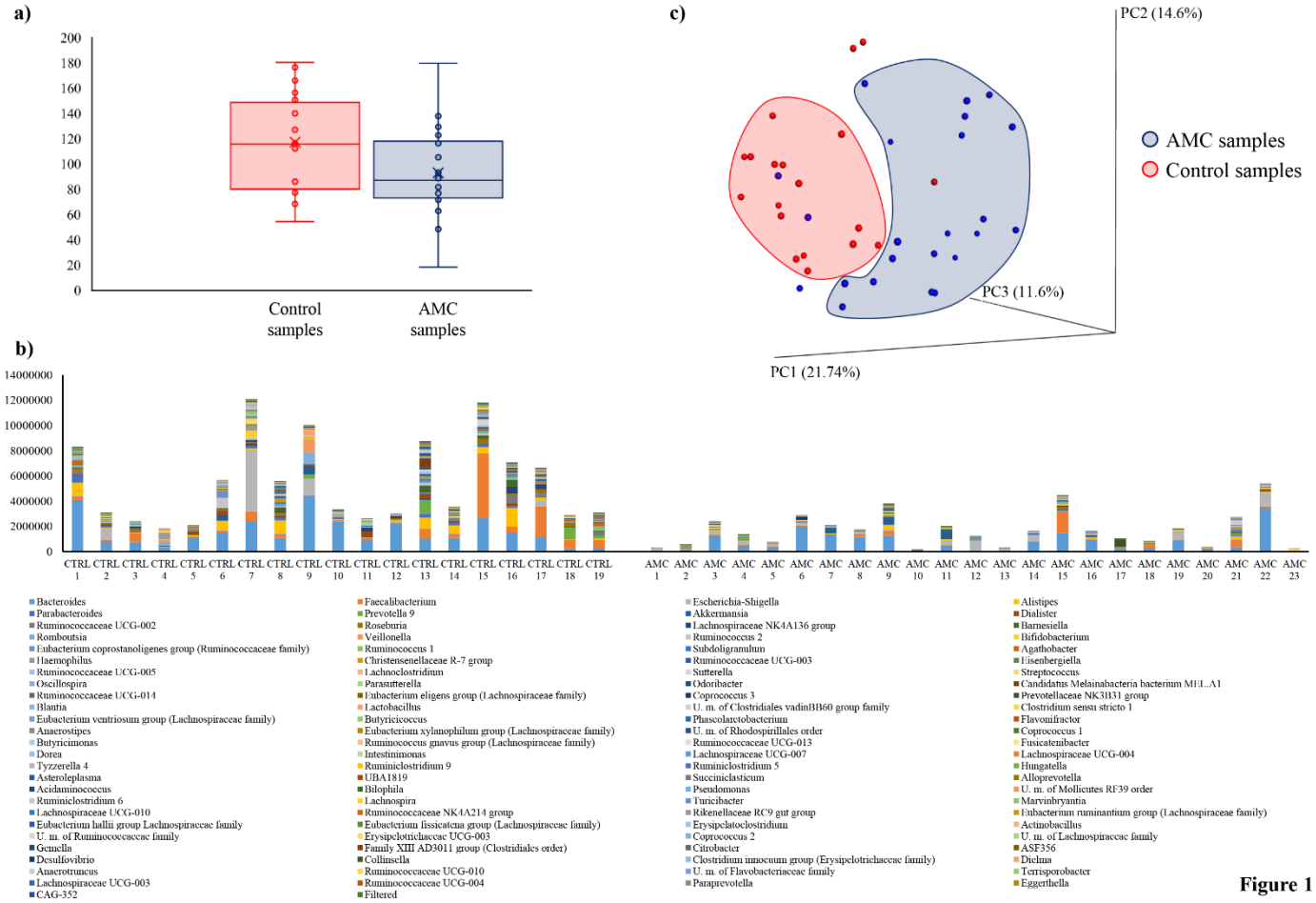


Figure 1. Evaluation of the microbiota composition of AMC and CTRL samples. Panel a shows a Whiskers plot based on observed OTUs identified from AMC and CTRL samples. The *x* axis represents the different groups, while the *y* axis indicates the number of observed OTUs - the boxes represent 50 % of the data set, distributed between the first and the third quartiles. The median divides the boxes into the interquartile range, while "X" represents the mean. Panel b displays the microbiota composition of AMC and Control samples based on 16S rRNA profiling normalized with a quantitative microbiome profiling approach employing flow cytometric enumeration of microbial cells for each sample. The *x* axis represents the different analyzed samples, while the *y* axis indicates normalized OTUs for each sample. Panel c depicts the beta diversity in AMC and Control samples. The predicted PCoA encompassing all 43 analyzed samples are reported and the different clusters are represented by different colors.

Isolation of novel bifidobacterial strains and assessment of AMC susceptibility.

For the purpose of evaluating AMC sensitivity of different members of the *Bifidobacterium* genus, we performed MIC assays involving 237 different strains, including 63 type strains of currently recognized bifidobacterial (sub)species and 174 different strains belonging to four *Bifidobacterium* species, i.e., *Bifidobacterium adolescentis*, *Bifidobacterium bifidum*, *Bifidobacterium breve* and *Bifidobacterium longum* spp. (Table S2), representing common colonizers of the human gut (84, 87, 89, 302, 303). Moreover, in order to isolate bifidobacterial strains with enhanced insensitivity to AMC, we applied a bifidobacterial isolation protocol on fecal samples belonging to individuals subjected to AMC therapy (Table S1). The above-mentioned analysis (178, 179) (see Material & Methods) allowed the isolation of 24 novel strains (Table 1), of which 18 were shown to belong to the *B. longum* species, three to the *B. breve* species and three to the *B. pseudocatenulatum* species. All tested strains showed a unimodal distribution of MIC_{AMC} breakpoint values, ranging from 0.125 μg/mL to 32 μg/mL (Fig. S1 and Table S2). 96.5 % of the strains exhibited a MIC ≤ 1 μg/mL, five strains displayed a breakpoint value equal to 2 μg/mL and, notably, four bifidobacterial strains showed a MIC equal to or higher than 4 μg/mL (Fig. S1 and Table S2). Interestingly, the newly isolated strain *B. breve* 1891B displayed the highest MIC_{AMC} value, i.e. 32 μg/mL (Table 1), which was 11-fold higher than the average bifidobacterial MIC_{AMC}. Moreover, *B. breve* strain M1D showed a MIC value of 16 μg/mL, which was five-fold higher compared to the *Bifidobacterium* MIC average (Table 1). This antibiotic insensitivity is unusual for *B. breve* species, which in general exhibit a high sensitivity to various antibiotics, including AMC (232). Furthermore, two isolates belonging to *B. longum* species, i.e., *B. longum* subsp. *longum* 1898B and *B. longum* subsp. *longum* 39B, were shown to exhibit MIC values of 8 μg/mL and 4 μg/mL (Table 1), being eight- and four-fold higher, respectively, compared to the average MIC value of other *B. longum* strains. These findings suggest that different levels of resistance/insensitivity to AMC reflect strain-specific characteristics rather than a species-specific feature. Notably, the MIC_{AMC} breakpoints identified in 98.5 % of the bifidobacterial strains tested were very low compared to MIC_{AMC} values previously identified for

other members of the human gut microbiota such as *Escherichia coli*, *Citrobacter* spp., *Bacteroides* spp. and *Parabacteroides* spp., which showed MIC_{AMC} breakpoint values higher than 8 µg/mL (304, 305). Interestingly, these data indicate that AMC resistance/insensitivity in bifidobacteria does not appear to follow a vertical route of evolution but may have been acquired through Horizontal Gene Transfer (HGT), in a similar way to that described for other gut commensal microorganisms (306). Finally, we evaluated the susceptibility to the antibiotic amoxicillin alone for the four strains of *Bifidobacterium* insensitive to AMC. Interestingly, each strain showed higher MIC values, i.e. *B. breve* 1891B showed a MIC value of 64 µg/mL, *B. breve* M1D strain exhibited a MIC breakpoint of 32 µg/mL, *B. longum* subsp. *longum* 1898B and *B. longum* subsp. *longum* 39B presented MIC values of 8 µg/mL. These data confirm the resistance of these strains to amoxicillin, both as the sole antibiotic and in combination with clavulanic acid.

Comparative genomics and identification of putative resistance genes in AMC-insensitive *Bifidobacterium* strains.

In order to identify the genetic features that are involved in AMC insensitivity in the four identified bifidobacterial strains, we sequenced, annotated and *in silico* analyzed the genomes of *B. breve* 1891B, *B. breve* M1D, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B, which were shown to elicit the highest MIC_{AMC} (Table 1). The coverage depth of these four newly isolated *Bifidobacterium* chromosomes ranged from 72- to 388-fold, which upon assembly generated one to 25 contigs (Table 2). The number of predicted ORFs ranged from 1842 of *B. longum* subsp. *longum* 39B to 2102 of *B. breve* 1891B (Table 2). We then evaluated the genetic similarities of *B. breve* 1891B, *B. breve* M1D, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B with other chromosome sequences that are publicly available for *B. breve* and *B. longum* subsp. *longum* (sub)species through a comparative genomics analysis. Specifically, the genome sequences of *B. breve* 1891B, *B. breve* M1D, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B were compared with the chromosome sequences of six *B. breve* and five *B. longum* subsp. *longum*

strains, including AMC insensitive strains (Table S3). *In silico* analysis revealed that 1272 genes are commonly shared among the analyzed *B. breve* strains, representing the core genome of this taxon (Fig. 2), whereas the core genome of the assessed *B. longum* subsp. *longum* strains was shown to be composed of 1106 genes. Moreover, there are variable numbers of truly unique genes (TUGs) in either of the two investigated species, ranging between 175 for *B. breve* UCC2003 and 36 for *B. breve* 12L, and ranging from 585 for *B. longum* subsp. *longum* 1897B and 113 for *B. longum* subsp. *longum* 39B (Fig. 2 and Table S3). Interestingly, this analysis revealed the presence of 323 accessory genes shared between *B. breve* AMC-insensitive strains, i.e. *B. breve* 1891B and *B. breve* M1D, and absent in AMC-sensitive *B. breve* strains (Fig. 2). Furthermore, the genomes of the AMC-resistant strains *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B were shown to contain 25 genes that are shared between them while being absent in other analyzed *B. longum* strains (Fig. 2). In addition, the proteome of each of the four identified AMC-insensitive *Bifidobacterium* strains was screened for putative Antibiotic Resistance (AR) genes, using the MEGAREs database (227). This *in silico* analysis showed that the predicted resistome of *B. breve* 1891B, *B. breve* M1D, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B ranged from 167 genes in the case of *B. breve* M1D to 175 genes for *B. longum* subsp. *longum* 1898B (Fig. 2). We included in our analyses both sequences coding for antibiotic-removing transporters and sequences specifying antibiotic-inactivating enzymes. The number of predicted antibiotic-removing transporters ranged from 109 for *B. longum* subsp. *longum* 39B to 127 genes for *B. breve* 1891B and *B. breve* M1D (Fig. 2). Interestingly, the analyzed strains were shown to harbor genes putatively encoding beta-lactamases as part of their resistome; *B. longum* strains were predicted to encode five distinct beta-lactamases, whereas the assessed *B. breve* strains were predicted to specify four beta-lactamases. However, since these putative beta-lactamases are conserved among different strains of the same species, it appears that they are not involved in the observed AMC insensitivity phenotype observed for some strains.

Moreover, further analysis of the *in silico* resistome data and comparison between genes found in the insensitive strains but absent from the sensitive strains allowed the identification 11 predicted AR

genes that are present only in the *B. breve* AMC-insensitive strains. Eight of these identified genes were predicted to encode transporters, whereas two genes seemed to encode enzymes that provide protection against glycopeptide antibiotics (307). Finally, one gene was predicted to encode a glycosyl hydrolase involved in the modification of the lipopolysaccharide (LPS) core and lipid A region with ethanolamine and add amino-arabinose to the 4′ phosphate of lipid A (308). Furthermore, carrying out the same analysis for members of the *B. longum* subsp. *longum* taxon, we identified just one gene uniquely shared between AMC-resistant *B. longum* subsp. *longum* strains, predicted to encode a putative AR transporter. These identified genes are indeed credible genetic candidates responsible for the high MIC_{AMC} breakpoint values identified in *B. breve* 1891B, *B. breve* M1D, *B. longum* subsp. *longum* 39B and *B. longum* subsp. *longum* 1898B.

Table 1. *Bifidobacterium* strains isolated from AMC samples and corresponding MIC_{AMC} values.

Species	Strain	MIC_{AMC} (µg/mL)
Bifidobacterium breve	1891B	32 µg/mL
Bifidobacterium breve	M1D	16 µg/mL
Bifidobacterium breve	M5B	1 µg/mL
Bifidobacterium longum subsp. *longum*	1898B	8 µg/mL
Bifidobacterium longum	2195B	1 µg/mL
Bifidobacterium longum	2196B	1 µg/mL
Bifidobacterium longum	2197B	1 µg/mL
Bifidobacterium longum	2198B	0.5 µg/mL
Bifidobacterium longum	2199B	1 µg/mL
Bifidobacterium longum	2202B	0.125 µg/mL
Bifidobacterium longum subsp. *longum*	39B	4 µg/mL
Bifidobacterium longum	AD12C	1 µg/mL
Bifidobacterium longum	E2C	1 µg/mL
Bifidobacterium longum	E4F	2 µg/mL
Bifidobacterium longum	E6H	1 µg/mL
Bifidobacterium longum	F2A	0.25 µg/mL
Bifidobacterium longum	G7G	0.25 µg/mL
Bifidobacterium longum	G8F	0.5 µg/mL
Bifidobacterium longum	L5G	2 µg/mL
Bifidobacterium longum	MISS1F	0.25 µg/mL
Bifidobacterium longum	T3	2 µg/mL
Bifidobacterium pseudocatenulatum	AN3D	2 µg/mL
Bifidobacterium pseudocatenulatum	L3G	0.5 µg/mL
Bifidobacterium pseudocatenulatum	M8H	2 µg/mL

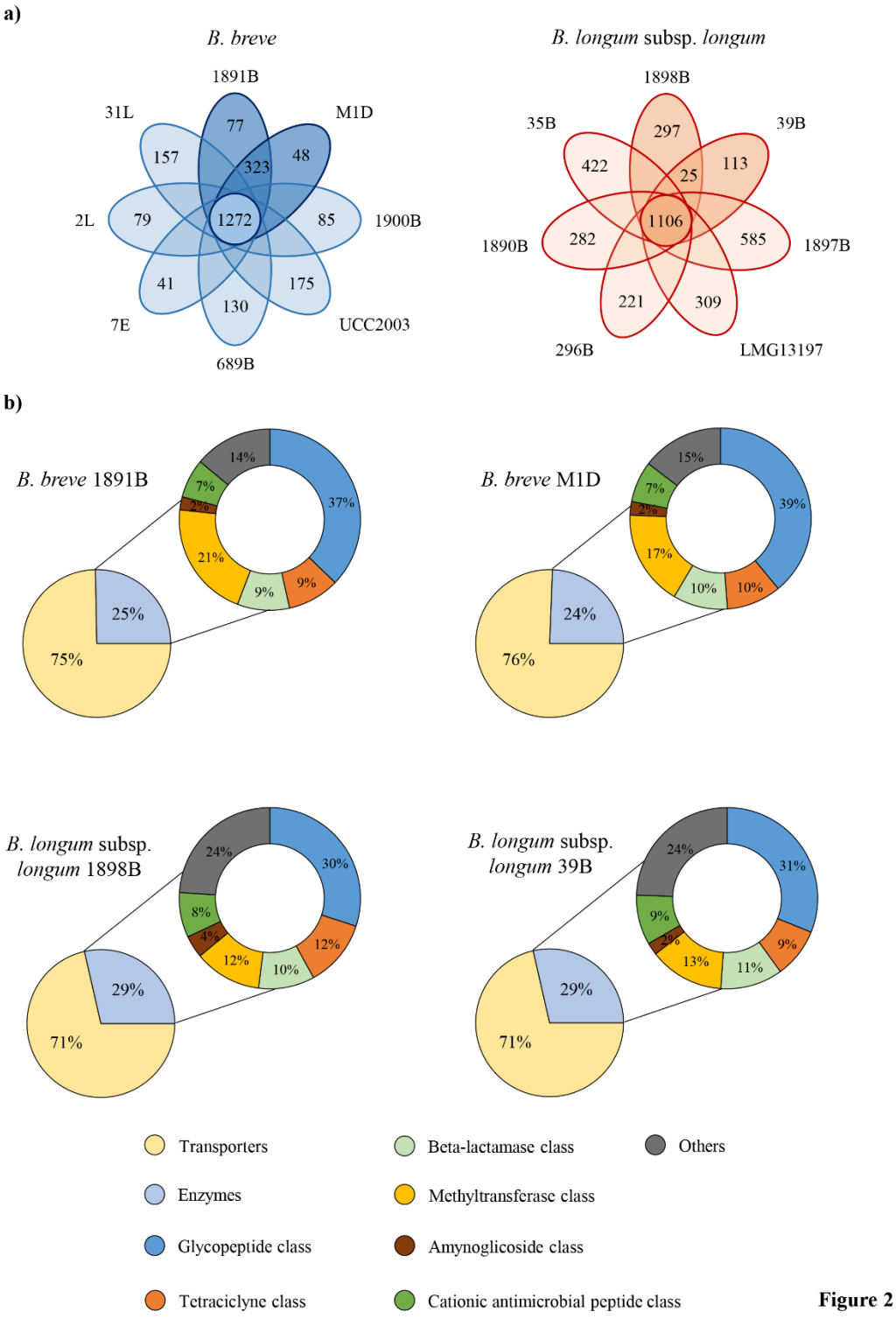


Figure 2. Comparative genomic and resistome analyses of AMC-resistant *Bifidobacterium* strains. Panel a shows two Venn diagrams. The numbers in the central circle represent the core genome of the *B. breve* and *B. longum* subsp. *longum* strains analyzed. The number in the insertion section represents the shared genes between AMC-resistant strains, while the number in the oval circles depict the TUGs for each strain. Panel b exhibits the predicted resistome of the AMC-resistant *Bifidobacterium* strains.

Table 2. Genetic features of AMC-resistant *Bifidobacterium* strains.

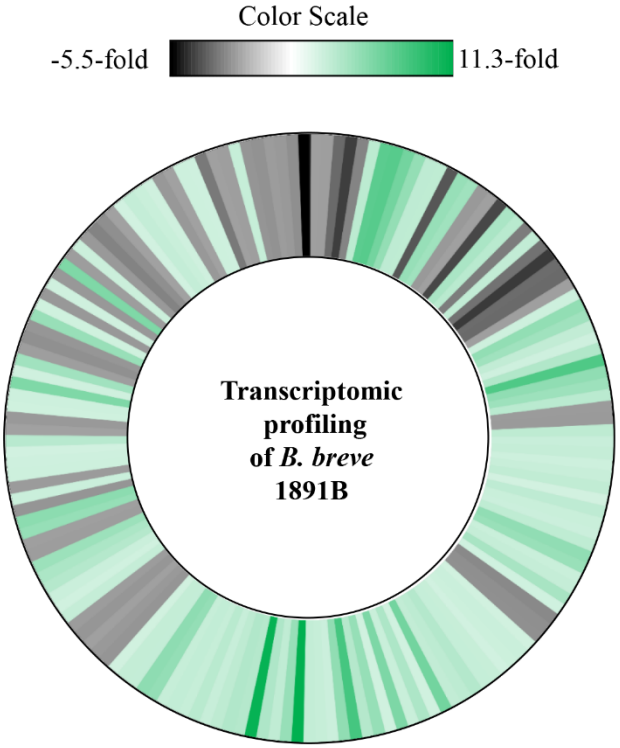
Strains	Genome Lenght (Bp)	Nº of contigs	Nº of ORFs	Nº of tRNA	Nº of rRNA Loci	Accession number
B. breve 1891B	2427222	1	2102	54	3	JACZEM000000000
B. breve M1D	2421612	7	2053	54	3	JACZEL000000000
B. longum subsp. *longum* 1898B	2464874	9	2065	56	4	JACZEK000000000
B. longum subsp. *longum* 39B	2287000	25	1842	56	1	JACZEJ000000000

Transcriptomic analysis of *Bifidobacterium breve* 1891B.

In order to investigate if and how the presence of AMC in the growth medium modulates the transcription of genes putatively involved in the AR described above, we explored the transcriptome of *B. breve* 1891B when cultivated in the presence or absence of AMC in MRS medium. This analysis showed significant modulation of 163 genes (> 2-fold induction, p-value ≤ 0.05), 110 of which were upregulated in the MRS_{AMC} medium, while 53 genes were downregulated when compared to the reference condition (Fig. 3). Functional assignment of the upregulated and downregulated transcribed genes based on the eggNOG database (187) revealed that the overexpressed transcripts referred to genes involved in defensive mechanisms and in the metabolism and transport of inorganic ions (20.2 % and 9.2 %, respectively) (Fig. 3). Conversely, downregulated transcripts concerned genes involved in the transport and metabolism of carbohydrates and nucleotides (13.7 % and 25.5 %, respectively). Interestingly, among the upregulated genes, one particular ORF, associated with locus tag 1891B_1181 and encoding a predicted ABC transporter, exhibited a 3.6-fold increased transcription (p-value < 0.001) (Fig. 3). This gene was previously characterized in this study in the resistome of strain *B. breve* 1891B, shared with other *B. breve* AMC resistant strains and not present in *B. breve* AMC sensitive strains (Table 1). These findings suggest that this ABC transporter-encoding gene is involved in the observed insensitivity of *B. breve* 1891B to the AMC antibiotic. In addition, two TUGs observed in the genome of the strain *B. breve* 1891B, i.e., ORFs 1891B_1167 and 1891B_1282, were shown to be transcriptionally induced in the presence of AMC, exhibiting an upregulation of 3.4-fold and 2.7-fold, respectively (p-value < 0.05) (Fig. 3), and in both cases encoding hypothetical

proteins. These data indicate that these unique genes are involved in the high resistance of *B. breve* 1891B toward AMC antibiotic.

a)



b)

Color Scale

2-fold — 11.3-fold

Genes	EggNOG function	Function	Fold	*p*-value
1891B_0035	G	ABC superfamily ATP binding cassette transporter solute-binding protein		0.005
1891B_0094	S	cell filamentation protein Fic		0.003
1891B_0099	S	hemolysin III		0.008
1891B_0200	S	hypothetical protein		0.01
1891B_0212	S	phage infection protein		0.023
1891B_0266	V	Putative manganese efflux pump		0.01
1891B_0323	V	toxin		0.008
1891B_0393	R	Plasmid stability protein		0.035
1891B_0412	P	Permease of ABC transporter system		0.05
1891B_0463	L	Transposase		0.02
1891B_0524	J	RNA pseudouridine synthase		0.004
1891B_0554	K	Fe-S cluster assembly protein HesB		0.041
1891B_0590	S	hypothetical protein		0.047
1891B_0599	E	diaminopimelate aminotransferase		0
1891B_0602	V	Transporter, Sodium/bile acid symporter family, macrolide-resistance protein		0.05
1891B_0621	S	hypothetical protein		0.035
1891B_0636	T	diguanylate cyclase		0.022
1891B_0676	S	Arylsulfate sulfotransferase		0.012
1891B_0701	H	thiamine biosynthesis protein ThiS		0.02
1891B_0712	S	hypothetical protein		0.045
1891B_0749	S	peptidase		0.005
1891B_0858	S	hypothetical protein		0.05
1891B_0884	P	calcium-transporting ATPase		0.039
1891B_0949	J	rRNA methyltransferase		0.001
1891B_0991	S	hypothetical protein		0.05
1891B_0996	R	transporter		0.029
1891B_1000	I	lipase		0.028
1891B_1002	R	phosphoesterase		0.007
1891B_1003	K	TenA family transcriptional regulator		0.01
1891B_1004	S	hypothetical protein		0.05
1891B_1059	S	phage tail tape measure protein		0.004
1891B_1070	S	Phage portal protein gp6-like protein		0.01
1891B_1073	S	hypothetical protein		0.018
1891B_1074	S	phage holin		0.004
1891B_1125	R	COF family hydrolase		0.05
1891B_1130	S	electron transporter		0.01
1891B_1131	R	hydrolase or phosphatase		0.001
1891B_1132	S	hypothetical protein		0.001
1891B_1149	V	Permease		0.007
1891B_1155	V	DNA-binding protein		0.007
1891B_1159	T	histidine protein kinase		0.031
1891B_1167	S	hypothetical protein		0.039
1891B_1168	K	RNA polymerase sigma24 factor		0.05
1891B_1176	V	cytochrome C oxidase subunit IV		0.003
1891B_1181	V	ABC transporter		0
1891B_1187	P	transporter		0.047
1891B_1190	V	Transcriptional regulator, MarR family		0.009
1891B_1216	S	hypothetical protein		0.039
1891B_1219	S	hypothetical protein		0.003
1891B_1282	S	hypothetical protein		0.049
1891B_1300	S	DNA-binding protein		0.008
1891B_1308	L	transposase		0.044
1891B_1314	S	Transposase subunit B		0.037
1891B_1336	V	lipolytic protein G-D-S-L family		0.019
1891B_1338	S	hypothetical protein		0.037
1891B_1339	L	DEAD/DEAH box helicase		0.034
1891B_1398	S	spermidine synthase		0.025
1891B_1410	V	Cell wall biosynthesis-associated protein		0.002
1891B_1421	J	D-tyrosyl-tRNA(Tyr) deacylase		0.049
1891B_1433	S	division/cell wall cluster transcriptional repressor MraZ		0.027
1891B_1453	C	acylphosphatase		0.032
1891B_1455	S	alpha-galactosidase		0.029
1891B_1456	G	alpha-L-fucosidase		0.02
1891B_1466	V	DNA-binding protein		0.006
1891B_1485	K	LacI family transcriptional regulator		0.01
1891B_1486	G	oligosaccharide:H+ symporter		0.034
1891B_1487	G	beta-fructofuranosidase		0.016
1891B_1495	V	glycerophosphodiester phosphodiesterase		0.045
1891B_1504	K	Transcriptional regulator, GntR family		0.009
1891B_1506	I	PAP2 family phosphoesterase		0.002
1891B_1568	K	Periplasmic binding protein/LacI transcriptional regulator		0.01
1891B_1581	R	nucleotidyltransferase		0.05
1891B_1594	S	nucleoside deoxyribosyltransferase		0.017
1891B_1658	S	hypothetical protein		0.05
1891B_1662	V	chromosome partitioning protein ParA		0.016
1891B_1663	V	transcriptional regulator		0.023
1891B_1667	S	ATPase for DNA repair		0.009
1891B_1670	V	Lanthionine synthetase C family protein		0.013
1891B_1695	R	ABC transporter		0.013
1891B_1696	R	GNAT family N-acetyltransferase		0.037
1891B_1700	V	excisionase family DNA-binding domain-containing protein		0.018
1891B_1702	V	MFS transporter		0.002
1891B_1712	S	hypothetical protein		0.018
1891B_1723	V	ABC transporter permease		0.001
1891B_1724	V	ABC transporter ATP-binding protein		0.01
1891B_1725	V	multidrug ABC transporter ATP-binding protein		0.004
1891B_1726	E	Branched-chain amino acid permease AzlC		0.002
1891B_1853	P	phosphate ABC transporter ATP-binding protein		0.05
1891B_1854	P	phosphate ABC transporter substrate-binding protein		0.05
1891B_1855	P	phosphate ABC transporter permease		0.021
1891B_1856	P	phosphate ABC transporter substrate-binding protein		0.024
1891B_1858	T	Phosphate regulon sensor protein PhoR		0.018
1891B_1874	G	kinase, PfkB family		0.009
1891B_1878	V	translation repressor RelB		0.05
1891B_1881	S	alpha/beta hydrolase		0
1891B_1883	R	oxidoreductase		0.036
1891B_1893	S	hypothetical protein		0.034
1891B_1913	S	hypothetical protein		0.034
1891B_1933	J	23S rRNA N-6-methyltransferase		0.004
1891B_1934	L	transposase		0.01
1891B_2008	V	ATPase		0.014
1891B_2025	S	hypothetical protein		0.021
1891B_2040	P	ABC transporter permease		0.038
1891B_2042	G	ABC transporter substrate-binding protein		0.035
1891B_2051	G	alpha-amylase		0.006
1891B_2052	P	ABC transporter permease		0.001
1891B_2053	P	sugar ABC transporter permease		0.005
1891B_2054	G	ABC transporter substrate-binding protein		0.032
1891B_2056	G	alpha-galactosidase		0

Figure 3

Figure 3. Transcriptional modulation of *B. breve* 1819B genes in presence of AMC. Panel a shows the statistically significant differential transcription observed for *B. breve* 1891B when growth in MRS medium with an addition of 32 µg/mL of AMC when compared to the reference condition. Panel b displays the subset of significantly upregulated encoding genes. Red boxes highlight the ORFs 1891B_1181, ORFs 1891B_1167 and 1891B_1282, which are putatively involved in the AMC resistance mechanism. Color legends at the top of the images indicate increased transcription levels (green) compared to those of the reference samples. The EggNOG letter for each significant up-regulated gene is reported. Each letter stands for the following function: **S**, Function unknown, **L**, Replication, recombination and repair, **R**, General function prediction only, **G**, Carbohydrate transport and metabolism, **J**, Translation, ribosomal structure and biogenesis, **E**, Amino acid transport and metabolism, **K**, Transcription, **P**, Inorganic ion transport and metabolism, **C**, Energy production and conversion, **V**, Defense mechanisms, **T**, Signal transduction mechanisms, H, Coenzyme transport and metabolism, **I**, Lipid transport and metabolism.

Effect of resistant *Bifidobacterium* strains on the human gut microbiota in the presence of AMC.

The GIT microbiota composition is known to be influenced by antibiotic compounds (293, 300). In particular, antibiotic exposure will alter the microbial composition in the infant gut community (301, 309). In order to evaluate the effects of AMC on the human gut microbiota in the presence of AMC-insensitive *Bifidobacterium* strains retrieved in this study, we performed four different co-culture experiments, in which each AMC-insensitive bifidobacterial strain was cultivated with a fresh fecal sample and in the presence or absence of the AMC antibiotic (20 µM) in the growth medium (see Materials & Methods) (293). The co-cultures were monitored at four different time points, i.e. 12h, 18h, 24h and 36h after the inoculum. For each time point and for each cultivation condition tested, the changes in the microbiota composition were assessed by shallow-shotgun metagenomics analysis. The analyses resulted in a total of 2,603,786 reads with an average of 63,507 ± 23,113 reads per sample (Table S4). The analysis of the species richness, i.e. the number of species identified in each sample, revealed a significant decrease in complexity in the AMC-containing sample group (average of 414 ± 58) compared to the untreated control group (average of 474 ± 34) (p-value < 0.05) (Fig. S2). In order to obtain a comprehensive biological interpretation of the analyzed batch-culture microbiome complexity, we performed a quantitative microbiome profiling experiment based on flow

cytometric analyses for the enumeration of microbial cells present in each co-culture condition at each time point (287). Interestingly, comparison between each co-culture experiment with its own control revealed a decrease of the number of microbial cells in 85% of the samples to which AMC was added (Fig. 4). Moreover, statistical analysis revealed a significant decrease (p-value < 0.05) of 1.35 fold in the number of cells in the AMC treated group (average of 1.97E+05 ± 6.11E+04) compared to the control sample group (average of 2.65E+05 ± 9.41E+04).

Focusing on the absolute abundance of the species corresponding to the identified AMC-resistant *Bifidobacterium* strains, the only strain that seemed to increase in abundance in the presence or absence of AMC was *B. breve* 1891B (Fig. 4). In fact, the fecal samples co-cultivated with 1891B revealed a high abundance of *B. breve* species in both untreated and treated conditions, with an average abundance of 24.56 % ± 8.90 % and 34.08 % ± 6.32 %, respectively. Moreover, alignment analysis based on Bowtie2 software revealed that all these reads classified as *B. breve* align to the *Bifidobacterium breve* 1891B reference genome (alignment identity of 99 %). Intriguingly, in both conditions in which we inoculated the fecal sample with *B. breve* 1891B, this strain appeared to outcompete other bacteria present in the co-cultivation set-up, already 12h after the inoculum (Fig. 4). In contrast, the other strains tested displayed low competitive performances even in the absence of AMC (Fig. 4 and Table S5).

Interestingly, the fecal samples inoculated with *B. breve* 1891B and treated with AMC showed a higher abundance of *Bifidobacterium* species (average abundance of 7.72 % ± 2.78 % evaluated excluding *B. breve*) compared to other AMC co-culture experiments (total average abundance of 0.89 % ± 0.48 % evaluated excluding the inoculated strain). Therefore, these results support the apparent ability of *B. breve* 1891B to grow in the presence of AMC antibiotic, making this strain an interesting candidate for the development of *Bifidobacterium*-containing probiotic products.

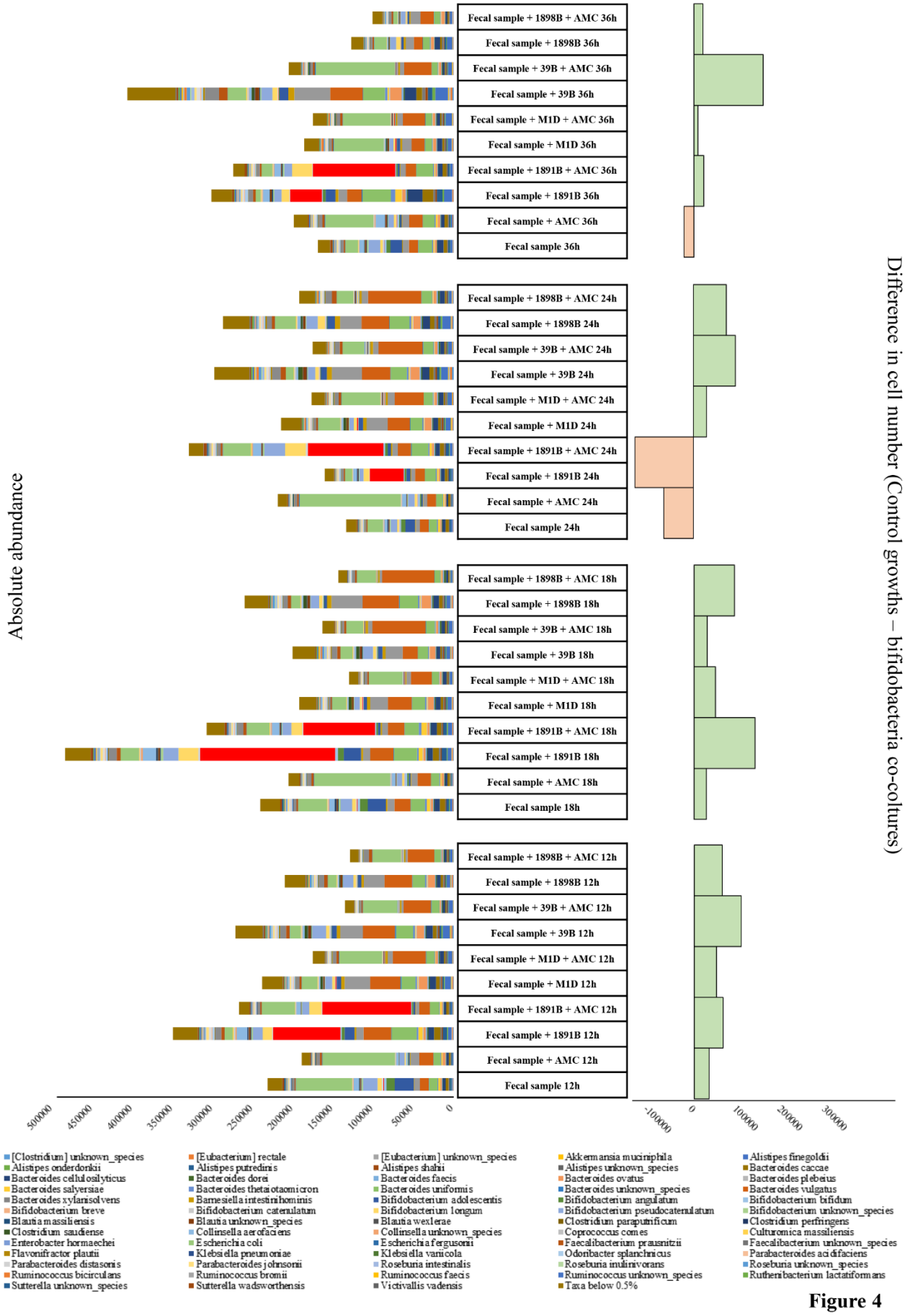

Absolute abundance
Difference in cell number (Control growths – bifidobacteria co-coltures)
Fecal sample + 1898B + AMC 36h
Fecal sample + 1898B 36h
Fecal sample + 39B + AMC 36h
Fecal sample + 39B 36h
Fecal sample + M1D + AMC 36h
Fecal sample + M1D 36h
Fecal sample + 1891B + AMC 36h
Fecal sample + 1891B 36h
Fecal sample + AMC 36h
Fecal sample 36h
Fecal sample + 1898B + AMC 24h
Fecal sample + 1898B 24h
Fecal sample + 39B + AMC 24h
Fecal sample + 39B 24h
Fecal sample + M1D + AMC 24h
Fecal sample + M1D 24h
Fecal sample + 1891B + AMC 24h
Fecal sample + 1891B 24h
Fecal sample + AMC 24h
Fecal sample 24h
Fecal sample + 1898B + AMC 18h
Fecal sample + 1898B 18h
Fecal sample + 39B + AMC 18h
Fecal sample + 39B 18h
Fecal sample + M1D + AMC 18h
Fecal sample + M1D 18h
Fecal sample + 1891B + AMC 18h
Fecal sample + 1891B 18h
Fecal sample + AMC 18h
Fecal sample 18h
Fecal sample + 1898B + AMC 12h
Fecal sample + 1898B 12h
Fecal sample + 39B + AMC 12h
Fecal sample + 39B 12h
Fecal sample + M1D + AMC 12h
Fecal sample + M1D 12h
Fecal sample + 1891B + AMC 12h
Fecal sample + 1891B 12h
Fecal sample + AMC 12h
Fecal sample 12h
500000
450000
400000
350000
300000
250000
200000
150000
100000
50000
0
-100000
0
100000
200000
300000
[Clostridium] unknown_species
[Eubacterium] rectale
[Eubacterium] unknown_species
Akkermansia muciniphila
Alistipes finegoldii
Alistipes onderdonkii
Alistipes putredinis
Alistipes shahii
Alistipes unknown_species
Bacteroides caccae
Bacteroides cellulosilyticus
Bacteroides dorei
Bacteroides faecis
Bacteroides ovatus
Bacteroides plebeius
Bacteroides salyersiae
Bacteroides thetaiotaomicron
Bacteroides uniformis
Bacteroides unknown_species
Bacteroides vulgatus
Bacteroides xylanisolvens
Barnesiella intestinihominis
Bifidobacterium adolescentis
Bifidobacterium angulatum
Bifidobacterium bifidum
Bifidobacterium breve
Bifidobacterium catenulatum
Bifidobacterium longum
Bifidobacterium pseudocatenulatum
Bifidobacterium unknown_species
Blautia massiliensis
Blautia unknown_species
Blautia wexlerae
Clostridium paraputrificum
Clostridium perfringens
Clostridium saudiense
Collinsella aerofaciens
Collinsella unknown_species
Coprococcus comes
Culturomica massiliensis
Enterobacter hormaechei
Escherichia coli
Escherichia fergusonii
Faecalibacterium prausnitzii
Faecalibacterium unknown_species
Flavonifractor plautii
Klebsiella pneumoniae
Klebsiella variicola
Odoribacter splanchnicus
Parabacteroides acidifaciens
Parabacteroides distasonis
Parabacteroides johnsonii
Roseburia intestinalis
Roseburia inulinivorans
Roseburia unknown_species
Ruminococcus bicirculans
Ruminococcus bromii
Ruminococcus faecis
Ruminococcus unknown_species
Ruthenibacterium lactatiformans
Sutterella unknown_species
Sutterella wadsworthensis
Victivallis vadensis
Taxa below 0.5%

Figure 4

Figure 4. Growth experiments in fecal media.

The histogram on the left displays the taxonomic absolute abundance of the co-culture experiments. The absolute abundance of *B. breve* 1891B is indicated with the color red in the bar plot. The histogram on the right shows a comparison between the absolute abundance of the co-culture experiments in MRS_{AMC} and the reference condition.

CONCLUSIONS

Many pharmaceuticals are well known to influence the human gut microbiota. Since the combination of amoxicillin and clavulanic acid is one of the most frequently prescribed antibiotics in the Western world, especially during infancy and adolescence, we decided to explore the impact of this antibiotic formulation on the gut microbiota of 23 children and compare this to a control group who had not been treated with AMC. Interestingly, results indicated a drastic reduction of bacteria in AMC-treated children when compared to the CTRL group, including a reduction in absolute abundance of the *Bifidobacterium* genus. We decided to perform MIC experiments of several strains belonging to our collection and novel strains that had been isolated from children treated with AMC. The determination of sensitivity/resistance of these intestinal *Bifidobacterium* strains showed that 98.5 % of them are sensitive to this antibiotic. Nonetheless, we isolated four strains that showed higher resistance, of which *B. breve* 1891B strain displayed the highest MIC_{AMC} value. Transcriptional analyses of this strain, growth in MRS added with AMC antibiotic, highlighted one interesting gene that was shown to be considerably upregulated compared to the reference condition and belonging to the predicted resistome, suggesting its involvement in the AMC resistance of *B. breve*. Finally, by simulating gut-microbiota experiments it emerged that the 1891B strain is able to survive and compete in the presence of a complex microbiota combined with AMC antibiotic, opening up the possibility of using this strain in a probiotic product when AMC therapy is prescribed.

SUPPLEMENTAL MATERIALS

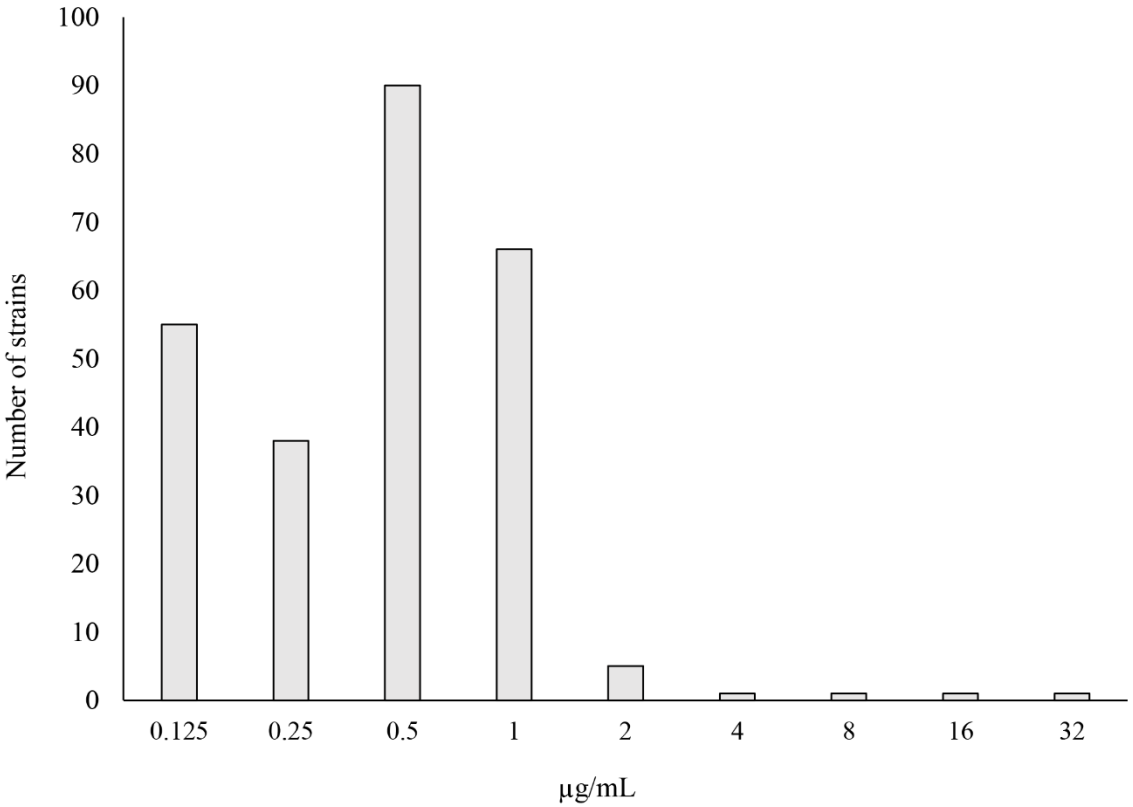


Figure S1

Figure S1. AMC breakpoints values of Bifidobacterium genus.
Distribution of breakpoint values for AMC in 261 different Bifidobacterium strains tested. The x axis reports the antibiotic concentrations (microgram per milliliter), while the y axis represents the number of strains.

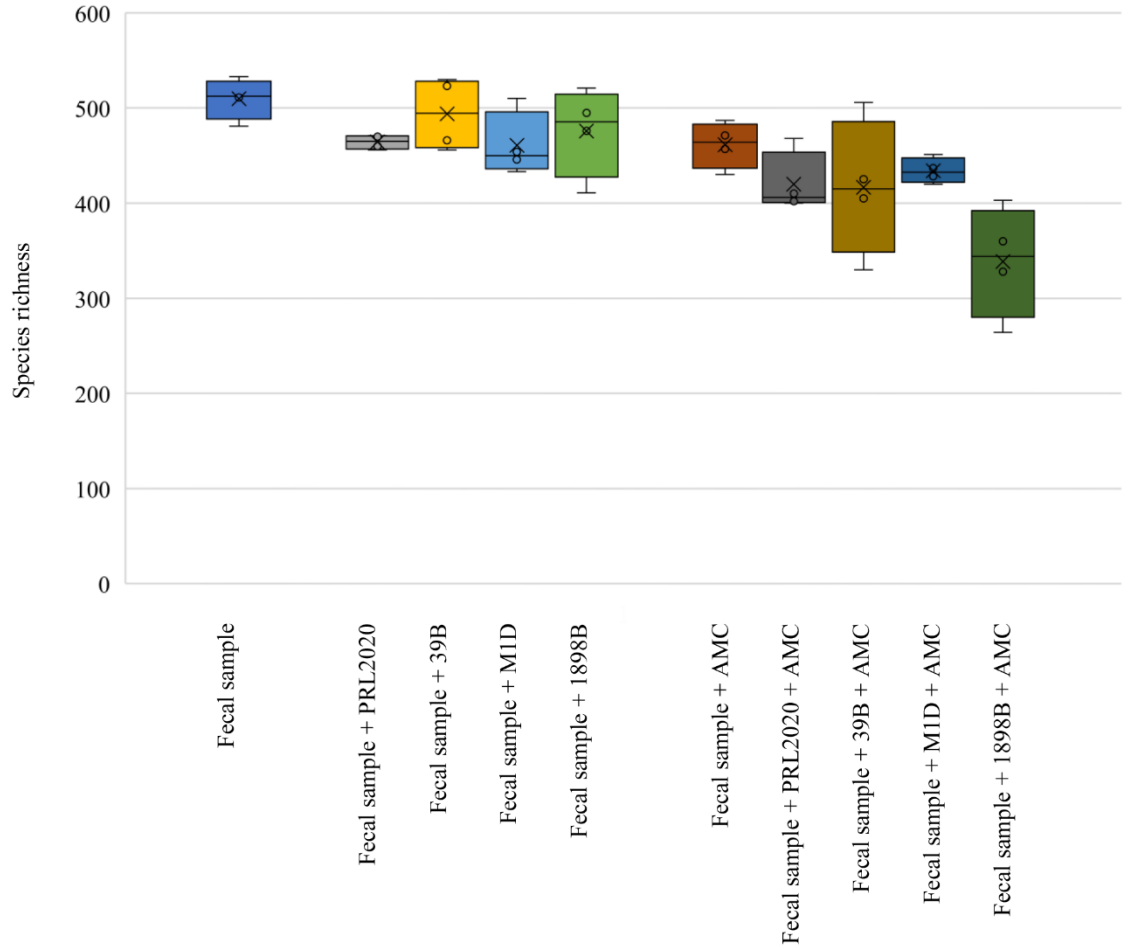


Figure S2

Figure S2. Examination of the species richness of the co-culture experiments.
The species richness is represented by the box and whisker plot calculated through the number of species identified in each samples. The bottom and top of the box represent the first and third quartiles, and the band inside the box is the median. Moreover, the ends of the whiskers represent the minimum and maximum of all the data of the sample.

Table S1. Samples used in this study.

AMC group			Control group	
Samples	Age	Days of AMC treatment	Samples	Age
AMC 1	5 Years	10 days	CTRL 1	8 Years
AMC 2	5 Years	10 days	CTRL 2	4 Years
AMC 3	5 Years	10 days	CTRL 3	4 Years
AMC 4	7 Years	10 days	CTRL 4	1 Year
AMC 5	6 Years	10 days	CTRL 5	10 Years
AMC 6	5 Years	10 days	CTRL 6	4 Years
AMC 7	11 Years	10 days	CTRL 7	1 Year
AMC 8	6 Years	10 days	CTRL 8	6 Years
AMC 9	4 Years	8 days	CTRL 9	1 Year
AMC 10	12 Years	8 days	CTRL 10	4 Years
AMC 11	9 Years	10 days	CTRL 11	7 Years
AMC 12	4 Years	10 days	CTRL 12	1 Year
AMC 13	4 Years	7 days	CTRL 13	3 Years
AMC 14	3 Years	7 days	CTRL 14	2 Years
AMC 15	4 Years	7 days	CTRL 15	4 Years
AMC 16	11 Years	10 days	CTRL 16	4 Years
AMC 17	6 Years	10 days	CTRL 17	6 Years
AMC 18	9 Years	8 days	CTRL 18	3 Years
AMC 19	7 Years	10 days	CTRL 19	8 Years
AMC 20	4 Years	7 days		
AMC 21	12 Years	8 days		
AMC 22	3 Years	8 days		
AMC 23	1 Year	7 days		

Table S2. MIC breackpoint values for members of the *Bifidobacterium* genus

Species	Strains	MIC (µg/mL)
B. acticoloniiforme	DSM22766	0,5
B. adolescentis	125B	0,5
B. adolescentis	14B	0,125
B. adolescentis	153B	0,25
B. adolescentis	1892B	0,125
B. adolescentis	22L	0,125
B. adolescentis	235B	0,125
B. adolescentis	236B	0,125
B. adolescentis	369B	0,125
B. adolescentis	382B	0,5
B. adolescentis	388B	0,5
B. adolescentis	42B	0,25
B. adolescentis	487B	0,125
B. adolescentis	50B	0,125
B. adolescentis	532B	0,25
B. adolescentis	537B	0,25
B. adolescentis	53B	0,25
B. adolescentis	55B	0,125
B. adolescentis	56B	0,5
B. adolescentis	57B	0,125
B. adolescentis	58B	0,125
B. adolescentis	61B	0,125
B. adolescentis	62B	0,125
B. adolescentis	679B	0,125
B. adolescentis	703B	0,5
B. adolescentis	70B	0,5
B. adolescentis	711B	0,5
B. adolescentis	712B	0,125
B. adolescentis	713B	0,125
B. adolescentis	714B	0,25
B. adolescentis	716B	0,125
B. adolescentis	723B	0,5
B. adolescentis	731B	1
B. adolescentis	734B	1
B. adolescentis	740B	0,125
B. adolescentis	74B	0,125
B. adolescentis	75B	0,125
B. adolescentis	76B	0,25
B. adolescentis	77B	0,125
B. adolescentis	780B	0,125
B. adolescentis	793B	0,125
B. adolescentis	796B	0,25
B. adolescentis	79B	0,25

B. adolescentis	809B	0,25
B. adolescentis	80B	0,5
B. adolescentis	831B	0,125
B. adolescentis	841B	0,5
B. adolescentis	848B	0,5
B. adolescentis	856B	1
B. adolescentis	859B	0,125
B. adolescentis	898B	0,125
B. adolescentis	924B	0,125
B. adolescentis	951B	1
B. adolescentis	952B	0,25
B. adolescentis	954B	0,25
B. adolescentis	970B	0,25
B. adolescentis	971B	0,125
B. adolescentis	972B	1
B. adolescentis	973B	0,125
B. adolescentis	AD2-8	0,125
B. adolescentis	AL12-4	0,25
B. adolescentis	AL46-2	0,25
B. adolescentis	AL46-7	0,25
B. adolescentis	ATCC15703	0,25
B. adolescentis	LMG10733	0,125
B. adolescentis	LMG10734	0,5
B. adolescentis	LMG11579	0,5
B. adolescentis	LMG18897	0,125
B. aesculapii	DSM26737	0,125
B. angulatum	LMG11039	1
B. animalis subsp. *animalis*	LMG10508	0,5
B. animalis subsp. *lactis*	DSM10140	0,5
B. anseris	LMG30189	0,125
B. aquikefiri	LMG28769	0,125
B. asteroides	LMG10735	1
B. biavatii	DSM23969	0,25
B. bifidum	13200	1
B. bifidum	155B	0,125
B. bifidum	156B	0,5
B. bifidum	184B	0,125
B. bifidum	1887B	0,125
B. bifidum	1A	0,125
B. bifidum	2G	0,125
B. bifidum	313B	0,125
B. bifidum	316B	0,125
B. bifidum	323B	1
B. bifidum	324B	0,25
B. bifidum	352B	1

B. bifidum	361B	1
B. bifidum	364B	0,125
B. bifidum	81B	0,125
B. bifidum	85B	0,125
B. bifidum	98B	0,125
B. bifidum	A8	0,5
B. bifidum	IF10/10	0,25
B. bifidum	IF23/2	0,25
B. bifidum	LMG11041	0,5
B. bifidum	LMG11582	1
B. bifidum	LMG11583	1
B. bifidum	LMG13195	1
B. bifidum	PRL2010	0,5
B. bohemicum	DSM22767	1
B. bombi	DSM19703	0,5
B. boum	LMG10736	1
B. breve	158B	1
B. breve	1891B	32
B. breve	1900B	0,25
B. breve	2L	1
B. breve	31L	1
B. breve	675B	1
B. breve	689B	0,125
B. breve	691B	1
B. breve	7E	1
B. breve	M1D	16
B. breve	M5B	1
B. breve	UCC2003	0,5
B. callitrichos	DSM23973	1
B. catenulatum	LMG11043	0,5
B. choerinum	LMG10510	1
B. coryneforme	LMG18911	1
B. criceti	LMG30188	0,25
B. crudilactis	LMG23609	1
B. cuniculi	LMG10738	0,5
B. dentium	LMG11045	0,5
B. eulemuris	DSM100216	0,5
B. gallicum	LMG11596	0,5
B. pullorum subsp. *gallinarum*	LMG11586	0,125
B. hapali	DSM100202	0,25
B. imperatoris	LMG30297	0,125
B. indicum	LMG11587	0,5
B. italicum	LMG30187	0,5
B. catenulatum subsp. *kashiwanohense*	DSM21854	0,5
B. lemurum	DSM28807	0,125

B. longum	103B	0,5
B. longum	105B	0,5
B. longum	107B	1
B. longum	108B	0,5
B. longum	109B	0,5
B. longum	10B	1
B. longum	111B	1
B. longum	114B	0,5
B. longum	115B	1
B. longum	116B	1
B. longum	120B	0,5
B. longum	123B	0,5
B. longum	127B	0,5
B. longum	12A	0,5
B. longum	130B	0,5
B. longum	134B	0,5
B. longum	139B	0,25
B. longum	142B	0,5
B. longum	159B	0,5
B. longum	168B	0,5
B. longum	169B	0,5
B. longum	178B	0,5
B. longum	1890B	0,5
B. longum	1897B	0,5
B. longum	1898B	8
B. longum	18B	1
B. longum	18L	0,5
B. longum	19L	0,5
B. longum	1B	0,5
B. longum	207B	0,5
B. longum	209B	0,125
B. longum	20B	0,5
B. longum	20L	0,5
B. longum	211B	0,5
B. longum	216B	0,5
B. longum	217B	1
B. longum	2195B	1
B. longum	2196B	1
B. longum	2197B	1
B. longum	2198B	0,5
B. longum	2199B	1
B. longum	21B	1
B. longum	21L	0,5
B. longum	2202B	0,125
B. longum	223B	1

B. longum	224B	1
B. longum	227B	1
B. longum	234B	1
B. longum	23B	0,5
B. longum	26B	0,25
B. longum	27B	1
B. longum	27L	0,5
B. longum	28B	0,5
B. longum	28L	0,5
B. longum	296B	1
B. longum	2B	0,5
B. longum	314B	0,5
B. longum	32L	0,5
B. longum	339B	0,5
B. longum	33B	0,25
B. longum	33L	0,5
B. longum	349B	1
B. longum	35B	0,25
B. longum	39B	4
B. longum	3B	0,5
B. longum	40B	0,5
B. longum	411B	1
B. longum	439B	0,125
B. longum	446B	0,25
B. longum	453B	0,5
B. longum	458B	0,5
B. longum	45B	0,125
B. longum	460B	0,5
B. longum	46B	0,5
B. longum	47B	0,5
B. longum	49B	0,5
B. longum	531B	1
B. longum	66B	0,5
B. longum	67B	0,5
B. longum	71B	0,5
B. longum	7F	0,5
B. longum	87B	0,25
B. longum	94B	0,5
B. longum	9B	0,5
B. longum	AD12C	1
B. longum	Bl29	0,25
B. longum	E2C	1
B. longum	E4F	2
B. longum	E6H	1
B. longum	F2A	0,25

B. longum	G7G	0,25
B. longum	G8F	0,5
B. longum	L5G	2
B. longum	MISS1F	0,25
B. longum	T3	2
B. longum subsp. *infantis*	ATCC15697	0,5
B. longum subsp. *longum*	LMG13197	1
B. longum subsp. *suis*	LMG21814	0,5
B. magnum	LMG11591	0,125
B. margollesii	LMG30296	1
B. meryciucm	LMG11341	0,25
B. minimum	LMG11592	0,5
B. mongoliense	DSM21395	1
B. moukabalense	DSM27321	0,5
B. myositis	DSM100196	0,5
B. parmae	LMG30295	0,25
B. pseudocatenulatum	AN3D	2
B. pseudocatenulatum	L3G	0,5
B. pseudocatenulatum	LMG10505	0,5
B. pseudocatenulatum	M8H	2
B. pseudolongum subsp. *globosum*	LMG11659	0,5
B. pseudolongum subsp. *pseudolongum*	LMG11571	1
B. psychraerophilum	LMG21775	0,25
B. pullorum subsp. *pullorum*	DSM20433	0,5
B. reuteri	DSM23975	1
B. ruminantium	LMG21811	1
B. pullorum subsp. *saeculare*	LMG14934	1
B. saguini	DSM23967	1
B. scardovii	LMG21589	1
B. stellenboschense	DSM23968	0,25
B. stercoris	DSM24849	1
B. subtile	LMG11597	0,5
B. porcinum	LMG21689	1
B. thermacidophilum	LMG21395	1
B. thermophilum	DSM20212	1
B. tissieri	DSM100201	1
B. tsurumiense	JCM13495	0,5
B. vansinderenii	LMG30126	1

Table S3. General genome features of *Bifidobacterium breve* and *Bifidobacterium longum* strains used for the comparative genomic analysis.

Strains	Nº of ORFs	Nº tRNA	Nº rRNA loci	Nº TUGs	AMC MIC (μg/mL)
B. breve 1900B	1929	52	3	85	0,25
B. breve UCC2003	2040	53	2	175	0,5
B. breve 689B	1937	52	2	130	0,125
B. breve 7E	1876	52	3	41	1
B. breve 2L	1888	51	2	79	1
B. breve 31L	1921	52	3	157	1
B. breve 1891B	2118	54	3	77	32
B. breve M1D	2053	54	3	48	16

Strains	Nº of ORFs	Nº tRNA	Nº rRNA loci	Nº TUGs	AMC MIC (μg/mL)
B. longum subsp. *longum* 35B	2219	57	5	460	0,25
B. longum subsp. *longum* 1890B	1988	59	2	284	0,5
B. longum subsp. *longum* 296B	1874	54	4	226	1
B. longum subsp. *longum* LMG13197	1950	73	4	216	1
B. longum subsp. *infantis* ATCC15697	2564	77	4	972	0,5
B. longum subsp. *longum* 39B	1842	56	1	112	4
B. longum subsp. *longum* 1898B	2065	56	4	289	8

Table S4. Filtered reads of co-culture experiments.

Sample	Reads classified
Fecal sample 12h	75308
Fecal sample + AMC 12h	76483
Fecal sample + 1891B 12h	82948
Fecal sample + 1891B + AMC 12h	95128
Fecal sample + M1D 12h	55240
Fecal sample + M1D + AMC 12h	61718
Fecal sample + 39B 12h	70871
Fecal sample + 39B + AMC 12h	52053
Fecal sample + 1898B 12h	40994
Fecal sample + 1898B + AMC 12h	49033
Fecal sample 18h	60001
Fecal sample + AMC 18h	72953
Fecal sample + 1891B 18h	92393
Fecal sample + 1891B + AMC 18h	91941
Fecal sample + M1D 18h	40295
Fecal sample + M1D + AMC 18h	49669
Fecal sample + 39B 18h	63771
Fecal sample + 39B + AMC 18h	21970
Fecal sample + 1898B 18h	40994
Fecal sample + 1898B + AMC 18h	10829
Fecal sample 24h	87558
Fecal sample + AMC 24h	92955
Fecal sample + 1891B 24h	92365
Fecal sample + 1891B + AMC 24h	60316
Fecal sample + M1D 24h	66221
Fecal sample + M1D + AMC 24h	47072
Fecal sample + 39B 24h	36264
Fecal sample + 39B + AMC 24h	85101
Fecal sample + 1898B 24h	72085
Fecal sample + 1898B + AMC 24h	16775
Fecal sample 36h	88153
Fecal sample + AMC 36h	87836
Fecal sample + 1891B 36h	71966
Fecal sample + 1891B + AMC 36h	81918
Fecal sample + M1D 36h	42000
Fecal sample + M1D + AMC 36h	52063
Fecal sample + 39B 36h	44863
Fecal sample + 39B + AMC 36h	62806
Fecal sample + 1898B 36h	83624
Fecal sample + 1898B + AMC 36h	21678

Chapter 7

The Stepwise Assembly of the Human Vaginal Microbiota is shaped by *Lactobacillus crispatus*

Mancino W*, Mancabelli L*, Lugli GA, Milani C, Viappiani A, Anzalone R, Longhi G, van Sinderen D, Ventura M, Turroni F.

The results of this chapter are submitted to Applied and Environmental Microbiology.

*These authors contribuited equally.

ABSTRACT

The vaginal microbiota is defined as the community of bacteria residing in the human vaginal tract. Recent studies have demonstrated that the high prevalence of *Lactobacillus crispatus* species is commonly associated with a healthy vaginal environment. In the current study, we assessed the microbial composition of 94 public healthy vaginal samples through shotgun metagenomics analyses. Results showed that *L. crispatus* was the most representative species and correlated negatively with bacteria involved in vaginal infections. Moreover, we isolated and sequenced the genome of new 15 *L. crispatus* strains from different environments and the comparative genomics analysis revealed a genetic adaptation of strains to their ecological niche. In addition, *in-vitro* growth experiments display the capability of this species to modulate the composition of the vaginal microbial consortia. Overall, our findings suggest an ecological role exploited by *L. crispatus* in reducing the complexity of the vaginal microbiota toward a depletion of pathogenic bacteria.

MATERIAL AND METHODS

Selection and metagenomics analyses of public datasets.

In this study, we performed a metagenomics analysis based on three publicly available WMS datasets based on Illumina sequencing technology, corresponding to a total of 94 vaginal samples (Table S1). Specific metadata regarding age, diet and any therapies was not available. In order to classify the reads to the lowest possible taxonomic rank, the downloaded fastq files were analyzed with the METAnnotatorX bioinformatics platform (310).

Isolation of new *L. crispatus* strains.

In order to investigate the occurrence of *L. crispatus* in vaginal microbiota and in poultry fecal samples, seven human healthy vaginal fluid samples and 10 feral chicken fecal samples were explored. Vaginal samples were collected through vaginal swabs mixed with 4 mL of phosphate-buffered saline (PBS - pH 6.5), while fecal samples were composed of 10 g of fresh fecal material, which is a sufficient quantity to represent the whole biodiversity of fecal microbiota, as reported in a previous study (163). Serial dilution and subsequent plating were performed using MRS agar supplemented with 1 % of lactose (wt/vol). Morphologically different colonies that developed on MRS plates were picked and streaked in order to isolate purified bacterial strains.

Taxonomic identification of new isolated strains.

Identification of each isolate was performed by PCR amplification of a portion of the 16S rRNA gene through primer pair P0 (5'-GAAGAGTTTGATCCTGGCTCAG-3') and P6 (5'-CTACGGCTACCTTGTTACGA-3'). Each 16S rRNA gene thus generated from individual strains

was sequenced and was analyzed by BLAST against the GenBank database. Moreover, the genome of the *L. crispatus* strains identified were also characterized using *L. crispatus* species-specific primers: Lbc_fw: 5'-AGGATATGGAGAGCAGGAAT-3', Lbc_rv: 5'-CAACTATCTCTTACACACGCC-3' (311-314). Each sample was subjected to the following thermal cycling conditions: 5 min at 94ºC for one cycle, then 20 s at 94ºC, 30 sec at 57ºC and 40 sec at 72ºC for 30 cycles, followed by a 5-min elongation period at 72ºC. The *L. crispatus* isolated strains are listed in Table 1.

pH and Sodium Chloride tolerance tests.

The ability of isolated strains to tolerate varying pH levels or NaCl concentrations was evaluated by monitoring Optical density (OD) values on 96-well plates. In different wells, MRS medium enhanced with 1 % of lactose was supplemented with 2 %, 6 % or 10 % NaCl (wt/vol). In addition, isolated strains were also cultivated in MRS medium supplemented with 1 % of lactose set by addition of HCl at pH 2.0, pH 3.0 or pH 4.0. Microliter plates were incubated under aerobic conditions for 48h at 37°C. Cell density was monitored by OD measurements at OD600 using a plate reader (BioTek, VT, USA). The resistance level to the imposed sodium chloride or pH stress was calculated in each case by comparing the maximum OD600 reached in a particular medium with that of the control medium (MRS + 1 % of lactose). Assays were performed in triplicate as independent experiments.

Glycogen growth assays.

Lactobacillus strains were cultivated on semi-synthetic MRS medium without sugar supplemented with 1 % (wt/vol) of glycogen (Merck, Germany) as the sole carbon source. OD measurements at a wave length of 600 nm were detected using a plate reader (BioTek, VT, USA). The plate reader was read in intermittent mode, with absorbance readings performed at three minutes intervals for three times after 12h, 15h, 18h, 24h, 36h, 40h and 48h of growth, where each reading was ahead of 30 s

shaking at medium speed. Cultures were grown in biologically independent triplicates in aerobic conditions at 37ºC. The resulting growth data were expressed as the mean of these replicates. Control growth experiments were carried out with strains growth in MRS supplemented with 1 % of lactose (wt/vol).

Detection of hydrogen peroxide production.

Lactobacillus strains were tested for the production of H_2O_2. Each strain was plated on TMB-plus agar, as previously described (315). In brief, lactobacilli strains were plated on TMB-plus agar and incubated for 48 h at 37ºC in anaerobic conditions (2.99 % H_2, 17.01 % CO_2 and 80 % N_2) in a chamber (Concept 400, Ruskin). After exposure to ambient air, hydrogen peroxide-producing colonies turning blue. Change of color was assessed after 30 minutes, as previously described (316). Color intensity was defined based on the transposition of the intensity of the colonies blue color on Adobe Photoshop software (Adobe Incorporated, California, USA) after a photography of each plate. The color grade was designated with the combination of the Red, Green and Blue (RGB) color method lights, maintaining the R and the G grade to 0. Each experiment was performed in triplicate.

Genome sequencing and assemblies.

DNA extracted from lactobacilli isolates was submitted to whole-genome sequencing using MiSeq (Illumina, UK) at GenProbio srl (Parma, Italy) according to the supplier's protocol (Illumina, UK). Furthermore, DNA isolated from *L. crispatus* PRL2021 was also subjected to whole-genome sequencing using a MinION (Oxford Nanopore, UK) at GenProbio srl (Parma, Italy) according to the supplier's protocol (Oxford Nanopore, UK). Fastq files of the paired-end reads obtained from targeted genome sequencing of isolated strains were used as input for genome assemblies through the MEGAnnotator pipeline (166). SPAdes software was used for de novo assembly of each *L. crispatus* genome sequence (167, 168), while protein-encoding ORFs were predicted using Prodigal (169). The

coverage depth of these newly isolated 16 *L. crispatus* chromosomes ranged from 73- to 259-fold, which upon assembly generated 36 to 259 contigs (Table 1).

Comparative genomics.

A pan-genome calculation was performed using the pan-genome analysis pipeline PGAP (293), including each *L. crispatus* genome collected from this study and 94 *L. crispatus* genomes public available (NCBI database) (Table S1). 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 (171) (cutoff, E value of $< 1 \times 10^{-4}$ and 50 % identity over at least 80 % of both protein sequences). The resulting output was then clustered into protein families by means of MCL (graph theory-based Markov clustering algorithm), using the gene family method. A pan-genome 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 (173) and phylogenetic trees were constructed using ClustalW (174). Based on these comparative analyses, a *L. crispatus* supertree was visualized using FigTree (http://tree.bio.ed.ac.uk/software/figtree/). The average nucleotide identity (ANI) values among *L. crispatus* genomes analyzed was calculate with fastANI software (317). Functional annotation of each protein of *L. crispatus* strains was performed employing the eggNOG database (187).

Prediction of putative bacteriocin-encoding genes.

The genome sequences of the 16 isolated and of the 94 public available *L. crispatus* genomes were screened for bacteriocin-encoding genes. Moreover, 116 genomes publicly available belonged to *L. iners*, *L. jensenii* and *L. gasseri* species were explored for the bacteriocins prediction. The screening was carried out using the BAGEL3 database (318) through BLASTP analysis (E-value cutoff of $1e^{-}$

^{5}) (171). Afterwards, a manual examination of the sequences predicted to encode a bacteriocin-like protein was performed. The predicted bacteriocin genes were classified according to the bacteriocin classification reported in the BAGEL3 database (23677608).

Growth experiment of the isolated vaginal *L. crispatus* in Simulated Vaginal Fluid (SVF) and shallow shotgun metagenomics.

The SVF was prepared as mentioned by Pan et al. (319). For growth experiments, overnight lactobacilli cultures were diluted since an OD value of 1.0. Each culture was inoculated at 1 % (vol/vol) into SVF. We performed nine different experiments in which were inoculated *L. gasseri* ATCC 9857, *L. iners* LMG 14328 and *L. jensenii* V94G (isolated in this study) together with one new vaginal-isolated *L. crispatus* strain. Batch cultures were incubated under aerobic conditions for 12h at 37°C. After 12h of growth, the cultures were centrifuged at 3000 rpm for 8 minutes and the pellets were harvested. The pellets were subjected to DNA extraction using the GeneElute^{TM} Bacterial Genomic DNA Kit (Sigma, Germany) following the manufacturer's instruction. The extracted DNA was prepared following the Illumina Nextera XT protocol. Briefly, DNA samples were enzymatically fragmented, barcoded and purified involving magnetic beads. Then, samples were quantified using fluorometric Qubit quantification system (Life Technologies, USA), loaded on a 2200 Tape Station Instrument (Agilent Technologies, USA) and normalized to 4 nM. Sequencing was performed paired-end using an Illumina NextSeq 500 sequencer with NextSeq High Output v2 Kit Chemicals (Illumina Inc., San Diego, USA). The retained reads were analyzed with the METAnnotatorX bioinformatics platform (310).

Growth experiment of the isolated *L. crispatus* PRL2021 in SVF and Fecal Media with fecal samples.

An overnight *L. crispatus* PRL2021 culture was diluted to reach an OD value of 1.0 and was then inoculated at 1 % (vol/vol) into SVF and Fecal media. The SVF was prepared as mentioned by Pan

et al. (319), whereas the fecal media was prepared as described by Macfarlane et al. (292). We performed four different experiments for each growth media in which were inoculated a woman fecal sample at 1 % (wt/vol) or 5 % (wt/vol) together with isolated *L. crispatus* PRL2021 strain. Control growth experiments were carried out without *L. crispatus* PRL2021. After 12h of growth, the cultures were centrifuged at 3000 rpm for 8 minutes and the pellets were harvested. The pellets were subjected to DNA extraction using the QIAamp DNA Stool Mini kit (Qiagen) following the manufacturer's instructions. The extracted DNA was prepared following the Illumina Nextera XT protocol as mentioned previously for co-culture experiments.

Evaluation of cell density by flow cytometry assay.

For bacterial cell counting, 500 µL of batch culture was diluted in physiological solution (Phosphate-buffered saline, PBS). Subsequently, bacterial cells were stained with one µL SYBR®Green I and incubated in the dark for at least 15 min before measurement. All count experiments were performed using an Attune NxT Flow flow cytometer (Invitrogen, ThermoFisher Scientific) equipped with a Blue Laser set at 50 mWatt and tuned to an excitation wavelength of 488 nm. Multiparametric analyses were performed on both scattering signals (FSC, SSC) and SYBR Green I fluorescence was detected on FL1 channel. Cell debris and eukaryotic cells were excluded from acquisition analysis by a sample-specific FL1 threshold. All data sets were statistically analyzed with Attune NxT Flow Cytometer Software. Utilizing these cell counts to normalize the sequencing data into absolute abundance of each profiled taxa, we were capable to perform quantitative microbiome profiling using a previously described method (287)(320).

Statistical analysis.

The hierarchical clustering (HCL) of samples was obtained using bacterial composition at species level and was calculated through TMeV 4.8.1 software using Pearson correlation as a distance metric based on information at genus level. The data obtained was represented by a cladogram. SPSS

software (IBM, Italy) was used to perform statistical analysis by Student t-test and by Kendall Tau-rank co-occurrence.

Data Deposition.

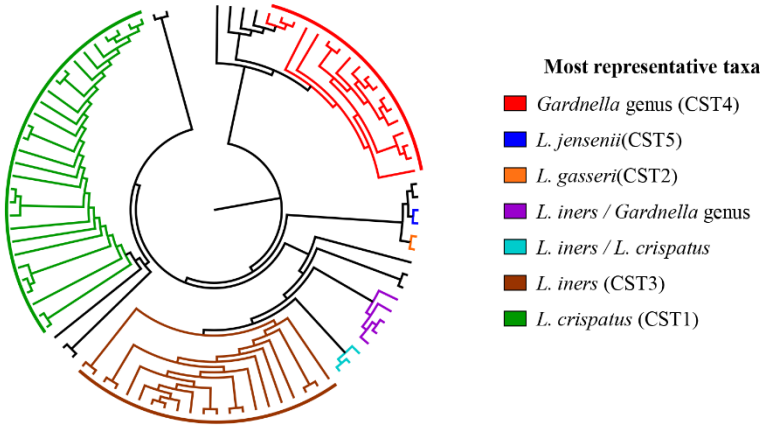
Raw sequences of the Shallow-shotgun metagenomics profiling experiments are accessible through SRA study accession numbers PRJNA641015. Newly isolated *Lactobacillus* genomes were sequenced and deposited with BioSample number reported in Table 1

RESULTS & DISCUSSION

Evaluation of VM composition of healthy women. In 2011 Ravel et al. identified and classified vaginal microbiota in five Community State Types (CSTs) (18). In detail, four CSTs were dominated by different species of *Lactobacillus*, i.e. *L. crispatus* (CST 1), *L. gasseri* (CST 2), *L. iners* (CST 3) and *L. jensenii* (CST 5), whereas the fourth (CST 4) possesses lower proportions of lactic acid bacteria and higher proportions of strictly anaerobic organisms. Subsequent studies, mostly based on 16S rRNA gene-based amplicon sequencing, have confirmed the presence of these CSTs (321-325). Metagenomic shotgun techniques, i.e., whole metagenome sequencing (WMS), have allowed overcoming some of the limitations of 16S rRNA gene-based amplicon sequencing, such as DNA extraction method and primer pair efficiency. In particular, WMS allows the evaluation of the bacterial communities at species level in a more precise and effective method compared to 16S rRNA gene-based amplicon sequencing. Consequently, in order to assess the VM composition of healthy women, we decided to perform a thorough literature search for vaginal shotgun metagenomic datasets based on Illumina sequencing technology. In detail, we collected a total of 94 publicly available data sets corresponding to vaginal samples of healthy women (Table S1) that were re-analyzed using a shallow shotgun metagenomics approach (326), thereby achieving a very accurate taxonomic cataloguing down to species level resolution of the VM. The metagenomic analysis included a total of 2,361,660 reads with an average of 25,124 ± 23,730 reads per sample (Table S1). In detail, bacterial species analysis showed that 60.64 % of the samples presented a vaginal microbiota dominated by *Lactobacillus* species, confirming that in the majority of women a healthy vaginal microbiota is dominated by *Lactobacillus* species (327). Interestingly, 32.98 % and 28.72 % of the samples showed *L. crispatus* and *L. iners*, respectively, as the most abundant vaginal bacterium. In addition, 29.79 % and 23.40 % of the samples showed a relative abundance of > 40 % of *L. crispatus* and *L. iners*, respectively. Furthermore, we evaluated the presence of CST in the 94 samples collected through hierarchical clustering (HCL) analysis (Figure 1). The HCL analysis of these publicly available datasets showed the presence of all previously identified CSTs, confirming the predominance of

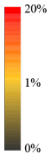
CST1 and CST3, which represented 30.85 % and 22.34 % of the samples, respectively. Furthermore, the 19.15 % of the samples displayed a microbial profile corresponding to CST4 characterized by the presence of reads belonging to the *Gardnerella* spp. In order to detect the main bacterial taxa belonging to CST1, CST3 and CST4 identified in this study, taxonomic profiling at species level was assessed to reconstruct the core microbiota members, i.e. core-microbiota, of the VM of each CST (22647042). For this purpose, we reconstructed the core-microbiota by selecting bacterial species that occur with a prevalence greater than 80 % among the collected samples of a particular CST (328, 329). Interestingly, the CST1 and CST3 showed a core-microbiota characterized by bacteria belonging to the *Lactobacillus* genus, while CST4 showed a more complex core-microbiota characterized by *Gardnerella* genus, potential vaginal pathogens, such as *Atopobium vaginae*, as well as *L. iners*. Furthermore, the identification of the accessory microbiota, i.e. bacterial species different from core-microbiota and with a prevalence greater than 30 % and an average relative abundance greater than 0.5 % (330), revealed a lower biodiversity of the CST1 microbiota. In detail, comparison of the accessory microbiota of CST1 with those of CST3 and CST4 showed a 50 % and 76.9 % decrease in the number species, respectively. Moreover, the observed concurrence of *L. iners* and *Gardnerella* genus suggested investigating the possible correlations between the characteristic vaginal bacteria. Consequently, the relative abundance of the fourth CST *Lactobacillus* species and the most common vaginal pathogens, composing the 94 samples, were employed in a co-variance analysis based on Kendall tau-rank (Figure 1). Interestingly, *L. crispatus*, *L. jensenii* and *L. gasseri* correlated negatively with at least 50 % of the potential vaginal pathogens, highlighting a possible ability of these microorganisms to outcompete and outnumber bacteria involved in vaginal infections. Particularly, these results could reinforce the hypothesis that the *L. crispatus* species reduces the biodiversity of VM and consequently protects this environment from viral and bacterial infections (20, 21, 35-37).

a)



b)

Taxonomy	CST1		CST3		CST4	
	Prevalence	Average	Prevalence	Average	Prevalence	Average
Atopobium vaginae	20.69%	0.24%	38.10%		88.89%	8.01%
Gardnerella unknown_species	27.59%	1.35%	33.33%	0.30%	94.44%	10.40%
Gardnerella vaginalis	48.28%	1.01%	66.67%	0.92%	100.00%	45.08%
Lactobacillus crispatus	100.00%	83.11%	66.67%	1.14%	44.44%	0.93%
Lactobacillus gasseri	93.10%	0.63%	95.24%	0.95%	66.67%	4.80%
Lactobacillus helveticus	96.55%	0.24%	42.86%		11.11%	
Lactobacillus iners	65.52%	2.87%	100.00%	79.15%	94.44%	1.05%
Lactobacillus jensenii	93.10%	2.01%	76.19%	3.32%	61.11%	0.04%
Lactobacillus kefiranofaciens	96.55%	0.39%	38.10%		11.11%	
Lactobacillus unknown_species	93.10%	1.15%	100.00%	4.07%	83.33%	0.19%
Streptococcus cristatus	34.48%		71.43%	0.87%	83.33%	0.11%



c)

	Lactobacillus crispatus	*Lactobacillus iners*	*Lactobacillus jensenii*	*Lactobacillus* unknown_species	*Lactobacillus gasseri*	*Gardnerella vaginalis*	*Gardnerella* unknown_species	*Atopobium vaginae*	*Escherichia coli*	*Prevotella* unknown_species	*Prevotella bivia*	*Chlamydia trachomatis*	*Megasphaera stantonii*	*Streptococcus agalactiae*	*Megasphaera* unknown_species	*Mobiluncus curtisii*	*Ureaplasma parvum*	*Mycoplasma genitalium*	*Enterococcus faecalis*	*Mycoplasma hominis*	*Mobiluncus mulieris*
Lactobacillus crispatus	1.000	-0.302	0.296	0.203	0.073	-0.347	-0.273	-0.315	0.002	-0.042	-0.194	-0.045	-0.162	-0.184	-0.281	-0.133	-0.111	-0.111	-0.089	-0.247	-0.135
Lactobacillus iners	-0.302	1.000	-0.038	0.155	0.026	0.027	0.071	-0.015	-0.017	-0.165	-0.082	0.006	-0.018	0.029	-0.035	0.031	0.305	0.002	0.048	0.132	0.044
Lactobacillus jensenii	0.296	-0.038	1.000	0.344	0.265	-0.352	-0.256	-0.274	0.054	-0.167	-0.145	-0.069	-0.151	-0.262	-0.217	-0.128	0.001	-0.117	-0.069	-0.185	-0.171
Lactobacillus unknown_species	0.203	0.155	0.344	1.000	0.296	-0.289	-0.166	-0.204	0.097	-0.159	-0.153	-0.065	-0.188	-0.124	-0.216	-0.056	0.132	-0.024	0.019	-0.151	-0.044
Lactobacillus gasseri	0.073	0.026	0.265	0.296	1.000	-0.228	-0.146	-0.236	0.169	-0.277	-0.169	-0.163	-0.358	-0.191	-0.312	-0.21	-0.104	-0.128	-0.030	-0.107	-0.182
Gardnerella vaginalis	-0.347	0.027	-0.352	-0.289	-0.228	1.000	0.69	0.565	-0.111	0.172	0.342	0.061	0.256	0.339	0.318	0.273	0.172	0.084	0.147	0.187	0.304
Gardnerella unknown_species	-0.273	0.071	-0.256	-0.166	-0.146	0.69	1.000	0.559	-0.175	0.177	0.381	0.115	0.171	0.286	0.307	0.365	0.233	0.136	0.159	0.244	0.451
Atopobium vaginae	-0.315	-0.015	-0.274	-0.204	-0.236	0.565	0.559	1.000	-0.118	0.229	0.279	0.154	0.323	0.26	0.437	0.318	0.208	0.099	0.128	0.306	0.334
Escherichia coli	0.002	-0.017	0.054	0.097	0.169	-0.111	-0.175	-0.118	1.000	-0.307	-0.229	-0.111	-0.217	-0.097	-0.181	-0.112	-0.109	-0.101	-0.091	-0.176	-0.102
Prevotella unknown_species	-0.042	-0.165	-0.167	-0.159	-0.277	0.172	0.177	0.229	-0.307	1.000	0.423	0.133	0.47	0.335	0.443	0.21	0.016	0.176	0.047	0.21	0.244
Prevotella bivia	-0.194	-0.082	-0.145	-0.153	-0.169	0.342	0.381	0.279	-0.229	0.423	1.000	0.076	0.381	0.372	0.46	0.283	0.009	0.150	0.031	0.244	0.269
Chlamydia trachomatis	-0.045	0.006	-0.069	-0.065	-0.163	0.061	0.115	0.154	-0.111	0.133	0.076	1.000	0.051	0.023	0.179	0.208	0.135	-0.022	0.178	0.244	0.244
Megasphaera stantonii	-0.162	-0.018	-0.151	-0.188	-0.358	0.256	0.171	0.323	-0.217	0.47	0.381	0.051	1.000	0.311	0.731	0.130	0.124	0.229	0.007	0.216	0.107
Streptococcus agalactiae	-0.184	0.029	-0.262	-0.124	-0.191	0.339	0.286	0.26	-0.097	0.335	0.372	0.023	0.311	1.000	0.339	0.229	0.143	0.107	0.044	0.269	0.093
Megasphaera unknown_species	-0.281	-0.035	-0.217	-0.216	-0.312	0.318	0.307	0.437	-0.181	0.443	0.46	0.179	0.731	0.339	1.000	0.206	0.028	0.265	-0.004	0.302	0.197
Mobiluncus curtisii	-0.133	0.031	-0.128	-0.056	-0.21	0.273	0.365	0.318	-0.112	0.21	0.283	0.208	0.130	0.229	0.206	1.000	0.114	-0.046	0.264	0.426	0.502
Ureaplasma parvum	-0.111	0.305	0.001	0.132	-0.104	0.172	0.233	0.208	-0.109	0.016	0.009	0.135	0.124	0.143	0.028	0.114	1.000	0.026	0.099	0.088	0.140
Mycoplasma genitalium	-0.111	0.002	-0.117	-0.024	-0.128	0.084	0.136	0.099	-0.101	0.176	0.150	-0.022	0.229	0.107	0.265	-0.046	0.026	1.000	-0.060	-0.038	0.298
Enterococcus faecalis	-0.089	0.048	-0.069	0.019	-0.030	0.147	0.159	0.128	-0.091	0.047	0.031	0.178	0.007	0.044	-0.004	0.264	0.099	-0.060	1.000	0.088	0.153
Mycoplasma hominis	-0.247	0.132	-0.185	-0.151	-0.107	0.187	0.244	0.306	-0.176	0.21	0.244	0.244	0.216	0.269	0.302	0.426	0.088	-0.038	0.088	1.000	0.153
Mobiluncus mulieris	-0.135	0.044	-0.171	-0.044	-0.182	0.304	0.451	0.334	-0.102	0.244	0.269	0.244	0.107	0.093	0.197	0.502	0.140	0.298	0.153	0.153	1.000

Figure 1

Figure 1. Evaluation of vaginal microbiota composition of healthy women. Panel a illustrates a circular cladogram of the 94 publicly samples based on hierarchical clustering (HCL) analysis. The samples were subdivided according to the most representative taxa and CST. Panel b reports the prevalence and average abundance of the bacteria that correspond to CST1, CST3 and CST4. The values that represent the core-microbiota are highlighted in green. Panel c reports the results of Kendall tau-rank co-variance analysis based on the fourth CST *Lactobacillus* species and the most common vaginal pathogens. The significant positive correlations are highlighted in green, while significant negative correlations are evidenced in red.

Isolation of *L. crispatus* strains from human and chicken samples.

In order to isolate *L. crispatus* strains, we collected seven vaginal swab samples of healthy women and 10 feral chicken fecal samples, following a protocol for the isolation of members of the *Lactobacillus* genus (see Materials & Methods section). We decided to screen healthy vaginal swabs and chicken fecal samples based on the previously described high prevalence of *L. crispatus* strains in these environments (316, 331, 332). In total 18 different strains were isolated, all belonging to the *Lactobacillus* genus. In this context, 15 strains were determined to represent *L. crispatus* species, two were shown to be *L. jensenii* species and one isolate was demonstrated to belong to the *L. gasseri* species. Moreover, strain *L. crispatus* M247 was isolated from a commercial probiotic product (Table 1). Regarding the novel *L. crispatus* isolated, seven strains were isolated from poultry fecal samples, whereas eight strains were isolated from vaginal fluid (Table 1). The relative ease by which such a high number of *L. crispatus* was isolated is consistent with the high abundance level of this species in the human VM and in the GIT of chickens, highlighting its ecological specialization for these environments (316, 331, 332).

Evaluation of novel *L. crispatus* strains to vaginal environment challenges.

In order to evaluate the ecological adaptation of the isolated strains to the vaginal environment, all identified *L. crispatus* strains were subjected to growth experiments in the presence of environmental stress conditions, i.e., acidic- and osmotic stress, simulating those naturally occurring in the vaginal tract (6, 17, 324, 333-335). Interestingly, all isolates showed comparable growth performances at varying NaCl concentrations (Fig. 2). Specifically, *L. crispatus* strains showed a survival rate of > 50 % when the concentration of NaCl in growth media was 2 % (wt/vol). Furthermore, three isolated strains from chicken fecal samples, i.e. LB67, LB69 and LB70, were not able to tolerate this level of osmotic stress, even at lower NaCl concentration (Fig.2).

Table 1. *Lactobacillus* strains used in this study and general genome features.

Strain	Size (Bp)	ORFs number	rRNA loci	tRNA	Coverage depth	Contig number	Origin	BioSample
L. crispatus PRL2021	2329621	2340	6	64	126-	43	Vaginal Swab	SAMN15357510
L. crispatus LB56	2160961	2215	3	55	137-	195	Vaginal Swab	SAMN15357362
L. crispatus LB57	2382340	2496	3	65	132-	259	Vaginal Swab	SAMN15357363
L. crispatus LB58	2236149	2232	3	65	184-	199	Vaginal Swab	SAMN15357364
L. crispatus LB59	2135374	2172	3	66	182-	189	Vaginal Swab	SAMN15357365
L. crispatus M247	2098939	2083	3	57	256-	156	Probiotic product	-
L. crispatus LB61	2081602	2087	3	66	228-	140	Vaginal Swab	SAMN15357366
L. crispatus LB62	2160518	2235	3	63	165-	172	Vaginal Swab	SAMN15357367
L. crispatus LB63	2132629	2178	4	60	141-	182	Vaginal Swab	SAMN15357368
L. crispatus LB64	2051142	2040	3	57	259-	36	Chicken Fecal Sample	SAMN15357369
L. crispatus LB65	2061272	2035	3	61	156-	44	Chicken Fecal Sample	SAMN15357370
L. crispatus LB66	2041845	2006	3	56	209-	44	Chicken Fecal Sample	SAMN15357371
L. crispatus LB67	1967651	1912	3	60	193-	60	Chicken Fecal Sample	SAMN15357372
L. crispatus LB68	1966006	1921	2	56	73-	111	Chicken Fecal Sample	SAMN15357373
L. crispatus LB69	1977839	1959	3	55	79-	95	Chicken Fecal Sample	SAMN15357374
L. crispatus LB70	2021157	1988	3	54	76-	199	Chicken Fecal Sample	SAMN15357375
L. gasseri V105C	-	-	-	-	-	-	Vaginal Swab	Not Sequenced
L. gasseri ATCC 9857	-	-	-	-	-	-	ATCC collection	Not Sequenced
L. jensenii V79H	-	-	-	-	-	-	Vaginal Swab	Not Sequenced
L. jensenii V94G	-	-	-	-	-	-	Vaginal Swab	Not Sequenced
L. iners LMG 14328	-	-	-	-	-	-	LMG collection	Not Sequenced

Notably, strains isolated from vaginal swabs displayed a survival rate that was significantly higher than those isolated from chicken fecal samples (p-value = 0.012), suggesting a higher tolerance to osmotic stress. In this context, it is worth mentioning that the highest survival rate in the presence of osmotic stress was displayed by *L. crispatus* PRL2021 (Fig. 2). In addition, all isolates do not tolerate harsh acidic conditions, displaying survival rates of < 20 % when exposed to a pH 2.0 or pH 3.0 for 48h. Conversely, six strains, i.e. PRL2021, LB57, LB61, LB62, LB63 and LB66, reached a survival rate of > 85 % when the pH of the growth medium was 4.0. Notably, also under this growth condition, strain PRL2021 reached the highest survival rate compared to other assessed *L. crispatus* isolates, with a percentage of 96.81 % (Fig. 2). Interestingly, five of the strains mentioned above, that were shown to elicit the highest survival rate at pH 4.0, had been isolated from vaginal fluid, suggesting

an adaptation of these strains to their natural ecological niche due to the acidic conditions of the vaginal environment (336).

Different studies showed that the presence of luminal glycogen in the vaginal tract of healthy women is associated with a high prevalence of lactobacilli (337, 338). In order to investigate the utilization of glycogen by *Lactobacillus* strains, *in vitro* cultivation assays were performed in growth medium, i.e., MRS without sugar, containing glycogen as the unique carbon source. The control assay was carried out with MRS medium supplemented with lactose and the cultures were monitored for 48h (see Materials & Methods section). Interestingly, *L. crispatus* strains isolated from vaginal swab grew significantly better when compared to strains of poultry origin (p-value < 0.01), confirming the adaptation of these strains for the utilization of a carbon source naturally present in their ecological niche (Figure S1). Among strains of vaginal origin, PRL2021 and LB63 showed the highest growth performance, reaching OD values of > 0.42 after 12h of growth (Figure S1). In contrast, strains of poultry origin were not able to use glycogen as a carbon source showing low OD values (average OD 0.33 ± 0.15) even after 48h of growth (Figure S1). In addition, other tested *Lactobacillus* strains belonging to *L. gasseri*, *L. jensenii* and *L. iners* species displayed growth levels 1.5-fold lower when compared to *L. crispatus* strains (p-value < 0.05) (Figure S1).

Production of H_2O_2 by *Lactobacillus* strains.

Hydrogen peroxide production is generally considered to be a key selective factor allowing *Lactobacillus* spp. to dominate the vaginal environment (315, 339, 340). Moreover, it has been demonstrated that women vaginally colonized by lactobacilli producing hydrogen peroxide were less exposed to BV compared to women colonized by *Lactobacillus* strains that did not produce H_2O (341, 342). For these reasons, in order to investigate the production of hydrogen peroxide, 21 *Lactobacillus* strains involving *L. crispatus*, *L. iners*, *L. jensenii* and *L. gasseri* species (Table 2), were tested by the qualitative TMB-plus peroxidase assay (see Material & Methods) (315, 316). Interestingly, all the *L.*

crispatus strains tested, except for *L. crispatus* LB63, produced variable levels of hydrogen peroxide without any correlation to their ecological origin. The colony-associated blue color intensity of H_2O_2 producers was different in the examined *L. crispatus* strains. In particular, four strains, i.e. *L. crispatus* PRL2021, *L. crispatus* LB58, *L. crispatus* LB65 and *L. crispatus* LB66, produced colonies with an intense dark blue colour, suggesting strong H_2O_2production (Table 2). In contrast, other H_2O_2-producing strains produced colonies with a less blue intensity (Table 2). Furthermore, 60 % of the strains of other typical vaginal *Lactobacillus* species tested did not produce H_2O_2 (Table 2), reflecting data reported in previous studies, highlighting that *L. crispatus* species is one of the highest H_2O_2 generating species among typical vaginal lactobacilli (316, 343). The H_2O_2-producing feature is a strain-specific feature, as demonstrated by our results and by previously published studies (316, 343). Strains producing hydrogen peroxide could display a higher ecological fitness (315) and their presence could be important in the modulation of the vaginal microbiota.

Table 2. H_2O_2 production by lactobacilli strains isolated in this study.

Strains	H_2O_2 production	RGB Code Color[a]
L. crispatus PRL2021	+	0-0-100
L. crispatus LB56	+	0-0-215
L. crispatus LB57	+	0-0-160
L. crispatus LB58	+	0-0-115
L. crispatus LB59	+	0-0-175
L. crispatus M247	+	0-0-195
L. crispatus LB61	+	0-0-255
L. crispatus LB62	+	0-0-200
L. crispatus LB63	-	-
L. crispatus LB64	+	0-0-230
L. crispatus LB65	+	0-0-130
L. crispatus LB66	+	0-0-150
L. crispatus LB67	+	0-0-190
L. crispatus LB68	+	0-0-190
L. crispatus LB69	+	0-0-245
L. crispatus LB70	+	0-0-240
L. jensenii V79H	+	0-0-105
L. jensenii V94G	+	0-0-121
L. gasseri ATCC 9857	-	-
L. gasseri V105C	-	-
L. iners LMG 14328	-	-

[a]Red, Green and Blue color mode

General genome features of isolated *L. crispatus*.

Genome sequences of the 15 *L. crispatus* isolated strains were decoded through whole-genome sequencing employing the Illumina MiSeq platform. Additionally, DNA extracted from *L. crispatus* PRL2021 was also subjected to whole-genome sequencing using a MinION (Oxford Nanopore, UK). The achieved genome sequences were analyzed to evaluate the genetic content of these isolates. In this context, genome sequences ranged in size from 2,382,340 bp of *L. crispatus* LB57 to 1,966,006 bp of *L. crispatus* LB68 (Table 1). Interestingly, the genome sizes of novel isolates from vaginal fluid

samples were larger than the genome sizes of the isolated strains from poultry fecal samples (average size 2,190,904 bp ± 104,068 bp and 2,012,416 bp ± 41,196 bp in vaginal and poultry isolates, respectively; *p*-value < 0.001), suggesting increased genomic complexity of the strains from human origin. Remarkably, a similar difference in size was also identified in those genomes that are publicly available, confirming that vaginal *L. crispatus* strains possessed an expanded genetic makeup compared to those of poultry origin (average size 2,297,590 bp ± 163,268 bp and 2,056,164 bp ± 7,026 bp in vaginal and poultry origin isolated genomes, respectively; *p*-value < 0.001) (Table S2). Moreover, the Open Reading Frame (ORF) prediction revealed that the ORF number ranged from 2,496 of *L. crispatus* LB57 to 1,912 of *L. crispatus* LB67 (Table 1). Interestingly, 15 sequenced genomes encompassed a similar set of rRNA loci and tRNA genes, which consisted of three or four rRNA operons and between 66 and 54 tRNA genes distributed across the genome (Table 1) except for *L. crispatus* PRL2021 harboring six rRNA operons in its genome (Table 1). Notably, functional classification of *L. crispatus* ORFeome based on eggNOG database (187) was possible for the 82.1 % of the predicted ORFs. In particular, the 15.2 % of the predicted proteins are involved in cellular process and signaling (categories D, M, O, T, U and V), the 26.4 % referred to proteins related to information storage and processing (categories J, K and L) and the 31.6 % described proteins involved in metabolism processes (categories C, E, F, G, H, I, P and Q) (Fig. 2). These data underlined the high number of proteins related to carbohydrate and amino acid transport and metabolism, in according with the wide range of sugars metabolized by members of this species (319, 344) and with the putative weakness of lactobacilli to synthetize most amino acids (345). Recently, it has been identified an orthologous of the *glgX* gene of *Escherichia coli*, which is involved in glycogen metabolism, in *L. crispatus* genomes (338). Different studies demonstrated the correlation between high abundance of lactobacilli in vaginal environment and the presence of high level of luminal glycogen in vaginal tract of healthy women (337, 338, 346).

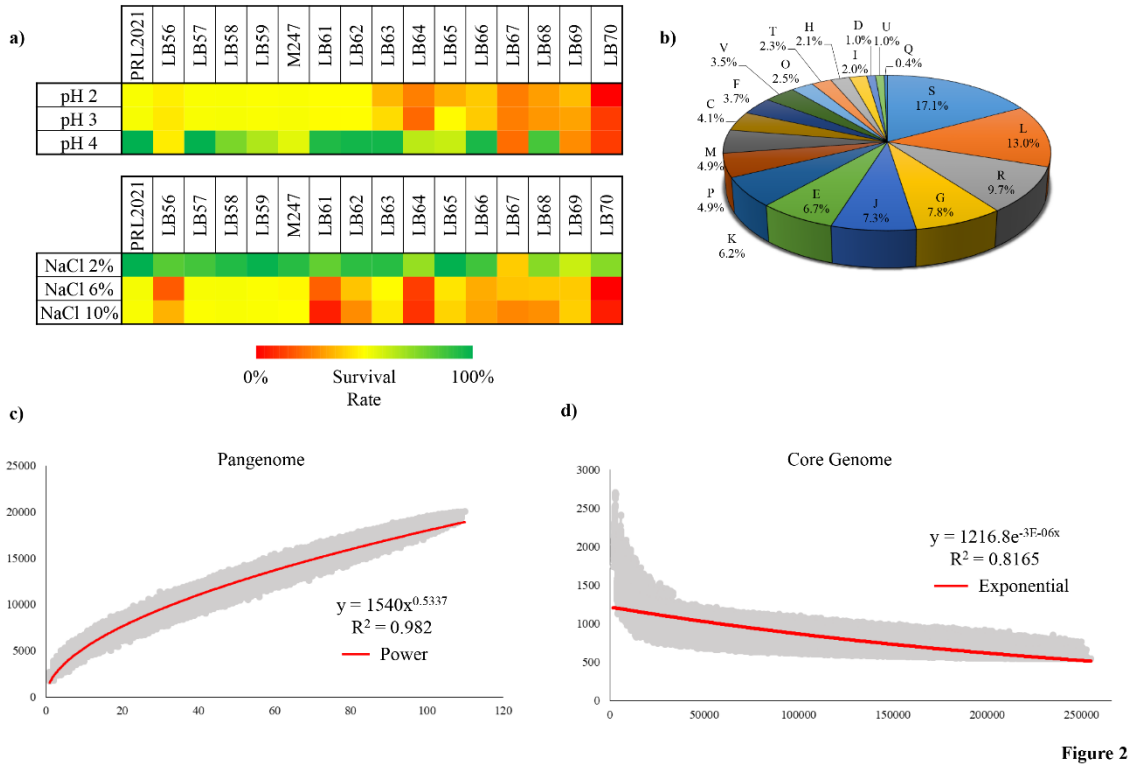


Figure 2

Figure 2. Functional classification of proteome, physiological features of isolated strains and pan-genome and core genome of *L. crispatus* species. Panel a displays the effect of the conditions that simulate vaginal tract for isolated *L. crispatus* strains. The two different heat maps represent the survival rate of each *L. crispatus* strain to NaCl and acidic conditions. Panel b shows functional assignment of the isolated *L. crispatus* proteome based on the eggNOG database. Each letter stands for the following function: **S**, Function unknown, **L**, Replication, recombination and repair, **R**, General function prediction only, **G**, Carbohydrate transport and metabolism, **J**, Translation, ribosomal structure and biogenesis, **E**, Amino acid transport and metabolism, **K**, Transcription, **P**, Inorganic ion transport and metabolism, **M**, Cell wall/membrane/envelope biogenesis, **C**, Energy production and conversion, **F,** Nucleotide transport and metabolism, **V**, Defense mechanisms, **O**, Posttranslational modification, protein turnover, chaperones, **T**, Signal transduction mechanisms, **H**, Coenzyme transport and metabolism, **I**, Lipid transport and metabolism, **D**, Cell cycle control, cell division, chromosome partitioning, **U**, Intracellular trafficking, secretion and vesicular transport, **Q**, Secondary metabolites biosynthesis, transport and catabolism. Panels c and d exhibit the pan-genome and the core genome of the *L. crispatus* species represent as variation of their gene pool sizes upon sequential addition of the 110 genomes analyzed.

Interestingly, we identified three orthologous of this gene in each *L. crispatus* genomes of vaginal origin, whereas genomes of *L. crispatus* from poultry origin encompassed one or two copies of this gene (Table S3). The presence of *glgX* orthologous in vaginal isolates suggests the genetic adaptation of these strains for the metabolism of glycogen.

Prediction of bacteriocin-encoding genes in *L. crispatus* species.

The production of antimicrobial compounds by lactobacilli, including hydrogen peroxide, organic acids, low molecular weight antimicrobial substances and bacteriocins, is considered an important feature of these microorganisms to counteract pathogenic bacteria (316, 347). For this reason, in order to predict bacteriocin-encoding genes, the genomes of isolated *L. crispatus* strains were screened for bacteriocin-encompassing gene clusters by means of the BAGEL3 software and associated data base (318). Notably, all tested genomes were predicted to contain bacteriocin-encoding loci, ranging from three genes of *L. crispatus* LB56 to nine genes of *L. crispatus* LB57 and *L. crispatus* LB70 (Table S4). Interestingly, according to BAGEL-mediated classification, all the predicted bacteriocin-encoding genes belonged to the classes II and III (318) (Table S4). Specifically, all poultry-derived strains, except for LB67 and LB70, only encoded class III-bacteriocins (Table S4), whereas vaginal isolated strains were predicted to produce bacteriocins of both class II and class III (Table S4). The presence of bacteriocin-encoding genes belonged to classes II and III in vaginal-isolated strains suggests a wider range of antimicrobial activity compared to strains of chicken origin. Intriguingly, all isolated strains presented at least one gene predicted to encode for a bacteriocin known as helveticin-J (Table S4). This type of bacteriocin was first identified in *Lactobacillus helveticus* strain 481 and it exhibited a bactericidal mode of action against sensitive indicators (348). Moreover, M23 family metallopeptidase-encoding genes were identified in all sequenced genomes (Table S4). This peptidase family includes proteins that act as a bacteriocin that degrades the peptidoglycan of other bacteria (349). Notably, three strains, i.e. PRL2021, LB57 and LB70, were shown to contain two adjacent genes in their genome encoding a two-component bacteriocin, known as bacteriocin LS2 (350). Interestingly, the genome of strain *L. crispatus* PRL2021, that exhibited a robust production of other antimicrobial compounds, such as H_2O_2 (see above), harbours two distinct loci each encoding helveticin-J (Table S4), suggesting that this strain produces a high level of antimicrobial activity, which may allow this strain to occupy and be highly competitive in particular ecological niches, such as the human vaginal tract. Moreover, two poultry isolated strains, i.e., LB67 and LB70, harbour a

garvicin-encoding gene. Garvicin is a large spectrum non-lantibiotic bacteriocin, previously characterized in *Lactococcus garvieae* species (351-353). Finally, *L. crispatus* strain LB70 was the only strain that encodes a lactacin F bacteriocin (Table S4). In order to explore if *L. crispatus* species exhibit superior antimicrobial activity compared to other *Lactobacillus* species that are typically present in the human vagina, bacteriocin-encoding genes were predicted also in 94 publicly available *L. crispatus* genomes and in 116 publicly available genomes belonged to *L. gasseri*, *L. jensenii* and *L. iners* species (Table S5). This analysis indeed showed that *L. crispatus* strains are predicted to harbour the highest number of bacteriocin-encoding genes in their genomes, ranging from four genes of five different strains to 10 genes of four different strains (Table S3). Notably, no bacteriocin-encoding genes were detected in 56 % and in 74 % of *L. jensenii* and *L. iners* genomes, respectively (Table S5). Moreover, 67 % of the investigated *L. gasseri* strains encompass one or two bacteriocin-encoding genes in their genomes (Table S5). Interestingly, the number of bacteriocin-encoding genes presented in *L. crispatus* genomes is significantly higher (p-value < 0.01) when compared to those identified in other analyzed vaginal *Lactobacillus* species, i.e. *L. iners*, *L. gasseri* and *L. jensenii* (Table S5). The *in silico* prediction of bacteriocin-encoding genes indicates the production of extensive antimicrobial activity by *L. crispatus* strains, in particular when compared to other vaginal *Lactobacillus* species, perhaps supporting a superior ecological fitness of this species.

Pan-genome and core genome analysis of *Lactobacillus crispatus* species.

In order to investigate the genomic differences between *L. crispatus* strains, we performed an extensive comparative genome analysis, which not only encompassed the genomes of the 16 newly isolated *L. crispatus* strains but also the 94 publicly available *L. crispatus* genomes (Table S2). For this purpose, we performed a re-annotation of all publicly available *L. crispatus* chromosomes using the same bioinformatics pipeline applied for the genomes of the 16 newly isolated strains. The reconstructed genomic datasets of the *L. crispatus* species, encompassing a total of 110 genomes sequences (i.e., 16 isolated in this study and 94 publicly available) represents the largest so far

reconstructed genetic database for this *Lactobacillus* species. These data were used to predict the pan-genome of *L. crispatus* species, i.e., the collection of genes of all strains of a species based on the cluster of orthologous groups (COGs) (185). Furthermore, these data were also used to predict the core genome, i.e. the collection of gene families shared between organisms of a given species, in this case the *L. crispatus* taxon (186). Plotting on a log-log scale as a function of the number of analyzed genomes, the pan-genome size was determined to consist of 20,247 COGs, suggesting that the power trend line has not reached a plateau (Fig. 2). The number of new genes discovered by sequential addition of genome sequences decreased from 594 COGs for the first two genomes, to 123 COGs for the final addition. These data suggests the existence of an open pan-genome within the *L. crispatus* species (Fig. 2), as already noted for other *Lactobacillus* species, including *L. gasseri*, *Lactobacillus paragasseri* and *L. helveticus* (354, 355). Moreover, the 110 *L. crispatus* genomes were examined to identify shared orthologous genes, as well as unique genes. *In silico,* analyses revealed 452 ORFs shared among all *L. crispatus* genomes analyzed, representing the core genome of this taxon. In addition, we identified a variable number of Truly Unique Genes (TUGs), which ranged from 233 of *L. crispatus* LR31 and *L. crispatus* LR33 to 29 of *L. crispatus* MGYG-HGUT-02348 and *L. crispatus* EM-LC1 (Fig. 3 and Table S2). Notably, due to the low quality of some genome sequence data, strains presenting more than 250 TUGs were not considered for this analysis (Table S2). Furthermore, an *in silico* approach was used to calculate the average nucleotide identity (ANI) values, defined as a measure of nucleotide-level genomic similarity between two genomes, between *L. crispatus* genomes (188). This analysis showed a highly syntenic genome structure among members of this species, with associated ANI values ranging from 95.57 % to 99.87 %. Interestingly, different range of ANI values were recognized between *L. crispatus* strains isolated from vaginal fluid and from poultry fecal samples. In this context, the lowest ANI value between vaginal isolated strains was 95.57 %, while for poultry isolated strains was 97.64 %. These data suggests that differences exist between *L. crispatus* strains isolated from different ecological niches, underlining highly syntenic genome structures among *L. crispatus* strains of poultry origin.

Phylogenetic analyses of the *Lactobacillus crispatus* species.

Recently, a taxonomic reclassification of the *Lactobacillus* genus was reflected in a more robust clade of this bacterial taxon, including microorganisms with shared metabolic and ecological features (344). Notably, these analyses revealed the presence in the *Lactobacillus* genus of species adapted to vertebrates or invertebrates, including *L. crispatus* species and typical human vaginal lactobacilli, such as *L. iners*, *L. jensenii* and *L. gasseri* species (344). The availability of genome sequences of members of *L. crispatus* species (Table S2), allowed a more robust phylogenetic reconstruction of this taxon. As already mentioned above, *in silico* analyses identified 452 ORFs shared among *L. crispatus* genomes analyzed, representing the core genome of this species. A concatenated protein sequence that included the product of each of these core genes was used to build a *L. crispatus* supertree (Fig. 3). This supertree showed that all 16 *L. crispatus* strains isolated in this study co-clustered with other publicly available *L. crispatus* genomes. Interestingly, the supertree gave rise to two clades, the first one composed by *L. crispatus* strains isolated from chicken as well as turkey GIT and the human GIT, while the second one included *L. crispatus* strains isolated from vaginal swabs and human urine samples (Fig. 3). Interestingly, *L. crispatus* DISK12 strain, which was isolated from human oral cavity, and the type strain *L. crispatus* DSM20584, which was isolated from human eye (356), co-clustered with strains isolated from the GIT of animals and humans (Fig. 3). Notably, strains CIP104459, C037 and OAB24-B, which were isolated from vaginal swabs (357, 358), did not cluster with other vaginal-isolated strains (Fig. 3). Moreover, three strains isolated from human urine samples, i.e., *L. crispatus* UMB0824, *L. crispatus* UMB0085 and *L. crispatus* UMB0040 (359), co-clustered with strains of animal and human GIT origin (Fig. 3). Despite some exceptions, due to the very low phylogenetic distance among members of *L. crispatus* species, these analyses showed different evolutionary developments amongst the strains analyzed, likely reflecting the ecological adaptation of such strains to their ecological niches.

Co-culture experiments in Simulated Vaginal Fluid (SVF).

In order to investigate the behavior of vaginal-isolated *L. crispatus* strains in a simplified VM composed by other typical vaginal-*Lactobacillus* species, we performed growth experiments on SVF that mimics the nutritional and chemical-physical conditions of the vaginal environment, as previously performed by Pan et al. (319). We assessed nine different parallel *in vitro* experiments, in which a single vaginal isolated *L. crispatus* strain was inoculated (1 % vol/vol) with a simple VM composed by *L. gasseri* ATCC 9857, *L. iners* LMG 14328 and *L. jensenii* V94G at 1 % (vol /vol) each. After 12h of growth, shallow-shotgun metagenomics was independently performed for each batch culture. Sequencing analyses resulted in a total of 837,443 reads with an average of 93,049 ± 8,520 reads per sample (Fig. 4 and Table S6).

Notably, after 12h the average relative abundance of *L. iners* LMG 14328 corresponded to 0.60 % ± 0.26 % (Fig. 4). This low abundance of *L. iners* LMG 14328 was probably due to the suboptimal growth conditions for this strain, perhaps as a result of its nutritional needs (360). Interestingly, *L. gasseri* ATCC 9857 seemed to take over other strains in seven out of nine experiments (Fig. 4) and in five cases, the relative abundance of this strain was greater than 65 % (Fig. 4). Moreover, in one case *L. crispatus* PRL2021 was shown to have grown better compared to the other lactobacilli, with a relative abundance of 77.4 % (Fig. 4). Other *L. crispatus* strains (e.g. LB57, LB58, LB59 and M247) were shown to be dramatically outnumbered by other lactobacilli after 12h of cultivation (Fig. 4).

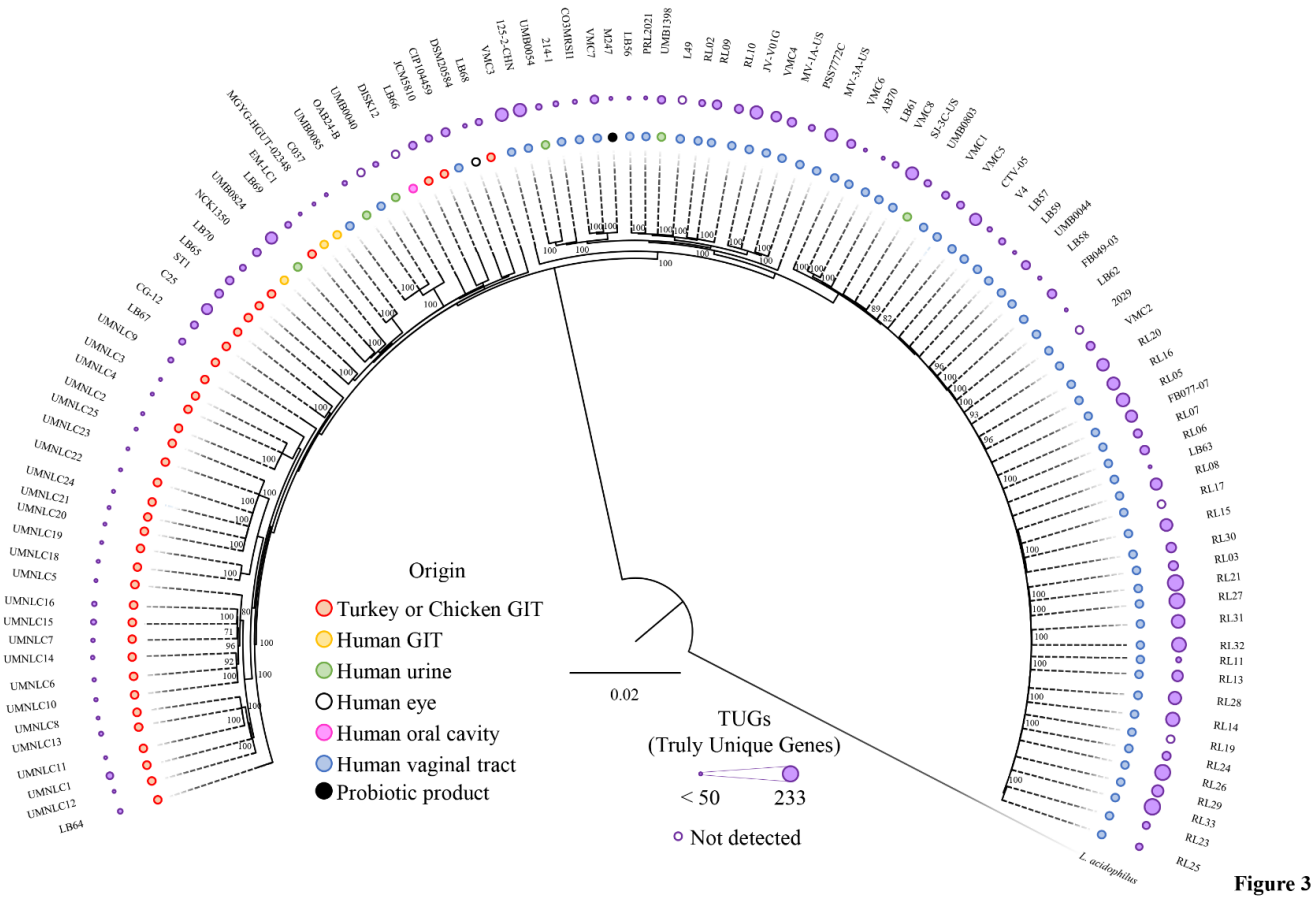


Figure 3

Figure 3. Phylogenomic tree of *L. crispatus* species. A proteomic tree was built on the concatenation of 452 *L. crispatus* core genes identified in the core genome analysis. This tree is constructed by the neighbor-joining method, and the genome sequence of *L. acidophilus* DSM 20079 was used as outgroup. Bootstrap percentage > 70 are show at node points, based on 1,000 replicates. Circles surrounding the tree represent the isolation origin of the strains and the number of TUGs (violet circles).

These findings suggest that *L. crispatus* PRL2021 possesses superior ecological fitness compared to the other assessed *L. crispatus* strains, and its growth performance highlights excellent adaptation to the vaginal environment.

Co-culture experiments using *L. crispatus* PRL2021 and human faeces.

The vaginal microbiota composition may be influenced by fecal/rectal bacteria since it is the closest anatomical site characterized by a complex microbiota. In this context, the fecal/rectal microbiota

may act as a source or reservoir of microorganisms (361). In fact, transmission of microorganisms from the rectum to the vagina may change the equilibrium of the vaginal microbiota or stimulate local inflammatory responses predisposing women to local infections, such as BV (362). Notably, recent studies have suggested that oral ingestion of *Lactobacillus*-containing products may play a possible role in the treatment or prevention of genital infections, acting as a gut-reservoir for vaginal (re)colonization by lactobacilli (363, 364). In order to evaluate the ecological effects of *L. crispatus* PRL2021, which was shown to have properties that were superior amongst the *L. crispatus* strains tested here, in shaping the vaginal microbiota, we performed a co-culture experiment of PRL2021 with human feces. In detail, we cultivated *L. crispatus* PRL2021 with four different fecal samples with an inoculum consisting of 1 % (wt/vol) or 5 % (wt/vol). The growth experiments were performed using both SVF and Fecal media. Control growths were carried out without *L. crispatus* PRL2021. The ecological effects of strain PRL2021 on the fecal microbiota in the co-culture experiments were investigated through a shallow shotgun metagenomics approach followed by an absolute quantification of the identified bacteria through flow-cytometry assays. Sequencing analyses resulted in a total of 1,566,052 reads with an average of 43,501 ± 15,086 reads per sample (Table S7). As expected, the metagenomic analysis highlighted the absence of *L. crispatus* in all four fecal samples and in control growths. In detail, samples grown in SFV and Fecal media with 1 % of stool sample showed an average abundance of *L. crispatus* of 12.33 % ± 8.00 % and 25.08 % ± 21.27 %, respectively. Similarly, the samples cultivated with 5 % of human stool sample indicated an average abundance of *L. crispatus* of 4.46 % ± 3.83 % and 7.40 % ± 9.04 % in SFV and Fecal media, respectively.

The quantitative microbiome profiling approach, employing flow cytometric enumeration of microbial cells present in each sample, allowed the identification of the microbial load of co-culture and control growths (287, 320). The comparison between each co-culture experiments with its own control showed a decrease of the number of microbial cells in all samples enriched with *L. crispatus* PRL2021 (Figure 4). More specifically, flow cytometric results showed a drop of cell number in the

co-culture samples ranging from a thousand- to hundred thousand-fold (Figure 4), highlighting the ability of *L. crispatus* PRL2021 to provoke a simplification of the bacterial community, such as the fecal or vaginal microbiota. Moreover, these results seem to show the capability of *L. crispatus* PRL2021 to grow even in a different and more complex environment than that of the vaginal ecosystem, such as the intestinal environment, making this strain an interesting candidate for the development of *Lactobacillus*-containing probiotic products.

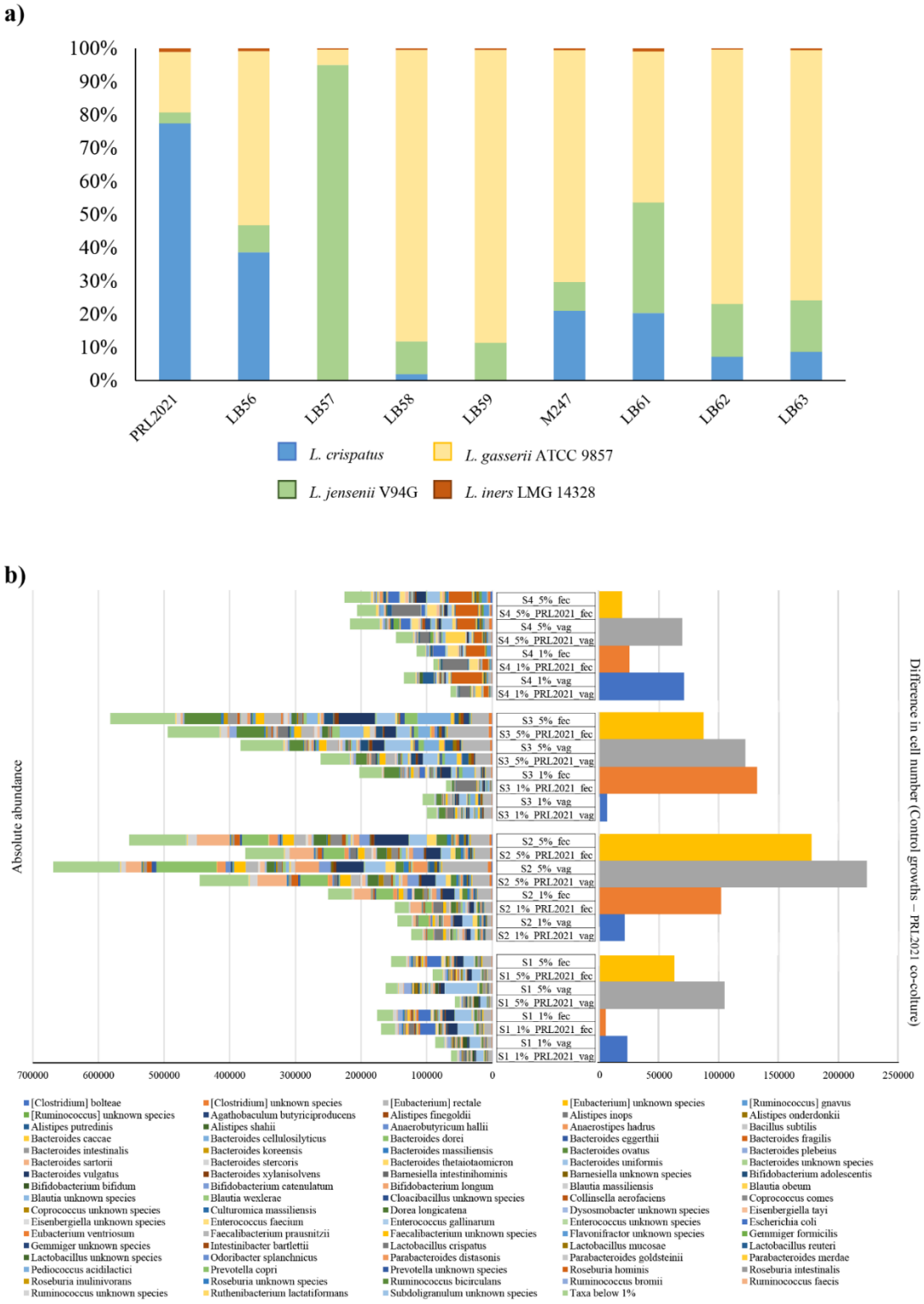


Figure 4

Figure 4. Growth experiment in Simulated Vaginal Fluid (SVF) and Fecal media. Panel a shows bacterial composition of batch-cultures based on shallow shotgun metagenomics profiling. The *x*-axis represents the different analyzed batch-cultures, indicate with the *L. crispatus* strain name tested in each experiment. The *y*-axis represents the absolute percentage of OTUs for each sample. Panel b reveals the taxonomic absolute abundances and the comparisons between the co-culture experiments and its own control.

CONCLUSIONS

The human VM is represented by bacteria colonizing the human vaginal tract and that are assumed to exploit an important role for the healthy status of the host. In this study, we analyzed the VM composition of healthy women and observed that *L. crispatus* is one of the most representative species and correlates negatively with bacteria involved in vaginal infections. Moreover, this species seems to be able to modulate the vaginal microbiota and appears to exploit a role in reducing biodiversity of VM. However, in this metagenomics analysis, 49 % of the samples are of Chinese origin and this could represent a limitation of this study because it can introduce a geographical bias. Certainly, a more in-depth investigation that includes multiple samples of different geographic origins could expand and refine the knowledge on the composition of CST and the healthy VM. Furthermore, the isolation of 15 *L. crispatus* strains from vaginal swabs and chicken fecal samples confirmed that these two ecological niches are frequently and abundantly colonized by members of this *Lactobacillus* species. The genome sequencing of the new isolates allowed reconstructing the largest genomic data set of the *L. crispatus* taxa, including 94 publicly available *L. crispatus* genomes. Phylogenomic analyses based on the core genome sequences showed the existence of two different clades in which, with just a few exceptions, *L. crispatus* strains reflected the ecological niche from which they were isolated, suggesting a genetic adaptation to different environments. Furthermore, VM simulated *in vitro* experiments revealed that *L. crispatus*, and in particular PRL2021 strain, is able to take over other typical vaginal *Lactobacillus* species and that possessed high growth capacity in the fecal environment, also highlighting the ability of this microbial taxon to decrease the complexity of a bacterial community. This analysis and *in silico* prediction of an extensive putative antimicrobial activity of this species suggested that *L. crispatus* PRL2021 could be consider as an interesting strain for preventing vaginosis and vaginal dysbiosis.

SUPPLEMENTAL MATERIALS

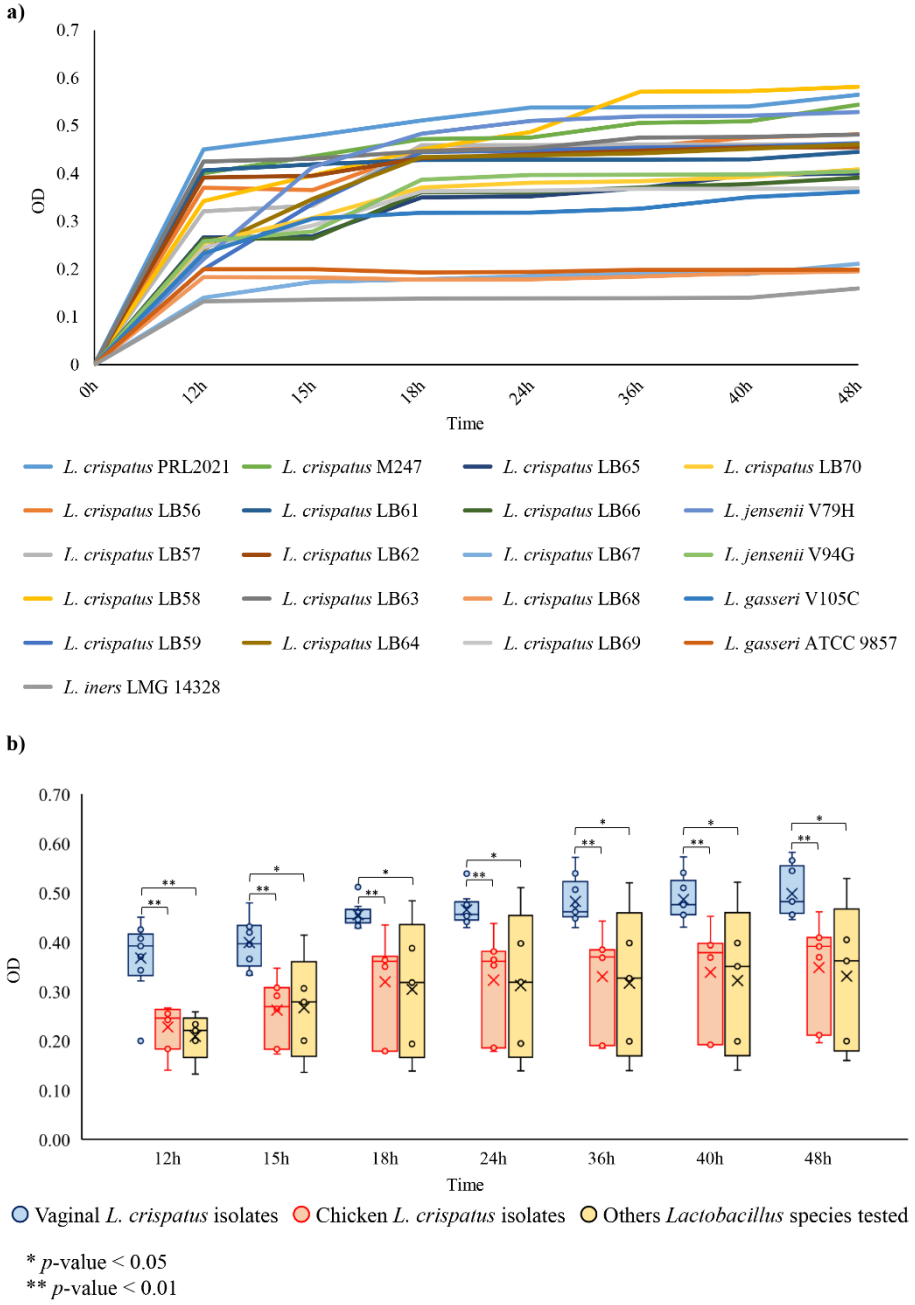


Figure S1

Figure S1. Growth curves of *Lactobacillus* strains with glycogen as sole carbon source. Panel a exhibits the growth curves of tested strains. On the *x*-axis OD values are reported at a wavelength of 600 nm, whereas the *y*-axis showed times (h). Each *Lactobacillus* strain is indicated with a different color line as reported in the legend. Panel b shows a Whiskers plot based on *Lactobacillus* growth with glycogen as sole carbon source. The *x*-axis represents the OD600, while the *y*-axis shows the times (h)- the boxes represent 50% of the data set, distributed between the first and the third quartiles. The median divides the boxes into the interquartile range, while the "X" represents the mean. Statistical analysis were performed through Student t-test between vaginal *L. crispatus* isolates and chicken *L. crispatus* isolates and between vaginal origin *L. crispatus* compare to others *Lactobacillus* species tested.

Table S1. Samples used in this experiment

Sample Name	BioProject	Nationality	Number of final reads
ERR2238774	PRJEB24147	China	10910
ERR2238775	PRJEB24147	China	85474
ERR2238777	PRJEB24147	China	44930
ERR2244420	PRJEB24147	China	43748
ERR2244423	PRJEB24147	China	27611
ERR2244501	PRJEB24147	China	61550
ERR2244503	PRJEB24147	China	54340
ERR2244505	PRJEB24147	China	72466
ERR2244507	PRJEB24147	China	55382
ERR2244509	PRJEB24147	China	41363
ERR2244511	PRJEB24147	China	37004
ERR2244513	PRJEB24147	China	46358
ERR2244516	PRJEB24147	China	19470
ERR2244518	PRJEB24147	China	24954
ERR2244520	PRJEB24147	China	51563
ERR2244522	PRJEB24147	China	55004
ERR2244524	PRJEB24147	China	16445
ERR2244526	PRJEB24147	China	58603
ERR2244528	PRJEB24147	China	61207
ERR2244530	PRJEB24147	China	88352
ERR2244532	PRJEB24147	China	63261
ERR2244534	PRJEB24147	China	65063
ERR2244536	PRJEB24147	China	50669
ERR2244538	PRJEB24147	China	13506
ERR2244540	PRJEB24147	China	7445
ERR2244542	PRJEB24147	China	8851
ERR2244515	PRJEB24147	China	93793
SRR5946030	PRJNA361427	China	33366
SRR5946031	PRJNA361427	China	28645
SRR5946032	PRJNA361427	China	33651
SRR5946033	PRJNA361427	China	39511
SRR5946034	PRJNA361427	China	33027
SRR5946035	PRJNA361427	China	40517
SRR5946036	PRJNA361427	China	38330
SRR5946037	PRJNA361427	China	34633
SRR5946038	PRJNA361427	China	37450
SRR5946039	PRJNA361427	China	42444
SRR5946040	PRJNA361427	China	32400
SRR5946041	PRJNA361427	China	34630
SRR5946042	PRJNA361427	China	83767

SRR5946043	PRJNA361427	China	34244
SRR5946044	PRJNA361427	China	34904
SRR5946045	PRJNA361427	China	31973
SRR5946046	PRJNA361427	China	31721
SRR5946047	PRJNA361427	China	40130
SRR5946048	PRJNA361427	China	40071
SRR064532	PRJNA48479	-	2867
SRR1031541	PRJNA48479	-	24936
SRR1031704	PRJNA48479	-	4193
SRR1031942	PRJNA48479	-	1430
SRR1031948	PRJNA48479	-	2871
SRR1175002	PRJNA48479	-	9164
SRR1564338	PRJNA48479	-	529
SRR1565338	PRJNA48479	-	4856
SRR1565461	PRJNA48479	-	8273
SRR1565840	PRJNA48479	-	346
SRR346709	PRJNA48479	-	2150
SRR512770	PRJNA48479	-	32231
SRR513151	PRJNA48479	-	3769
SRR513452	PRJNA48479	-	38455
SRR513769	PRJNA48479	-	3466
SRR513782	PRJNA48479	-	34528
SRR513785	PRJNA48479	-	15869
SRR513792	PRJNA48479	-	1415
SRR513805	PRJNA48479	-	3508
SRR513807	PRJNA48479	-	2362
SRR513811	PRJNA48479	-	1842
SRR513818	PRJNA48479	-	2525
SRR513825	PRJNA48479	-	3247
SRR514004	PRJNA48479	-	10722
SRR514205	PRJNA48479	-	4157
SRR514207	PRJNA48479	-	3357
SRR514213	PRJNA48479	-	1151
SRR514217	PRJNA48479	-	2505
SRR514260	PRJNA48479	-	970
SRR514267	PRJNA48479	-	5865
SRR514272	PRJNA48479	-	2159
SRR514310	PRJNA48479	-	2840
SRR514318	PRJNA48479	-	5846
SRR514847	PRJNA48479	-	2578
SRR514852	PRJNA48479	-	5993
SRR520269	PRJNA48479	-	11663
SRR520270	PRJNA48479	-	3444
SRR539903	PRJNA48479	-	1621
SRR540620	PRJNA48479	-	7043

SRR545403	PRJNA48479	-	14909
SRR545443	PRJNA48479	-	1952
SRR549537	PRJNA48479	-	32565
SRR628268	PRJNA48479	-	8592
SRR628269	PRJNA48479	-	7403
SRR628270	PRJNA48479	-	2795
SRR650075	PRJNA48479	-	2288
SRR650082	PRJNA48479	-	21598
SRR650085	PRJNA48479	-	8076

Table S2. *Lactobacillus crispatus* strains used for the comparative genomic analysis

Strains	Origin	Scaffolds	Assembly	Genome Size (Mb)	TUGs	Bacteriocin-encoding genes
L. crispatus 125-2-CHN	Vaginal swab	30	GCA_00016225.1	2,3	189	8
L. crispatus 2029	Vaginal swab	295	GCA_000466885.2	2,19	---	10
L. crispatus 214.1	Vaginal swab	187	GCA_000177575.1	2,07	74	6
L. crispatus LB56	Vaginal swab	195	This study	2,16	39	3
L. crispatus LB64	Chicken fecal sample	36	This study	2,05	73	5
L. crispatus LB65	Chicken fecal sample	44	This study	2,06	115	4
L. crispatus AB70	Vaginal swab	2	GCA_003971565.1	2,37	40	5
L. crispatus C037	Vaginal swab	96	GCA_001700475.1	2,15	39	8
L. crispatus C25	Chicken fecal sample	97	GCA_001704465.1	2,34	159	10
L. crispatus CG-12	Chicken fecal sample	62	GCA_004334905.1	2,04	112	10
L. crispatus CIP104459	Vaginal swab	124	GCA_008079315.1	1,99	121	7
L. crispatus CO3MRSI1	Vaginal swab	1	GCA_003795065.1	2,35	61	6
L. crispatus CTV-05	Vaginal swab	25	GCA_000165885.1	2,36	169	8
L. crispatus DISK12	Oral cavity	450	GCA_002811165.1	2,51	---	5
L. crispatus DSM20584	Human eye	150	GCA_001434005.1	2,06	51	9
L. crispatus EM-LC1	Human fecal sample	63	GCA_000497065.1	1,83	29	6
L. crispatus LB66	Chicken fecal sample	44	This study	2,04	117	4
L. crispatus FB049-03	Vaginal swab	5	GCA_000301115.1	2,46	129	7
L. crispatus FB077-07	Vaginal swab	10	GCA_000301135.1	2,7	172	8
L. crispatus LB67	Chicken fecal sample	60	This study	1,97	91	6
L. crispatus JCM 5810	Chicken fecal sample	7	GCA_001567095.1	2,05	100	8
L. crispatus JV-V01	Vaginal swab	86	GCA_000160515.1	2,22	186	5
L. crispatus L49	Vaginal swab	671	GCA_004103355.1	2,52	---	6
L. crispatus PRL2021	Vaginal swab	43	This study	2,33	38	5
L. crispatus LB57	Vaginal swab	259	This study	2,38	93	9
L. crispatus LB58	Vaginal swab	199	This study	2,24	58	5
L. crispatus LB59	Vaginal swab	189	This study	2,13	44	4
L. crispatus M247	Vaginal swab	156	This study	2,10	49	5
L. crispatus MGYG-HGUT-02348	Human fecal sample	63	GCA_902386155.1	1,83	29	6
L. crispatus MV-1A-US	Vaginal swab	7	GCA_000161915.2	2,25	133	5
L. crispatus MV-3A-US	Vaginal swab	76	GCA_000162315.1	2,44	185	7
L. crispatus NCK1350	Human gut	42	GCA_007713895.1	2,05	120	5
L. crispatus OAB24-B	Vaginal swab	752	GCA_001700485.1	2,01	---	6
L. crispatus LB68	Chicken fecal sample	111	This study	1,97	104	4
L. crispatus LB69	Chicken fecal sample	95	This study	1,98	93	4
L. crispatus LB70	Chicken fecal sample	199	This study	2,02	97	9
L. crispatus PSS7772C	Vaginal swab	215	GCA_001563615.1	2,39	94	6
L. crispatus RL02	Vaginal swab	291	GCA_004361375.1	2,24	92	6

L. crispatus RL03	Vaginal swab	330	GCA_004361385.1	2,55	143	5
L. crispatus RL05	Vaginal swab	431	GCA_004361345.1	2,58	198	5
L. crispatus RL06	Vaginal swab	371	GCA_004361635.1	2,19	133	6
L. crispatus RL07	Vaginal swab	387	GCA_004361355.1	2,18	125	7
L. crispatus RL08	Vaginal swab	364	GCA_004361315.1	2,3	174	5
L. crispatus RL09	Vaginal swab	289	GCA_004361575.1	2,28	131	6
L. crispatus RL10	Vaginal swab	297	GCA_004361565.1	2,32	121	6
L. crispatus RL11	Vaginal swab	280	GCA_004361295.1	2,19	78	7
L. crispatus RL13	Vaginal swab	434	GCA_004361265.1	2,25	166	7
L. crispatus RL14	Vaginal swab	375	GCA_004361555.1	2,61	203	5
L. crispatus RL15	Vaginal swab	386	GCA_004361245.1	2,33	177	6
L. crispatus RL16	Vaginal swab	350	GCA_004361545.1	2,6	188	8
L. crispatus RL17	Vaginal swab	616	GCA_004361215.1	2,46	---	6
L. crispatus RL19	Vaginal swab	541	GCA_004361205.1	2,31	---	7
L. crispatus RL20	Vaginal swab	359	GCA_004361515.1	2,44	182	8
L. crispatus RL21	Vaginal swab	489	GCA_004361195.1	2,57	233	5
L. crispatus RL23	Vaginal swab	306	GCA_004361475.1	2,33	107	6
L. crispatus RL24	Vaginal swab	319	GCA_004361185.1	2,39	121	5
L. crispatus RL25	Vaginal swab	313	GCA_004361465.1	2,35	102	6
L. crispatus RL26	Vaginal swab	443	GCA_004361175.1	2,27	223	7
L. crispatus RL27	Vaginal swab	510	GCA_004361445.1	2,6	228	6
L. crispatus RL28	Vaginal swab	438	GCA_004361125.1	2,24	185	6
L. crispatus RL29	Vaginal swab	401	GCA_004361115.1	2,25	169	5
L. crispatus RL30	Vaginal swab	407	GCA_004361455.1	2,32	147	6
L. crispatus RL31	Vaginal swab	483	GCA_004361095.1	2,4	194	7
L. crispatus RL32	Vaginal swab	417	GCA_004361075.1	2,37	211	7
L. crispatus RL33	Vaginal swab	436	GCA_004361395.1	2,4	233	6
L. crispatus SJ-3C-US	Vaginal swab	70	GCA_000176975.2	2,16	189	8
L. crispatus ST1	Chicken fecal sample	1	GCA_000091765.1	2,04	128	5
L. crispatus UMB0040	Human urine sample	99	GCA_002863245.1	2	77	7
L. crispatus UMB0044	Human urine sample	200	GCA_002861775.1	2,27	123	7
L. crispatus UMB0054	Human urine sample	163	GCA_002863485.1	2,14	85	4
L. crispatus UMB0085	Human urine sample	130	GCA_002861815.1	2,17	67	9
L. crispatus UMB0803	Human urine sample	156	GCA_002861765.1	2	102	7
L. crispatus UMB0824	Human urine sample	94	GCA_002861805.1	2,17	169	6
L. crispatus UMB1389	Human urine sample	223	GCA_002863505.1	2,27	115	6
L. crispatus UMNLC10	Turkey fecal sample	54	GCA_002218805.1	2,05	101	5
L. crispatus UMNLC11	Turkey fecal sample	73	GCA_002218815.1	2,04	51	5
L. crispatus UMNLC12	Turkey fecal sample	59	GCA_002218855.1	2,08	40	5
L. crispatus UMNLC13	Turkey fecal sample	110	GCA_002218845.1	2,02	48	4
L. crispatus UMNLC14	Turkey fecal sample	81	GCA_002218885.1	1,99	61	5

L. crispatus UMNLC15	Turkey fecal sample	81	GCA_002218895.1	2,03	54	4
L. crispatus UMNLC16	Turkey fecal sample	56	GCA_002218925.1	2,13	74	4
L. crispatus UMNLC18	Turkey fecal sample	76	GCA_002218945.1	2,04	70	5
L. crispatus UMNLC19	Turkey fecal sample	75	GCA_002218965.1	2,05	45	6
L. crispatus UMNLC1	Turkey fecal sample	69	GCA_002218615.1	2,19	47	5
L. crispatus UMNLC20	Turkey fecal sample	67	GCA_002218975.1	2,04	43	5
L. crispatus UMNLC21	Turkey fecal sample	68	GCA_002219005.1	2,04	41	5
L. crispatus UMNLC22	Turkey fecal sample	67	GCA_002219045.1	2,05	40	5
L. crispatus UMNLC23	Turkey fecal sample	63	GCA_002219085.1	2,04	34	5
L. crispatus UMNLC24	Turkey fecal sample	68	GCA_002219015.1	2,05	35	4
L. crispatus UMNLC25	Turkey fecal sample	69	GCA_002219055.1	2,05	37	5
L. crispatus UMNLC2	Turkey fecal sample	73	GCA_002218565.1	2,01	43	5
L. crispatus UMNLC3	Turkey fecal sample	64	GCA_002218645.1	2,02	44	5
L. crispatus UMNLC4	Turkey fecal sample	62	GCA_002218655.1	2,02	41	5
L. crispatus UMNLC5	Turkey fecal sample	50	GCA_002218765.1	2,01	42	5
L. crispatus UMNLC6	Turkey fecal sample	49	GCA_002218695.1	2,05	54	5
L. crispatus UMNLC7	Turkey fecal sample	71	GCA_002218775.1	2,1	50	5
L. crispatus UMNLC8	Turkey fecal sample	74	GCA_002218685.1	2,05	46	5
L. crispatus UMNLC9	Turkey fecal sample	68	GCA_002218735.1	1,95	79	6
L. crispatus V4	Vaginal swab	252	GCA_004681235.1	2,09	72	7
L. crispatus LB61	Vaginal swab	140	This study	2,08	38	4
L. crispatus LB62	Vaginal swab	172	This study	2,16	38	5
L. crispatus LB63	Vaginal swab	182	This study	2,13	43	6
L. crispatus VMC1	Vaginal swab	267	GCA_001546015.1	2,07	108	7
L. crispatus VMC2	Vaginal swab	258	GCA_001546025.1	2,2	128	10
L. crispatus VMC3	Vaginal swab	91	GCA_001541385.1	2,2	185	8
L. crispatus VMC4	Vaginal swab	244	GCA_001541405.1	2,31	148	6
L. crispatus VMC5	Vaginal swab	259	GCA_001541515.1	2,24	110	9
L. crispatus VMC6	Vaginal swab	253	GCA_001541505.1	2,34	123	6
L. crispatus VMC7	Vaginal swab	247	GCA_001541535.1	2,1	115	7
L. crispatus VMC8	Vaginal swab	255	GCA_001541585.1	2,33	91	5

Table S3. BLAST analysis of the glgX gene of *E. coli* in *L. crispatus* strains isolated.

Strain	Origin	Reference gene	Identity (%)	E-value	ORF
L. crispatus PRL2021	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	21,51	5,00E-08	PRL2021_0328
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	PRL2021_1423
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	PRL2021_2146
L. crispatus LB56	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	21,51	5,00E-08	LB56_0244
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB56_1077
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB56_1820
L. crispatus LB57	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB57_1200
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB57_2119
L. crispatus LB58	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB58_0305
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB58_1117
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB58_1832
L. crispatus LB59	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB59_0267
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB59_1082
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB59_1932
L. crispatus M247	Probiotic product	AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	-
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	-
L. crispatus LB61	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB61_0248
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB61_1082
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB61_1877
L. crispatus LB62	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB62_0273
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB62_1114
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB62_2023
L. crispatus LB63	Vaginal tract	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB63_0271
		AIZ89569.1_glgX_Escherichia_coli	23,53	9,00E-07	LB63_1134
		AIZ89569.1_glgX_Escherichia_coli	25,26	5,00E-12	LB63_1854
L. crispatus LB64	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB64_0301
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB64_1815
L. crispatus LB65	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB65_0233
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB65_1861
L. crispatus LB66	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB66_0314
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB66_1834
L. crispatus LB67	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	22,01	5,00E-08	LB67_0199
		AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB67_1714
L. crispatus LB68	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	25,52	2,00E-12	LB68_1694
L. crispatus LB69	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	23,46	6,00E-11	LB69_1663
		AIZ89569.1_glgX_Escherichia_coli	25,26	4,00E-12	LB69_1810
L. crispatus LB70	Chicken faecal sample	AIZ89569.1_glgX_Escherichia_coli	25,26	4,00E-12	LB70_1759

Table S4. Bacteriocin-encoding genes identified in new isolated *L. crispatus* strains.

Strain	Origin	Bacteriocin	Bagel Subject	Bacteriocin class
PRL2021	Vaginal	Bacteriocin LS2	bacteriocin_LS2chainb	Class II
		Bacteriocin LS2	bacteriocin_LS2chaina	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB56	Vaginal	Bacteriocin class II with double-glycine	bacteriocin_LS2chaina	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
LB57	Vaginal	Penocin	Penocin_A	Class II
		Bacteriocin LS2	bacteriocin_LS2chainb	Class II
		Bacteriocin LS2	bacteriocin_LS2chaina	Class II
		Penocin	Penocin_A	Class II
		Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
		Helveticin	Helveticin-J	Class III
LB58	Vaginal	Penocin	Penocin_A	Class II
		Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
LB59	Vaginal	Penocin	Penocin_A	Class II
		Helveticin	Helveticin	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
M247	Probiotic product	Penocin	Penocin_A	Class II
		Bacteriocin class II with double-glycine	bacteriocin_LS2chaina	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB61	Vaginal	Penocin	Penocin_A	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB62	Vaginal	Penocin	Penocin_A	Class II
		Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III

LB63	Vaginal	Penocin	Penocin_A	Class II
		Penocin	Penocin_A	Class II
		Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB64	Chicken GIT	Helveticin	Helveticin-J	Class III
		Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
LB65	Chicken GIT	M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB66	Chicken GIT	Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
LB67	Chicken GIT	Garvicin	Putative_bacteriocin	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
		Helveticin	Helveticin-J	Class III
LB68	Chicken GIT	M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III
LB69	Chicken GIT	Helveticin	Helveticin-J	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
LB70	Chicken GIT	Garvicin	Putative_bacteriocin	Class II
		Bacteriocin LS2	bacteriocin_LS2chainb	Class II
		Bacteriocin LS2	bacteriocin_LS2chaina	Class II
		Microcin	Amylovorin	Class II
		Lactacin	Lactacin_F,subunit_lafA	Class II
		M23 family metallopeptidase	Enterolysin_A	Class III
		M23 family metallopeptidase	Enterolysin_A	Class III
		Helveticin	Bacteriocin_helveticin_J	Class III
		Helveticin	Helveticin-J	Class III

Table S5. Bacteriocin encoding-genes predicted in *L. gasseri*, *L. jensenii* and *L. iners* species.

Strain	Predicted bacteriocin encoding-genes
L. gasseri ATCC 33323	2
L. gasseri MV-22	2
L. gasseri 202-4	2
L. gasseri SJ-9E-US	2
L. gasseri SV-16A-US	2
L. gasseri 224-1	2
L. gasseri CECT 5714	2
L. gasseri 2016	2
L. gasseri 240_LCRI	2
L. gasseri 249_LKEF	2
L. gasseri 497_LGAS	6
L. gasseri 459_LHEL	2
L. gasseri 770_LJOH	2
L. gasseri 987_LJOH	4
L. gasseri PSS7772D	2
L. gasseri 505	7
L. gasseri AL5	4
L. gasseri AL3	5
L. gasseri 4M13	6
L. gasseri DSM 14869	1
L. gasseri UMB0045	1
L. gasseri UMB0670	2
L. gasseri UMB0099	7
L. gasseri UMB0056	4
L. gasseri UMB0045b	1
L. gasseri JCM 1025	1
L. gasseri JG141	8
L. gasseri TF08-1	2
L. gasseri 1001175st1_D8	2
L. gasseri 7171	1
L. gasseri 7135	4
L. gasseri UMB1196	1
L. gasseri UMB2965	3
L. gasseri UMB0607	2
L. gasseri UMB4205	1
L. gasseri UMB3077	3
L. gasseri JCM 1131	2
L. gasseri BIO6369	2
L. gasseri NCTC13722	2
L. gasseri SV_Bg7063_mod2	1
L. gasseri MGYG-HGUT-02387	6

L. gasseri MGYG-HGUT-03690	2
L. jensenii 1153	0
L. jensenii JV-V16	0
L. jensenii 27-2-US	0
L. jensenii SJ-7A-US	1
L. jensenii 115-3-CHN	0
L. jensenii 269-3	0
L. jensenii MD IIE-70(2)	1
L. jensenii IM11	0
L. jensenii IM18-1	1
L. jensenii IM59	0
L. jensenii IM18-3	1
L. jensenii IM1	0
L. jensenii IM3	0
L. jensenii DSM 20557	0
L. jensenii TL2937	1
L. jensenii SNUV360	0
L. jensenii UMB0077	1
L. jensenii UMB0007	0
L. jensenii UMB0732	0
L. jensenii UMB0037	1
L. jensenii UMB8489	1
L. jensenii UMB0055	1
L. jensenii UMB0034	1
L. jensenii UMB1355	1
L. jensenii UMB1307	0
L. jensenii UMB1303	0
L. jensenii UMB1165	1
L. jensenii UMB8651	1
L. jensenii UMB8840	0
L. jensenii UMB4685	0
L. jensenii UMB3442	1
L. jensenii UMB572	0
L. jensenii UMB246	1
L. jensenii UMB4707	0
L. jensenii UMB639	0
L. jensenii FDAARGOS_749	0
L. jensenii VA04-2AN	0
L. jensenii M166	1
L. jensenii M167	1
L. jensenii UMB00847	0
L. jensenii UMB0836	0
L. jensenii UMB7766	0

L. jensenii MGYG-HGUT-02313	1
L. iners 11V1-d	2
L. iners 09V1-c	0
L. iners 03V1-c	0
L. iners 01V1-a	0
L. iners SPIN 2503V10-D	0
L. iners DSM 13335	0
L. iners AB-1	0
L. iners LEAF 2053A-b	2
L. iners LEAF 2052A-d	0
L. iners LEAF 2062A-h1	0
L. iners LEAF 3008A-a	0
L. iners ATCC 55195	0
L. iners UPII 143-D	0
L. iners UPII 60-B	2
L. iners SPIN 2503V10-B	2
L. iners 7_1_47FAA	0
L. iners LMG 18914	0
L. iners UMB0033	0
L. iners UMB1051	2
L. iners UMB0030	0
L. iners KA00186	0
L. iners LI335	0
L. iners Indica1	0
L. iners KY	0
L. iners C0094A1	2
L. iners C0059G1	2
L. iners C0011D1	0
L. iners C0210C1	2
L. iners C0322A1	0
L. iners C0254C1	0
L. iners MGYG-HGUT-01383	0

Table S6. Taxonomic recontruction of microbial population of the co-culture samples

	PRL2021	**LB56**	**LB57**	**LB58**	**LB59**	**M247**	**LB61**	**LB62**	**LB63**
Filtered reads	99781	94978	99733	78857	99748	82397	99438	97582	84929
Lactobacillus crispatus	77,4%	38,6%	0,06%	1,9%	0,03%	21,0%	20,36%	7,1%	8,6%
Lactobacillus jensenii V94G	3,2%	8,2%	94,96%	9,8%	11,28%	8,6%	33,24%	15,9%	15,6%
Lactobacillus gasseri ATCC 9857	18,2%	52,4%	4,58%	87,8%	88,28%	69,9%	45,51%	76,5%	75,3%
Lactobacillus iners LMG 14328	1,1%	0,8%	0,40%	0,5%	0,41%	0,5%	0,89%	0,4%	0,5%

Table S7. Shallow shotgun metagenomics data

	Filtered reads
S1	30573
S1_1%_PRL2021_vag	41322
S1_1%_vag	51854
S1_1%_PRL2021_fec	59203
S1_1%_fec	36119
S1_5%_PRL2021_vag	69306
S1_5%_vag	69329
S1_5%_PRL2021_fec	42155
S1_5%_fec	26668
S2	69217
S2_1%_PRL2021_vag	50707
S2_1%_vag	42040
S2_1%_PRL2021_fec	40628
S2_1%_fec	30976
S2_5%_PRL2021_vag	50700
S2_5%_vag	57575
S2_5%_PRL2021_fec	42459
S2_5%_fec	34748
S3	49870
S3_1%_PRL2021_vag	35088
S3_1%_vag	54149
S3_1%_PRL2021_fec	31245
S3_1%_fec	27824
S3_5%_PRL2021_vag	46585
S3_5%_vag	47981
S3_5%_PRL2021_fec	32784
S3_5%_fec	26958
S4	38531
S4_1%_PRL2021_vag	75292
S4_1%_vag	51495
S4_1%_PRL2021_fec	46333
S4_1%_fec	18571
S4_5%_PRL2021_vag	67402
S4_5%_vag	26390
S4_5%_PRL2021_fec	26435
S4_5%_fec	17540

Chapter 8
General Conclusion

This thesis aimed to investigate various genetic and functional characteristics of few important members of the human microbiota. In this context, the biology of the genus *Bifidobacterium* was analyzed, with particular attention to ecological aspects, which were studied using a combination of comparative and functional genomic approaches. Particular emphasis was placed on genomic variability achieved through horizontal and vertical routes, such as HGT events and vertical transmission between mothers and babies. Furthermore, the impact of antibiotic molecules on bifidobacteria is investigated through *in silico* and *in vitro* experiments. Moreover, the characteristic of another genus involved in human microbiota, such as *Lactobacillus* genus, has been considered. In particular, the composition of the vaginal microbiota and the role of *L. crispatus* species on the health status of vaginal environment has been examined.

8.1 The genetic differences between *B. animalis* subsp. *animalis* and *B. animalis* subsp. *lactis* taxa.

Among the recognized bifidobacterial species, different studies demonstrated the health-promoting characteristics of *B. animalis* species (365-368). Members of this bifidobacterial species are nowadays used as bacterial ingredients in different functional foods (367, 368). In chapter 3 of this thesis, a comparative genomic analysis of *B. animalis* species, including both *B. animalis* subsp. *animalis* and *B. animalis* subsp *lactis* taxa, is reported. This analysis revealed the existence of a closed pan-genome, in fact, 1098 genes are in common between the 56 strains analyzed, representing the *B. animalis* core genome. A detailed genetic analysis evidences the existence of subspecies-specific genes, representing the only real genetic difference between these two taxa. Besides, culture experiments using various sugars as sole carbon source revealed the ability of *B. animalis* subsp. *animalis* taxon to metabolize a larger set of sugars in respect to *B. animalis* subsp. *lactis* strains. These data were confirmed by the large number of genes involved in carbohydrate transport and metabolism found in the *B. animalis* subsp. *animalis* subspecies specific core genome. In conclusion, the

phylogenetic and the genomic analysis confirmed the evolutionary differentiation between these two subspecies, confirmed also by culture-experiments on different carbon sources.

8.2 The DNA bifidobacterial vertical transmission in an animal model.

Different studies have evidenced that the first colonizers of the newborns gut harbor bacteria that are vertically transmitted from the mother (151, 369-371). In particular, has been hypnotized that different bifidobacterial species are transmitted from mother to newborns (107, 151, 215). Nowadays, different studies focused on the vertical transmission route of bacteria between mothers and babies, but the transfer mechanism is not still entirely clear. In chapter 4, the DNA bifidobacterial transmission from female rats (*Rattus norvegicus*) to their puppies was discovered, after that three bifidobacterial strains, i.e., *Bifidobacterium longum*, *Bifidobacterium bifidum* and *Bifidobacterium breve* strains were administered to the pregnant rats. Interestingly, these data underlined that the DNA transmission is influenced by the co-presence of different bifidobacterial strains, suggesting a co-operative behavior of bifidobacterial strains diffused. Moreover, the presence of bifidobacterial DNA in the gut of newborns after Caesarian delivery suggested the existence of a pre-term microbiota inherited from the mother's gut microbial community. Nevertheless, the non-isolation of viable bifidobacterial cells from fetal samples cannot support the existence of a pre-term transmission of lively strains.

8.3 The safety of the *Bifidobacterium* genus.

In chapter 5, an *in silico* analysis concerning the resistome and the mobilome of the *Bifidobacterium* genus was discussed. Since bifidobacteria are recognized as health promoting bacteria, the safety aspect is very important. The study demonstrated that only a limited number of bifidobacterial genes are putatively involved in Antibiotic Resistance (AR) mechanisms. Interestingly, the AR bifidobacterial genetic arsenal seemed to be less complex in respect to other Gram-positive resistome,

as well as members of *Lactobacillus* genus or members of *Bacillus* genus (249, 251, 260, 261, 281, 282). Furthermore, the combination of the gathered data of *Bifidobacterium* resistome and mobilome revealed that only 20 % of the analyzed genes could be transferred to other microorganisms, representing Mobile Genetic Hotspots (MGHs) that could be involved in the HGT events. Furthermore, the vast majority of MGHs identified are unlikely to be transferred, due to the transposition mechanisms of the identified IS elements flanking putative AR genes. In addition, MIC experiments using bacitracin A antibiotic demonstrated that strains presented a *bac*A genes located in an MGH resulted resistant toward this antibiotic in respect bifidobacterial strains without a *bac*A genes, confirming the *in silico* prediction data. These analyses underlined the safety of *Bifidobacterium* genus in respect to other taxa, as well as *Escherichia coli* or members of the Gammaproteobacteria class, which are demonstrated to contribute to a high AR load in the human microbiota (283).

8.4 The Amoxicillin-Clavulanic acid antibiotic resistance of *Bifidobacterium breve* 1891B.

In Westernized countries, the combination antibiotic Amoxicillin-Clavulanic acid (AMC) is one of the most prescribed drugs against bacterial infections, in particular during the infancy and adolescence. In chapter 5, the composition of the gut microbiota of 23 children that have taken this antibiotic was compared to the gut microbiota of a control group that has not assumed AMC in the last six months. Interestingly, a drastic reduction of bacteria was discovered in children that have assumed AMC, especially in the absolute abundance of health-promoting bacteria, such as bifidobacteria. A complete MIC screening of several bifidobacterial strains was performed, showing that 98.5 % of them resulted to be sensitive toward this antibiotic. Interestingly, among the 1.5 % of strains with increased resistance toward AMC, the strain *B. breve* 1891B displayed the highest MIC_{AMC} value (32 μg/mL). The transcriptomic analysis of this strain, growth in presence of AMC,

underlined interesting genes which presented a considerably up-regulation. These data suggested the possible involvement of these genes in the AMC resistance mechanism of *B. breve* 1891B. Moreover, simulating gut-microbiota experiments showed that 1891B is able to take over other bacteria in the presence of AMC, suggesting the possible use of this strain in bifidobacteria-containing probiotic products when AMC therapy is prescribed.

8.5 The role of *Lactobacillus crispatus* in the human vaginal microbiota.

The last part of this thesis regards the study of the vaginal microbiota and in particular of the typical vaginal species *Lactobacillus crispatus*. A metagenomics analysis revealed that *L. crispatus* is one of the most representative species in healthy women. Data interestingly underlined the negative correlation between this species and bacteria typically involved in vaginal infections. Moreover, 15 newly different strains were isolated from vaginal swabs and chicken fecal samples, confirming the high abundance of this species in these two ecological niches. A phylogenomic analysis, including both 15 newly isolated genomes and 94 public available strains, revealed a genetic differentiation between strains of human origin and strains of animal origin. In addition, batch culture experiments simulating the vaginal microbiota showed that one vaginal isolated strain, i.e. *L. crispatus* PRL2021, seemed to take over other typical vaginal lactobacilli, as well as *Lactobacillus jhensenii*, *Lactobacillus gasseri* and *Lactobacillus iners*. Finally, all these data combining with the high capability of PRL2021 to growth in a fecal environment suggested the possible use of this strain as probiotic bacteria for preventing vaginosis and vaginal dysbiosis.

REFERENCES

1. Group NHW, Peterson J, Garges S, Giovanni M, McInnes P, Wang L, Schloss JA, Bonazzi V, McEwen JE, Wetterstrand KA, Deal C, Baker CC, Di Francesco V, Howcroft TK, Karp RW, Lunsford RD, Wellington CR, Belachew T, Wright M, Giblin C, David H, Mills M, Salomon R, Mullins C, Akolkar B, Begg L, Davis C, Grandison L, Humble M, Khalsa J, Little AR, Peavy H, Pontzer C, Portnoy M, Sayre MH, Starke-Reed P, Zakhari S, Read J, Watson B, Guyer M. 2009. The NIH Human Microbiome Project. Genome Res 19:2317-23.
2. Marchesi JR, Ravel J. 2015. The vocabulary of microbiome research: a proposal. Microbiome 3:31.
3. Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. 2005. Host-bacterial mutualism in the human intestine. Science 307:1915-20.
4. Turnbaugh PJ, Ley RE, Hamady M, Fraser-Liggett CM, Knight R, Gordon JI. 2007. The human microbiome project. Nature 449:804-10.
5. Ley RE, Peterson DA, Gordon JI. 2006. Ecological and evolutionary forces shaping microbial diversity in the human intestine. Cell 124:837-48.
6. Smith SB, Ravel J. 2017. The vaginal microbiota, host defence and reproductive physiology. J Physiol 595:451-463.
7. Byrd AL, Belkaid Y, Segre JA. 2018. The human skin microbiome. Nat Rev Microbiol 16:143-155.
8. Yamashita Y, Takeshita T. 2017. The oral microbiome and human health. J Oral Sci 59:201-206.
9. Rawls M, Ellis AK. 2019. The microbiome of the nose. Ann Allergy Asthma Immunol 122:17-24.
10. Wu BG, Segal LN. 2017. Lung Microbiota and Its Impact on the Mucosal Immune Phenotype. Microbiol Spectr 5.
11. Palmer RJ, Jr. 2014. Composition and development of oral bacterial communities. Periodontol 2000 64:20-39.
12. Abranches J, Zeng L, Kajfasz JK, Palmer SR, Chakraborty B, Wen ZT, Richards VP, Brady LJ, Lemos JA. 2018. Biology of Oral Streptococci. Microbiol Spectr 6.
13. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, Beck JM, Curtis JL, Huffnagle GB. 2015. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. mBio 6:e00037.
14. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, Flores SC, Fontenot AP, Ghedin E, Huang L, Jablonski K, Kleerup E, Lynch SV, Sodergren E, Twigg H, Young VB, Bassis CM, Venkataraman A, Schmidt TM, Weinstock GM, Lung HIVMP. 2013. Comparison of the respiratory microbiome in healthy nonsmokers and smokers. Am J Respir Crit Care Med 187:1067-75.
15. Wang J, Li F, Tian Z. 2017. Role of microbiota on lung homeostasis and diseases. Sci China Life Sci 60:1407-1415.
16. Rosenbaum M, Knight R, Leibel RL. 2015. The gut microbiota in human energy homeostasis and obesity. Trends Endocrinol Metab 26:493-501.
17. Linhares IM, Summers PR, Larsen B, Giraldo PC, Witkin SS. 2011. Contemporary perspectives on vaginal pH and lactobacilli. Am J Obstet Gynecol 204:120 e1-5.
18. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO, Brotman RM, Davis CC, Ault K, Peralta L, Forney LJ. 2011. Vaginal microbiome of reproductive-age women. Proc Natl Acad Sci U S A 108 Suppl 1:4680-7.
19. Wang S, Wang Q, Yang E, Yan L, Li T, Zhuang H. 2017. 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 8:564.
20. Nahui Palomino RA, Vanpouille C, Laghi L, Parolin C, Melikov K, Backlund P, Vitali B, Margolis L. 2019. Extracellular vesicles from symbiotic vaginal lactobacilli inhibit HIV-1 infection of human tissues. Nat Commun 10:5656.
21. Kwasniewski W, Wolun-Cholewa M, Kotarski J, Warchol W, Kuzma D, Kwasniewska A, Gozdzicka-Jozefiak A. 2018. Microbiota dysbiosis is associated with HPV-induced cervical carcinogenesis. Oncol Lett 16:7035-7047.
22. Uchihashi M, Bergin IL, Bassis CM, Hashway SA, Chai D, Bell JD. 2015. Influence of age, reproductive cycling status, and menstruation on the vaginal microbiome in baboons (Papio anubis). Am J Primatol 77:563-78.
23. Miller EA, Livermore JA, Alberts SC, Tung J, Archie EA. 2017. Ovarian cycling and reproductive state shape the vaginal microbiota in wild baboons. Microbiome 5:8.
24. Hashway SA, Bergin IL, Bassis CM, Uchihashi M, Schmidt KC, Young VB, Aronoff DM, Patton DL, Bell JD. 2014. Impact of a hormone-releasing intrauterine system on the vaginal microbiome: a prospective baboon model. J Med Primatol 43:89-99.
25. Hickey RJ, Abdo Z, Zhou X, Nemeth K, Hansmann M, Osborn TW, 3rd, Wang F, Forney LJ. 2013. Effects of tampons and menses on the composition and diversity of vaginal microbial communities over time. BJOG 120:695-704; discussion 704-6.
26. Hickey RJ, Zhou X, Pierson JD, Ravel J, Forney LJ. 2012. Understanding vaginal microbiome complexity from an ecological perspective. Transl Res 160:267-82.
27. Nunn KL, Forney LJ. 2016. Unraveling the Dynamics of the Human Vaginal Microbiome. Yale J Biol Med 89:331-337.
28. Martin DH, Marrazzo JM. 2016. The Vaginal Microbiome: Current Understanding and Future Directions. J Infect Dis 214 Suppl 1:S36-41.
29. Laniewski P, Barnes D, Goulder A, Cui H, Roe DJ, Chase DM, Herbst-Kralovetz MM. 2018. Linking cervicovaginal immune signatures, HPV and microbiota composition in cervical carcinogenesis in non-Hispanic and Hispanic women. Sci Rep 8:7593.
30. Onywera H, Williamson AL, Mbulawa ZZA, Coetzee D, Meiring TL. 2019. The cervical microbiota in reproductive-age South African women with and without human papillomavirus infection. Papillomavirus Res 7:154-163.
31. Borgdorff H, van der Veer C, van Houdt R, Alberts CJ, de Vries HJ, Bruisten SM, Snijder MB, Prins M, Geerlings SE, Schim van der Loeff MF, van de Wijgert J. 2017. The association between ethnicity and vaginal microbiota composition in Amsterdam, the Netherlands. PLoS One 12:e0181135.
32. Dols JA, Molenaar D, van der Helm JJ, Caspers MP, de Kat Angelino-Bart A, Schuren FH, Speksnijder AG, Westerhoff HV, Richardus JH, Boon ME, Reid G, de Vries HJ, Kort R. 2016. Molecular assessment of bacterial vaginosis by Lactobacillus abundance and species diversity. BMC Infect Dis 16:180.
33. Doerflinger SY, Throop AL, Herbst-Kralovetz MM. 2014. Bacteria in the vaginal microbiome alter the innate immune response and barrier properties of the human vaginal epithelia in a species-specific manner. J Infect Dis 209:1989-99.
34. Rose WA, 2nd, McGowin CL, Spagnuolo RA, Eaves-Pyles TD, Popov VL, Pyles RB. 2012. Commensal bacteria modulate innate immune responses of vaginal epithelial cell multilayer cultures. PLoS One 7:e32728.
35. Stapleton AE, Au-Yeung M, Hooton TM, Fredricks DN, Roberts PL, Czaja CA, Yarova-Yarovaya Y, Fiedler T, Cox M, Stamm WE. 2011. Randomized, placebo-controlled phase 2 trial of a Lactobacillus crispatus probiotic given intravaginally for prevention of recurrent urinary tract infection. Clin Infect Dis 52:1212-7.
36. Kort R. 2014. Personalized therapy with probiotics from the host by TripleA. Trends Biotechnol 32:291-3.

37. Li T, Liu Z, Zhang X, Chen X, Wang S. 2019. Local Probiotic Lactobacillus crispatus and Lactobacillus delbrueckii Exhibit Strong Antifungal Effects Against Vulvovaginal Candidiasis in a Rat Model. Front Microbiol 10:1033.
38. Lin L, Zhang J. 2017. Role of intestinal microbiota and metabolites on gut homeostasis and human diseases. BMC Immunol 18:2.
39. Marchesi JR, Adams DH, Fava F, Hermes GD, Hirschfield GM, Hold G, Quraishi MN, Kinross J, Smidt H, Tuohy KM, Thomas LV, Zoetendal EG, Hart A. 2016. The gut microbiota and host health: a new clinical frontier. Gut 65:330-9.
40. Clarke G, Stilling RM, Kennedy PJ, Stanton C, Cryan JF, Dinan TG. 2014. Minireview: Gut microbiota: the neglected endocrine organ. Mol Endocrinol 28:1221-38.
41. Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Nageshwar Reddy D. 2015. Role of the normal gut microbiota. World J Gastroenterol 21:8787-803.
42. Ventura M, Turroni F, Motherway MO, MacSharry J, van Sinderen D. 2012. Host-microbe interactions that facilitate gut colonization by commensal bifidobacteria. Trends Microbiol 20:467-76.
43. Nuriel-Ohayon M, Neuman H, Koren O. 2016. Microbial Changes during Pregnancy, Birth, and Infancy. Front Microbiol 7:1031.
44. Joossens M, Huys G, Cnockaert M, De Preter V, Verbeke K, Rutgeerts P, Vandamme P, Vermeire S. 2011. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut 60:631-7.
45. Sobhani I, Amiot A, Le Baleur Y, Levy M, Auriault ML, Van Nhieu JT, Delchier JC. 2013. Microbial dysbiosis and colon carcinogenesis: could colon cancer be considered a bacteria-related disease? Therap Adv Gastroenterol 6:215-29.
46. Gibson GR, Roberfroid MB. 1995. Dietary modulation of the human colonic microbiota: introducing the concept of prebiotics. J Nutr 125:1401-12.
47. Dethlefsen L, Eckburg PB, Bik EM, Relman DA. 2006. Assembly of the human intestinal microbiota. Trends Ecol Evol 21:517-23.
48. Rajilic-Stojanovic M, Smidt H, de Vos WM. 2007. Diversity of the human gastrointestinal tract microbiota revisited. Environ Microbiol 9:2125-36.
49. Turnbaugh PJ, Ridaura VK, Faith JJ, Rey FE, Knight R, Gordon JI. 2009. The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. Sci Transl Med 1:6ra14.
50. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, Deming C, Quinones M, Koo L, Conlan S, Spencer S, Hall JA, Dzutsev A, Kong H, Campbell DJ, Trinchieri G, Segre JA, Belkaid Y. 2012. Compartmentalized control of skin immunity by resident commensals. Science 337:1115-9.
51. Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature 444:1027-31.
52. Favier CF, Vaughan EE, De Vos WM, Akkermans AD. 2002. Molecular monitoring of succession of bacterial communities in human neonates. Appl Environ Microbiol 68:219-26.
53. Claesson MJ, Cusack S, O'Sullivan O, Greene-Diniz R, de Weerd H, Flannery E, Marchesi JR, Falush D, Dinan T, Fitzgerald G, Stanton C, van Sinderen D, O'Connor M, Harnedy N, O'Connor K, Henry C, O'Mahony D, Fitzgerald AP, Shanahan F, Twomey C, Hill C, Ross RP, O'Toole PW. 2011. Composition, variability, and temporal stability of the intestinal microbiota of the elderly. Proc Natl Acad Sci U S A 108 Suppl 1:4586-91.
54. Salazar N, Lopez P, Valdes L, Margolles A, Suarez A, Patterson AM, Cuervo A, de los Reyes-Gavilan CG, Ruas-Madiedo P, Gonzalez S, Gueimonde M. 2013. Microbial targets for the development of functional foods accordingly with nutritional and immune parameters altered in the elderly. J Am Coll Nutr 32:399-406.
55. Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T,

Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Dore J, Meta HITC, Antolin M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C, Cultrone A, Delorme C, Denariaz G, et al. 2011. Enterotypes of the human gut microbiome. Nature 473:174-80.

56. Turroni F, Peano C, Pass DA, Foroni E, Severgnini M, Claesson MJ, Kerr C, Hourihane J, Murray D, Fuligni F, Gueimonde M, Margolles A, De Bellis G, O'Toole PW, van Sinderen D, Marchesi JR, Ventura M. 2012. Diversity of bifidobacteria within the infant gut microbiota. PLoS One 7:e36957.
57. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Meier-Kolthoff JP, Klenk HP, Clement C, Ouhdouch Y, van Wezel GP. 2016. Taxonomy, Physiology, and Natural Products of Actinobacteria. Microbiol Mol Biol Rev 80:1-43.
58. Ventura M, Canchaya C, Fitzgerald GF, Gupta RS, van Sinderen D. 2007. Genomics as a means to understand bacterial phylogeny and ecological adaptation: the case of bifidobacteria. Antonie Van Leeuwenhoek 91:351-72.
59. Servin JA, Herbold CW, Skophammer RG, Lake JA. 2008. Evidence excluding the root of the tree of life from the actinobacteria. Mol Biol Evol 25:1-4.
60. Berdy J. 2005. Bioactive microbial metabolites. J Antibiot (Tokyo) 58:1-26.
61. Noens EE, Mersinias V, Willemse J, Traag BA, Laing E, Chater KF, Smith CP, Koerten HK, van Wezel GP. 2007. Loss of the controlled localization of growth stage-specific cell-wall synthesis pleiotropically affects developmental gene expression in an ssgA mutant of Streptomyces coelicolor. Mol Microbiol 64:1244-59.
62. Bennett JW. 1998. Mycotechnology: the role of fungi in biotechnology. J Biotechnol 66:101-7.
63. Hopwood DA, Chater KF, Bibb MJ. 1995. Genetics of antibiotic production in Streptomyces coelicolor A3(2), a model streptomycete. Biotechnology 28:65-102.
64. Philippe H, Douady CJ. 2003. Horizontal gene transfer and phylogenetics. Curr Opin Microbiol 6:498-505.
65. Ventura M, Canchaya C, Del Casale A, Dellaglio F, Neviani E, Fitzgerald GF, van Sinderen D. 2006. Analysis of bifidobacterial evolution using a multilocus approach. Int J Syst Evol Microbiol 56:2783-92.
66. Lugli GA, Milani C, Duranti S, Alessandri G, Turroni F, Mancabelli L, Tatoni D, Ossiprandi MC, van Sinderen D, Ventura M. 2019. Isolation of novel gut bifidobacteria using a combination of metagenomic and cultivation approaches. Genome Biol 20:96.
67. Lugli GA, Milani C, Duranti S, Mancabelli L, Mangifesta M, Turroni F, Viappiani A, van Sinderen D, Ventura M. 2018. Tracking the Taxonomy of the Genus Bifidobacterium Based on a Phylogenomic Approach. Appl Environ Microbiol 84.
68. Turroni F, van Sinderen D, Ventura M. 2011. Genomics and ecological overview of the genus Bifidobacterium. Int J Food Microbiol 149:37-44.
69. Modesto M, Watanabe K, Arita M, Satti M, Oki K, Sciavilla P, Patavino C, Camma C, Michelini S, Sgorbati B, Mattarelli P. 2019. Bifidobacterium jacchi sp. nov., isolated from the faeces of a baby common marmoset (Callithrix jacchus). Int J Syst Evol Microbiol 69:2477-2485.
70. Duranti S, Lugli GA, Napoli S, Anzalone R, Milani C, Mancabelli L, Alessandri G, Turroni F, Ossiprandi MC, van Sinderen D, Ventura M. 2019. Characterization of the phylogenetic diversity of five novel species belonging to the genus Bifidobacterium: Bifidobacterium castoris sp. nov., Bifidobacterium callimiconis sp. nov., Bifidobacterium goeldii sp. nov., Bifidobacterium samirii sp. nov. and Bifidobacterium dolichotidis sp. nov. Int J Syst Evol Microbiol 69:1288-1298.

71. Alberoni D, Gaggia F, Baffoni L, Modesto MM, Biavati B, Di Gioia D. 2019. Bifidobacterium xylocopae sp. nov. and Bifidobacterium aemilianum sp. nov., from the carpenter bee (Xylocopa violacea) digestive tract. Syst Appl Microbiol 42:205-216.
72. Trovatelli LD, Crociani F, Pedinotti M, Scardovi V. 1974. Bifidobacterium pullorum sp. nov.: a new species isolated from chicken feces and a related group of bifidobacteria isolated from rabbit feces. Arch Microbiol 98:187-98.
73. Biavati B, Mattarelli P. 1991. Bifidobacterium ruminantium sp. nov. and Bifidobacterium merycicum sp. nov. from the rumens of cattle. Int J Syst Bacteriol 41:163-8.
74. Okamoto M, Benno Y, Leung KP, Maeda N. 2008. Bifidobacterium tsurumiense sp. nov., from hamster dental plaque. Int J Syst Evol Microbiol 58:144-8.
75. Milani C, Mangifesta M, Mancabelli L, Lugli GA, James K, Duranti S, Turroni F, Ferrario C, Ossiprandi MC, van Sinderen D, Ventura M. 2017. Unveiling bifidobacterial biogeography across the mammalian branch of the tree of life. ISME J 11:2834-2847.
76. Turroni F, Foroni E, Pizzetti P, Giubellini V, Ribbera A, Merusi P, Cagnasso P, Bizzarri B, de'Angelis GL, Shanahan F, van Sinderen D, Ventura M. 2009. Exploring the diversity of the bifidobacterial population in the human intestinal tract. Appl Environ Microbiol 75:1534-45.
77. Schell MA, Karmirantzou M, Snel B, Vilanova D, Berger B, Pessi G, Zwahlen MC, Desiere F, Bork P, Delley M, Pridmore RD, Arigoni F. 2002. The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract. Proc Natl Acad Sci U S A 99:14422-7.
78. Sela DA, Chapman J, Adeuya A, Kim JH, Chen F, Whitehead TR, Lapidus A, Rokhsar DS, Lebrilla CB, German JB, Price NP, Richardson PM, Mills DA. 2008. The genome sequence of Bifidobacterium longum subsp. infantis reveals adaptations for milk utilization within the infant microbiome. Proc Natl Acad Sci U S A 105:18964-9.
79. Ventura M, Turroni F, Zomer A, Foroni E, Giubellini V, Bottacini F, Canchaya C, Claesson MJ, He F, Mantzourani M, Mulas L, Ferrarini A, Gao B, Delledonne M, Henrissat B, Coutinho P, Oggioni M, Gupta RS, Zhang Z, Beighton D, Fitzgerald GF, O'Toole PW, van Sinderen D. 2009. The Bifidobacterium dentium Bd1 genome sequence reflects its genetic adaptation to the human oral cavity. PLoS Genet 5:e1000785.
80. Turroni F, Bottacini F, Foroni E, Mulder I, Kim JH, Zomer A, Sanchez B, Bidossi A, Ferrarini A, Giubellini V, Delledonne M, Henrissat B, Coutinho P, Oggioni M, Fitzgerald GF, Mills D, Margolles A, Kelly D, van Sinderen D, Ventura M. 2010. Genome analysis of Bifidobacterium bifidum PRL2010 reveals metabolic pathways for host-derived glycan foraging. Proc Natl Acad Sci U S A 107:19514-9.
81. Turroni F, Milani C, van Sinderen D, Ventura M. 2011. Genetic strategies for mucin metabolism in Bifidobacterium bifidum PRL2010: an example of possible human-microbe co-evolution. Gut Microbes 2:183-9.
82. Turroni F, Duranti S, Bottacini F, Guglielmetti S, Van Sinderen D, Ventura M. 2014. Bifidobacterium bifidum as an example of a specialized human gut commensal. Front Microbiol 5:437.
83. Guglielmetti S, Tamagnini I, Minuzzo M, Arioli S, Parini C, Comelli E, Mora D. 2009. Study of the adhesion of Bifidobacterium bifidum MIMBb75 to human intestinal cell lines. Curr Microbiol 59:167-72.
84. Turroni F, Duranti S, Milani C, Lugli GA, van Sinderen D, Ventura M. 2019. Bifidobacterium bifidum: A Key Member of the Early Human Gut Microbiota. Microorganisms 7.
85. Podolsky DK. 1985. Oligosaccharide structures of isolated human colonic mucin species. J Biol Chem 260:15510-5.
86. Pokusaeva K, Fitzgerald GF, van Sinderen D. 2011. Carbohydrate metabolism in Bifidobacteria. Genes Nutr 6:285-306.
87. Duranti S, Milani C, Lugli GA, Turroni F, Mancabelli L, Sanchez B, Ferrario C, Viappiani A, Mangifesta M, Mancino W, Gueimonde M, Margolles A, van Sinderen D, Ventura M. 2015.

Insights from genomes of representatives of the human gut commensal Bifidobacterium bifidum. Environ Microbiol 17:2515-31.
88. Arboleya S, Bottacini F, O'Connell-Motherway M, Ryan CA, Ross RP, van Sinderen D, Stanton C. 2018. Gene-trait matching across the Bifidobacterium longum pan-genome reveals considerable diversity in carbohydrate catabolism among human infant strains. BMC Genomics 19:33.
89. O'Callaghan A, Bottacini F, O'Connell Motherway M, van Sinderen D. 2015. Pangenome analysis of Bifidobacterium longum and site-directed mutagenesis through by-pass of restriction-modification systems. BMC Genomics 16:832.
90. Liu S, Ren F, Zhao L, Jiang L, Hao Y, Jin J, Zhang M, Guo H, Lei X, Sun E, Liu H. 2015. Starch and starch hydrolysates are favorable carbon sources for bifidobacteria in the human gut. BMC Microbiol 15:54.
91. Lee JH, Karamychev VN, Kozyavkin SA, Mills D, Pavlov AR, Pavlova NV, Polouchine NN, Richardson PM, Shakhova VV, Slesarev AI, Weimer B, O'Sullivan DJ. 2008. Comparative genomic analysis of the gut bacterium Bifidobacterium longum reveals loci susceptible to deletion during pure culture growth. BMC Genomics 9:247.
92. Milani C, Lugli GA, Duranti S, Turroni F, Bottacini F, Mangifesta M, Sanchez B, Viappiani A, Mancabelli L, Taminiau B, Delcenserie V, Barrangou R, Margolles A, van Sinderen D, Ventura M. 2014. Genomic encyclopedia of type strains of the genus Bifidobacterium. Appl Environ Microbiol 80:6290-302.
93. Klappenbach JA, Dunbar JM, Schmidt TM. 2000. rRNA operon copy number reflects ecological strategies of bacteria. Appl Environ Microbiol 66:1328-33.
94. Ventura M, Canchaya C, Tauch A, Chandra G, Fitzgerald GF, Chater KF, van Sinderen D. 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol Mol Biol Rev 71:495-548.
95. Yildirim Z, Winters DK, Johnson MG. 1999. Purification, amino acid sequence and mode of action of bifidocin B produced by Bifidobacterium bifidum NCFB 1454. J Appl Microbiol 86:45-54.
96. Ventura M, Lee JH, Canchaya C, Zink R, Leahy S, Moreno-Munoz JA, O'Connell-Motherway M, Higgins D, Fitzgerald GF, O'Sullivan DJ, van Sinderen D. 2005. Prophage-like elements in bifidobacteria: insights from genomics, transcription, integration, distribution, and phylogenetic analysis. Appl Environ Microbiol 71:8692-705.
97. Ventura M, Turroni F, Lima-Mendez G, Foroni E, Zomer A, Duranti S, Giubellini V, Bottacini F, Horvath P, Barrangou R, Sela DA, Mills DA, van Sinderen D. 2009. Comparative analyses of prophage-like elements present in bifidobacterial genomes. Appl Environ Microbiol 75:6929-36.
98. Ventura M, Turroni F, Foroni E, Duranti S, Giubellini V, Bottacini F, van Sinderen D. 2010. Analyses of bifidobacterial prophage-like sequences. Antonie Van Leeuwenhoek 98:39-50.
99. Lugli GA, Milani C, Turroni F, Tremblay D, Ferrario C, Mancabelli L, Duranti S, Ward DV, Ossiprandi MC, Moineau S, van Sinderen D, Ventura M. 2016. Prophages of the genus Bifidobacterium as modulating agents of the infant gut microbiota. Environ Microbiol 18:2196-213.
100. Canchaya C, Proux C, Fournous G, Bruttin A, Brussow H. 2003. Prophage genomics. Microbiol Mol Biol Rev 67:238-76, table of contents.
101. Ventura M, Zomer A, Canchaya C, O'Connell-Motherway M, Kuipers O, Turroni F, Ribbera A, Foroni E, Buist G, Wegmann U, Shearman C, Gasson MJ, Fitzgerald GF, Kok J, van Sinderen D. 2007. Comparative analyses of prophage-like elements present in two Lactococcus lactis strains. Appl Environ Microbiol 73:7771-80.
102. Ammor MS, Florez AB, Mayo B. 2007. Antibiotic resistance in non-enterococcal lactic acid bacteria and bifidobacteria. Food Microbiol 24:559-70.

103. Kazimierczak KA, Flint HJ, Scott KP. 2006. Comparative analysis of sequences flanking tet(W) resistance genes in multiple species of gut bacteria. Antimicrob Agents Chemother 50:2632-9.
104. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA, Horvath P. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. Science 315:1709-12.
105. Barrangou R. 2013. CRISPR-Cas systems and RNA-guided interference. Wiley Interdiscip Rev RNA 4:267-78.
106. Milani C, Lugli GA, Duranti S, Turroni F, Mancabelli L, Ferrario C, Mangifesta M, Hevia A, Viappiani A, Scholz M, Arioli S, Sanchez B, Lane J, Ward DV, Hickey R, Mora D, Segata N, Margolles A, van Sinderen D, Ventura M. 2015. Bifidobacteria exhibit social behavior through carbohydrate resource sharing in the gut. Sci Rep 5:15782.
107. Milani C, Turroni F, Duranti S, Lugli GA, Mancabelli L, Ferrario C, van Sinderen D, Ventura M. 2016. Genomics of the Genus Bifidobacterium Reveals Species-Specific Adaptation to the Glycan-Rich Gut Environment. Appl Environ Microbiol 82:980-991.
108. Singh RP. 2019. Glycan utilisation system in Bacteroides and Bifidobacteria and their roles in gut stability and health. Appl Microbiol Biotechnol 103:7287-7315.
109. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490-5.
110. Ruas-Madiedo P, Gueimonde M, Fernandez-Garcia M, de los Reyes-Gavilan CG, Margolles A. 2008. Mucin degradation by Bifidobacterium strains isolated from the human intestinal microbiota. Appl Environ Microbiol 74:1936-40.
111. Ruiz L, Gueimonde M, Coute Y, Salminen S, Sanchez JC, de los Reyes-Gavilan CG, Margolles A. 2011. Evaluation of the ability of Bifidobacterium longum to metabolize human intestinal mucus. FEMS Microbiol Lett 314:125-30.
112. Lugli GA, Duranti S, Milani C, Mancabelli L, Turroni F, Alessandri G, Longhi G, Anzalone R, Viappinai A, Tarracchini C, Bernasconi S, Yonemitsu C, Bode L, Goran MI, Ossiprandi MC, van Sinderen D, Ventura M. 2020. Investigating bifidobacteria and human milk oligosaccharide composition of lactating mothers. FEMS Microbiol Ecol 96.
113. Yoshida E, Sakurama H, Kiyohara M, Nakajima M, Kitaoka M, Ashida H, Hirose J, Katayama T, Yamamoto K, Kumagai H. 2012. Bifidobacterium longum subsp. infantis uses two different beta-galactosidases for selectively degrading type-1 and type-2 human milk oligosaccharides. Glycobiology 22:361-8.
114. Asakuma S, Hatakeyama E, Urashima T, Yoshida E, Katayama T, Yamamoto K, Kumagai H, Ashida H, Hirose J, Kitaoka M. 2011. Physiology of consumption of human milk oligosaccharides by infant gut-associated bifidobacteria. J Biol Chem 286:34583-92.
115. Sela DA. 2011. Bifidobacterial utilization of human milk oligosaccharides. Int J Food Microbiol 149:58-64.
116. Duranti S, Turroni F, Milani C, Foroni E, Bottacini F, Dal Bello F, Ferrarini A, Delledonne M, van Sinderen D, Ventura M. 2013. Exploration of the genomic diversity and core genome of the Bifidobacterium adolescentis phylogenetic group by means of a polyphasic approach. Appl Environ Microbiol 79:336-46.
117. Duranti S, Turroni F, Lugli GA, Milani C, Viappiani A, Mangifesta M, Gioiosa L, Palanza P, van Sinderen D, Ventura M. 2014. Genomic characterization and transcriptional studies of the starch-utilizing strain Bifidobacterium adolescentis 22L. Appl Environ Microbiol 80:6080-90.
118. Ryan SM, Fitzgerald GF, van Sinderen D. 2006. Screening for and identification of starch-, amylopectin-, and pullulan-degrading activities in bifidobacterial strains. Appl Environ Microbiol 72:5289-96.
119. Ze X, Duncan SH, Louis P, Flint HJ. 2012. Ruminococcus bromii is a keystone species for the degradation of resistant starch in the human colon. ISME J 6:1535-43.

120. Ventola CL. 2015. The antibiotic resistance crisis: part 1: causes and threats. P T 40:277-83.
121. Munita JM, Arias CA. 2016. Mechanisms of Antibiotic Resistance. Microbiol Spectr 4.
122. Ouwehand AC, Forssten S, Hibberd AA, Lyra A, Stahl B. 2016. Probiotic approach to prevent antibiotic resistance. Ann Med 48:246-55.
123. Serafini F, Bottacini F, Viappiani A, Baruffini E, Turroni F, Foroni E, Lodi T, van Sinderen D, Ventura M. 2011. Insights into physiological and genetic mupirocin susceptibility in bifidobacteria. Appl Environ Microbiol 77:3141-6.
124. Morovic W, Roos P, Zabel B, Hidalgo-Cantabrana C, Kiefer A, Barrangou R. 2018. Transcriptional and Functional Analysis of Bifidobacterium animalis subsp. lactis Exposure to Tetracycline. Appl Environ Microbiol 84.
125. Sutherland R, Boon RJ, Griffin KE, Masters PJ, Slocombe B, White AR. 1985. Antibacterial activity of mupirocin (pseudomonic acid), a new antibiotic for topical use. Antimicrob Agents Chemother 27:495-8.
126. Hughes J, Mellows G. 1980. Interaction of pseudomonic acid A with Escherichia coli B isoleucyl-tRNA synthetase. Biochem J 191:209-19.
127. Milani C, Duranti S, Lugli GA, Bottacini F, Strati F, Arioli S, Foroni E, Turroni F, van Sinderen D, Ventura M. 2013. Comparative genomics of Bifidobacterium animalis subsp. lactis reveals a strict monophyletic bifidobacterial taxon. Appl Environ Microbiol 79:4304-15.
128. Turroni F, Serafini F, Foroni E, Duranti S, O'Connell Motherway M, Taverniti V, Mangifesta M, Milani C, Viappiani A, Roversi T, Sanchez B, Santoni A, Gioiosa L, Ferrarini A, Delledonne M, Margolles A, Piazza L, Palanza P, Bolchi A, Guglielmetti S, van Sinderen D, Ventura M. 2013. Role of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in modulating bacterium-host interactions. Proc Natl Acad Sci U S A 110:11151-6.
129. Turroni F, Foroni E, O'Connell Motherway M, Bottacini F, Giubellini V, Zomer A, Ferrarini A, Delledonne M, Zhang Z, van Sinderen D, Ventura M. 2010. Characterization of the serpin-encoding gene of Bifidobacterium breve 210B. Appl Environ Microbiol 76:3206-19.
130. Ferrario C, Milani C, Mancabelli L, Lugli GA, Duranti S, Mangifesta M, Viappiani A, Turroni F, Margolles A, Ruas-Madiedo P, van Sinderen D, Ventura M. 2016. Modulation of the eps-ome transcription of bifidobacteria through simulation of human intestinal environment. FEMS Microbiol Ecol 92:fiw056.
131. Milani C, Mangifesta M, Mancabelli L, Lugli GA, Mancino W, Viappiani A, Faccini A, van Sinderen D, Ventura M, Turroni F. 2017. The Sortase-Dependent Fimbriome of the Genus Bifidobacterium: Extracellular Structures with Potential To Modulate Microbe-Host Dialogue. Appl Environ Microbiol 83.
132. Freitas F, Alves VD, Reis MA. 2011. Advances in bacterial exopolysaccharides: from production to biotechnological applications. Trends Biotechnol 29:388-98.
133. Hall-Stoodley L, Costerton JW, Stoodley P. 2004. Bacterial biofilms: from the natural environment to infectious diseases. Nat Rev Microbiol 2:95-108.
134. Fux CA, Costerton JW, Stewart PS, Stoodley P. 2005. Survival strategies of infectious biofilms. Trends Microbiol 13:34-40.
135. Fanning S, Hall LJ, van Sinderen D. 2012. Bifidobacterium breve UCC2003 surface exopolysaccharide production is a beneficial trait mediating commensal-host interaction through immune modulation and pathogen protection. Gut Microbes 3:420-5.
136. Horn N, Wegmann U, Dertli E, Mulholland F, Collins SR, Waldron KW, Bongaerts RJ, Mayer MJ, Narbad A. 2013. Spontaneous mutation reveals influence of exopolysaccharide on Lactobacillus johnsonii surface characteristics. PLoS One 8:e59957.
137. Turroni F, Ventura M, Butto LF, Duranti S, O'Toole PW, Motherway MO, van Sinderen D. 2014. Molecular dialogue between the human gut microbiota and the host: a Lactobacillus and Bifidobacterium perspective. Cell Mol Life Sci 71:183-203.

138. Hidalgo-Cantabrana C, Nikolic M, Lopez P, Suarez A, Miljkovic M, Kojic M, Margolles A, Golic N, Ruas-Madiedo P. 2014. Exopolysaccharide-producing Bifidobacterium animalis subsp. lactis strains and their polymers elicit different responses on immune cells from blood and gut associated lymphoid tissue. Anaerobe 26:24-30.
139. Fanning S, Hall LJ, Cronin M, Zomer A, MacSharry J, Goulding D, Motherway MO, Shanahan F, Nally K, Dougan G, van Sinderen D. 2012. Bifidobacterial surface-exopolysaccharide facilitates commensal-host interaction through immune modulation and pathogen protection. Proc Natl Acad Sci U S A 109:2108-13.
140. Xu R, Shang N, Li P. 2011. In vitro and in vivo antioxidant activity of exopolysaccharide fractions from Bifidobacterium animalis RH. Anaerobe 17:226-31.
141. O'Connell Motherway M, Zomer A, Leahy SC, Reunanen J, Bottacini F, Claesson MJ, O'Brien F, Flynn K, Casey PG, Munoz JA, Kearney B, Houston AM, O'Mahony C, Higgins DG, Shanahan F, Palva A, de Vos WM, Fitzgerald GF, Ventura M, O'Toole PW, van Sinderen D. 2011. Functional genome analysis of Bifidobacterium breve UCC2003 reveals type IVb tight adherence (Tad) pili as an essential and conserved host-colonization factor. Proc Natl Acad Sci U S A 108:11217-22.
142. Ivanov D, Emonet C, Foata F, Affolter M, Delley M, Fisseha M, Blum-Sperisen S, Kochhar S, Arigoni F. 2006. A serpin from the gut bacterium Bifidobacterium longum inhibits eukaryotic elastase-like serine proteases. J Biol Chem 281:17246-52.
143. Roberts TH, Hejgaard J, Saunders NF, Cavicchioli R, Curmi PM. 2004. Serpins in unicellular Eukarya, Archaea, and Bacteria: sequence analysis and evolution. J Mol Evol 59:437-47.
144. Putignani L, Del Chierico F, Petrucca A, Vernocchi P, Dallapiccola B. 2014. The human gut microbiota: a dynamic interplay with the host from birth to senescence settled during childhood. Pediatr Res 76:2-10.
145. Aagaard K, Ma J, Antony KM, Ganu R, Petrosino J, Versalovic J. 2014. The placenta harbors a unique microbiome. Sci Transl Med 6:237ra65.
146. Perez-Munoz ME, Arrieta MC, Ramer-Tait AE, Walter J. 2017. A critical assessment of the "sterile womb" and "in utero colonization" hypotheses: implications for research on the pioneer infant microbiome. Microbiome 5:48.
147. Rodriguez JM, Murphy K, Stanton C, Ross RP, Kober OI, Juge N, Avershina E, Rudi K, Narbad A, Jenmalm MC, Marchesi JR, Collado MC. 2015. The composition of the gut microbiota throughout life, with an emphasis on early life. Microb Ecol Health Dis 26:26050.
148. Duranti S, Lugli GA, Mancabelli L, Armanini F, Turroni F, James K, Ferretti P, Gorfer V, Ferrario C, Milani C, Mangifesta M, Anzalone R, Zolfo M, Viappiani A, Pasolli E, Bariletti I, Canto R, Clementi R, Cologna M, Crifo T, Cusumano G, Fedi S, Gottardi S, Innamorati C, Mase C, Postai D, Savoi D, Soffiati M, Tateo S, Pedrotti A, Segata N, van Sinderen D, Ventura M. 2017. Maternal inheritance of bifidobacterial communities and bifidophages in infants through vertical transmission. Microbiome 5:66.
149. Penders J, Thijs C, Vink C, Stelma FF, Snijders B, Kummeling I, van den Brandt PA, Stobberingh EE. 2006. Factors influencing the composition of the intestinal microbiota in early infancy. Pediatrics 118:511-21.
150. Yatsunenko T, Rey FE, Manary MJ, Trehan I, Dominguez-Bello MG, Contreras M, Magris M, Hidalgo G, Baldassano RN, Anokhin AP, Heath AC, Warner B, Reeder J, Kuczynski J, Caporaso JG, Lozupone CA, Lauber C, Clemente JC, Knights D, Knight R, Gordon JI. 2012. Human gut microbiome viewed across age and geography. Nature 486:222-7.
151. Milani C, Mancabelli L, Lugli GA, Duranti S, Turroni F, Ferrario C, Mangifesta M, Viappiani A, Ferretti P, Gorfer V, Tett A, Segata N, van Sinderen D, Ventura M. 2015. Exploring Vertical Transmission of Bifidobacteria from Mother to Child. Appl Environ Microbiol 81:7078-87.

152. Avershina E, Storro O, Oien T, Johnsen R, Pope P, Rudi K. 2014. Major faecal microbiota shifts in composition and diversity with age in a geographically restricted cohort of mothers and their children. FEMS Microbiol Ecol 87:280-90.
153. Rautava S, Luoto R, Salminen S, Isolauri E. 2012. Microbial contact during pregnancy, intestinal colonization and human disease. Nat Rev Gastroenterol Hepatol 9:565-76.
154. Blaser MJ, Dominguez-Bello MG. 2016. The Human Microbiome before Birth. Cell Host Microbe 20:558-560.
155. Morrow AL, Rangel JM. 2004. Human milk protection against infectious diarrhea: implications for prevention and clinical care. Semin Pediatr Infect Dis 15:221-8.
156. Urbaniak C, Angelini M, Gloor GB, Reid G. 2016. Human milk microbiota profiles in relation to birthing method, gestation and infant gender. Microbiome 4:1.
157. Fernandez L, Langa S, Martin V, Jimenez E, Martin R, Rodriguez JM. 2013. The microbiota of human milk in healthy women. Cell Mol Biol (Noisy-le-grand) 59:31-42.
158. Hunt KM, Foster JA, Forney LJ, Schutte UM, Beck DL, Abdo Z, Fox LK, Williams JE, McGuire MK, McGuire MA. 2011. Characterization of the diversity and temporal stability of bacterial communities in human milk. PLoS One 6:e21313.
159. Jost T, Lacroix C, Braegger CP, Chassard C. 2012. New insights in gut microbiota establishment in healthy breast fed neonates. PLoS One 7:e44595.
160. Martin R, Jimenez E, Heilig H, Fernandez L, Marin ML, Zoetendal EG, Rodriguez JM. 2009. Isolation of bifidobacteria from breast milk and assessment of the bifidobacterial population by PCR-denaturing gradient gel electrophoresis and quantitative real-time PCR. Appl Environ Microbiol 75:965-9.
161. Marcobal A, Sonnenburg JL. 2012. Human milk oligosaccharide consumption by intestinal microbiota. Clin Microbiol Infect 18 Suppl 4:12-5.
162. Palmer C, Bik EM, DiGiulio DB, Relman DA, Brown PO. 2007. Development of the human infant intestinal microbiota. PLoS Biol 5:e177.
163. Ley RE, Hamady M, Lozupone C, Turnbaugh PJ, Ramey RR, Bircher JS, Schlegel ML, Tucker TA, Schrenzel MD, Knight R, Gordon JI. 2008. Evolution of mammals and their gut microbes. Science 320:1647-51.
164. Turroni F, Marchesi JR, Foroni E, Gueimonde M, Shanahan F, Margolles A, van Sinderen D, Ventura M. 2009. Microbiomic analysis of the bifidobacterial population in the human distal gut. ISME J 3:745-51.
165. Milani C, Lugli GA, Turroni F, Mancabelli L, Duranti S, Viappiani A, Mangifesta M, Segata N, van Sinderen D, Ventura M. 2014. Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. FEMS Microbiol Ecol 90:493-503.
166. Lugli GA, Milani C, Mancabelli L, van Sinderen D, Ventura M. 2016. MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation. FEMS Microbiol Lett 363.
167. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455-77.
168. Nurk S, Bankevich A, Antipov D, Gurevich AA, Korobeynikov A, Lapidus A, Prjibelski AD, Pyshkin A, Sirotkin A, Sirotkin Y, Stepanauskas R, Clingenpeel SR, Woyke T, McLean JS, Lasken R, Tesler G, Alekseyev MA, Pevzner PA. 2013. Assembling single-cell genomes and mini-metagenomes from chimeric MDA products. J Comput Biol 20:714-37.
169. Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics 11:119.
170. Zhao Y, Wu J, Yang J, Sun S, Xiao J, Yu J. 2012. PGAP: pan-genomes analysis pipeline. Bioinformatics 28:416-8.

171. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-10.
172. Vlietstra WJ, Zielman R, van Dongen RM, Schultes EA, Wiesman F, Vos R, van Mulligen EM, Kors JA. 2017. Automated extraction of potential migraine biomarkers using a semantic graph. J Biomed Inform 71:178-189.
173. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059-66.
174. Chenna R, Sugawara H, Koike T, Lopez R, Gibson TJ, Higgins DG, Thompson JD. 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res 31:3497-500.
175. Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5:e11147.
176. Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, Surovcik K, Meinicke P, Merkl R. 2006. Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics 7:142.
177. Csuros M. 2010. Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. Bioinformatics 26:1910-2.
178. Turroni F, Milani C, Duranti S, Mancabelli L, Mangifesta M, Viappiani A, Lugli GA, Ferrario C, Gioiosa L, Ferrarini A, Li J, Palanza P, Delledonne M, van Sinderen D, Ventura M. 2016. Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. ISME J 10:1656-68.
179. Ferrario C, Milani C, Mancabelli L, Lugli GA, Turroni F, Duranti S, Mangifesta M, Viappiani A, Sinderen D, Ventura M. 2015. A genome-based identification approach for members of the genus Bifidobacterium. FEMS Microbiol Ecol 91.
180. Ventura M, Reniero R, Zink R. 2001. Specific identification and targeted characterization of Bifidobacterium lactis from different environmental isolates by a combined multiplex-PCR approach. Appl Environ Microbiol 67:2760-5.
181. Ventura M, Zink R. 2003. Comparative sequence analysis of the tuf and recA genes and restriction fragment length polymorphism of the internal transcribed spacer region sequences supply additional tools for discriminating Bifidobacterium lactis from Bifidobacterium animalis. Appl Environ Microbiol 69:7517-22.
182. Briczinski EP, Loquasto JR, Barrangou R, Dudley EG, Roberts AM, Roberts RF. 2009. Strain-specific genotyping of Bifidobacterium animalis subsp. lactis by using single-nucleotide polymorphisms, insertions, and deletions. Appl Environ Microbiol 75:7501-8.
183. Odamaki T, Horigome A, Sugahara H, Hashikura N, Minami J, Xiao JZ, Abe F. 2015. Comparative Genomics Revealed Genetic Diversity and Species/Strain-Level Differences in Carbohydrate Metabolism of Three Probiotic Bifidobacterial Species. Int J Genomics 2015:567809.
184. Barrangou R, Briczinski EP, Traeger LL, Loquasto JR, Richards M, Horvath P, Coute-Monvoisin AC, Leyer G, Rendulic S, Steele JL, Broadbent JR, Oberg T, Dudley EG, Schuster S, Romero DA, Roberts RF. 2009. Comparison of the complete genome sequences of Bifidobacterium animalis subsp. lactis DSM 10140 and Bl-04. J Bacteriol 191:4144-51.
185. Tettelin H, Riley D, Cattuto C, Medini D. 2008. Comparative genomics: the bacterial pan-genome. Curr Opin Microbiol 11:472-7.
186. Tettelin H, Masignani V, Cieslewicz MJ, Donati C, Medini D, Ward NL, Angiuoli SV, Crabtree J, Jones AL, Durkin AS, Deboy RT, Davidsen TM, Mora M, Scarselli M, Margarit y Ros I, Peterson JD, Hauser CR, Sundaram JP, Nelson WC, Madupu R, Brinkac LM, Dodson RJ, Rosovitz MJ, Sullivan SA, Daugherty SC, Haft DH, Selengut J, Gwinn ML, Zhou L, Zafar N, Khouri H, Radune D, Dimitrov G, Watkins K, O'Connor KJ, Smith S, Utterback TR, White O, Rubens CE, Grandi G, Madoff LC, Kasper DL, Telford JL, Wessels MR, Rappuoli R, Fraser CM. 2005. Genome analysis of multiple pathogenic isolates of Streptococcus

agalactiae: implications for the microbial "pan-genome". Proc Natl Acad Sci U S A 102:13950-5.
187. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldon T, Rattei T, Creevey C, Kuhn M, Jensen LJ, von Mering C, Bork P. 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. Nucleic Acids Res 42:D231-9.
188. Richter M, Rossello-Mora R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126-31.
189. Jungersen M, Wind A, Johansen E, Christensen JE, Stuer-Lauridsen B, Eskesen D. 2014. The Science behind the Probiotic Strain Bifidobacterium animalis subsp. lactis BB-12((R)). Microorganisms 2:92-110.
190. Eales J, Gibson P, Whorwell P, Kellow J, Yellowlees A, Perry RH, Edwards M, King S, Wood H, Glanville J. 2017. Systematic review and meta-analysis: the effects of fermented milk with Bifidobacterium lactis CNCM I-2494 and lactic acid bacteria on gastrointestinal discomfort in the general adult population. Therap Adv Gastroenterol 10:74-88.
191. Lugli GA, Milani C, Turroni F, Duranti S, Mancabelli L, Mangifesta M, Ferrario C, Modesto M, Mattarelli P, Jiri K, van Sinderen D, Ventura M. 2017. Comparative genomic and phylogenomic analyses of the Bifidobacteriaceae family. BMC Genomics 18:568.
192. Darling AC, Mau B, Blattner FR, Perna NT. 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res 14:1394-403.
193. Turroni F, Strati F, Foroni E, Serafini F, Duranti S, van Sinderen D, Ventura M. 2012. Analysis of predicted carbohydrate transport systems encoded by Bifidobacterium bifidum PRL2010. Appl Environ Microbiol 78:5002-12.
194. Egan M, Bottacini F, O'Connell Motherway M, Casey PG, Morrissey R, Melgar S, Faurie JM, Chervaux C, Smokvina T, van Sinderen D. 2018. Staying alive: growth and survival of Bifidobacterium animalis subsp. animalis under in vitro and in vivo conditions. Appl Microbiol Biotechnol doi:10.1007/s00253-018-9413-7.
195. Landman OE. 1957. Properties and induction of beta-galactosidase in Bacillus megaterium. Biochim Biophys Acta 23:558-69.
196. Nanjo F, Katsumi R, Sakai K. 1990. Purification and characterization of an exo-beta-D-glucosaminidase, a novel type of enzyme, from Nocardia orientalis. J Biol Chem 265:10088-94.
197. Chinchetru MA, Cabezas JA, Calvo P. 1989. Purification and characterization of a broad specificity beta-glucosidase from sheep liver. Int J Biochem 21:469-76.
198. Weinstein L, Albersheim P. 1979. Structure of Plant Cell Walls: IX. Purification and Partial Characterization of a Wall-degrading Endo-Arabanase and an Arabinosidase from Bacillus subtilis. Plant Physiol 63:425-32.
199. Loquasto JR, Barrangou R, Dudley EG, Roberts RF. 2011. Short communication: the complete genome sequence of Bifidobacterium animalis subspecies animalis ATCC 25527(T) and comparative analysis of growth in milk with B. animalis subspecies lactis DSM 10140(T). J Dairy Sci 94:5864-70.
200. Uchimura Y, Wyss M, Brugiroux S, Limenitakis JP, Stecher B, McCoy KD, Macpherson AJ. 2016. Complete Genome Sequences of 12 Species of Stable Defined Moderately Diverse Mouse Microbiota 2. Genome Announc 4.
201. Kim JF, Jeong H, Yu DS, Choi SH, Hur CG, Park MS, Yoon SH, Kim DW, Ji GE, Park HS, Oh TK. 2009. Genome sequence of the probiotic bacterium Bifidobacterium animalis subsp. lactis AD011. J Bacteriol 191:678-9.
202. Loquasto JR, Barrangou R, Dudley EG, Stahl B, Chen C, Roberts RF. 2013. Bifidobacterium animalis subsp. lactis ATCC 27673 is a genomically unique strain within its conserved subspecies. Appl Environ Microbiol 79:6903-10.
203. Stahl B, Barrangou R. 2012. Complete genome sequences of probiotic strains Bifidobacterium animalis subsp. lactis B420 and Bi-07. J Bacteriol 194:4131-2.

204. Garrigues C, Johansen E, Pedersen MB. 2010. Complete genome sequence of Bifidobacterium animalis subsp. lactis BB-12, a widely consumed probiotic strain. J Bacteriol 192:2467-8.
205. Kang J, Chung WH, Lim TJ, Lim S, Nam YD. 2017. Complete genome sequence of the Bifidobacterium animalis subspecies lactis BL3, preventive probiotics for acute colitis and colon cancer. New Microbes New Infect 19:34-37.
206. Bottacini F, Dal Bello F, Turroni F, Milani C, Duranti S, Foroni E, Viappiani A, Strati F, Mora D, van Sinderen D, Ventura M. 2011. Complete genome sequence of Bifidobacterium animalis subsp. lactis BLC1. J Bacteriol 193:6387-8.
207. Chervaux C, Grimaldi C, Bolotin A, Quinquis B, Legrain-Raspaud S, van Hylckama Vlieg JE, Denariaz G, Smokvina T. 2011. Genome sequence of the probiotic strain Bifidobacterium animalis subsp. lactis CNCM I-2494. J Bacteriol 193:5560-1.
208. Carnevali L, Montano N, Statello R, Coude G, Vacondio F, Rivara S, Ferrari PF, Sgoifo A. 2017. Social stress contagion in rats: Behavioural, autonomic and neuroendocrine correlates. Psychoneuroendocrinology 82:155-163.
209. Turroni F, Ozcan E, Milani C, Mancabelli L, Viappiani A, van Sinderen D, Sela DA, Ventura M. 2015. Glycan cross-feeding activities between bifidobacteria under in vitro conditions. Front Microbiol 6:1030.
210. Alessandri G, Milani C, Duranti S, Mancabelli L, Ranjanoro T, Modica S, Carnevali L, Statello R, Bottacini F, Turroni F, Ossiprandi MC, Sgoifo A, van Sinderen D, Ventura M. 2019. Ability of bifidobacteria to metabolize chitin-glucan and its impact on the gut microbiota. Sci Rep 9:5755.
211. Ventura M, Elli M, Reniero R, Zink R. 2001. Molecular microbial analysis of Bifidobacterium isolates from different environments by the species-specific amplified ribosomal DNA restriction analysis (ARDRA). FEMS Microbiol Ecol 36:113-121.
212. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. 2010. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335-6.
213. Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460-1.
214. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. Microbiome 6:90.
215. Duranti S, Lugli GA, Milani C, James K, Mancabelli L, Turroni F, Alessandri G, Mangifesta M, Mancino W, Ossiprandi MC, Iori A, Rota C, Gargano G, Bernasconi S, Di Pierro F, van Sinderen D, Ventura M. 2019. Bifidobacterium bifidum and the infant gut microbiota: an intriguing case of microbe-host co-evolution. Environ Microbiol doi:10.1111/1462-2920.14705.
216. Duranti S, Mancabelli L, Mancino W, Anzalone R, Longhi G, Statello R, Carnevali L, Sgoifo A, Bernasconi S, Turroni F, Ventura M. 2019. Exploring the effects of COLOSTRONONI on the mammalian gut microbiota composition. PLoS One 14:e0217609.
217. Duranti S, Gaiani F, Mancabelli L, Milani C, Grandi A, Bolchi A, Santoni A, Lugli GA, Ferrario C, Mangifesta M, Viappiani A, Bertoni S, Vivo V, Serafini F, Barbaro MR, Fugazza A, Barbara G, Gioiosa L, Palanza P, Cantoni AM, de'Angelis GL, Barocelli E, de'Angelis N, van Sinderen D, Ventura M, Turroni F. 2016. Elucidating the gut microbiome of ulcerative colitis: bifidobacteria as novel microbial biomarkers. FEMS Microbiol Ecol 92.
218. Egan M, Motherway MO, Kilcoyne M, Kane M, Joshi L, Ventura M, van Sinderen D. 2014. Cross-feeding by Bifidobacterium breve UCC2003 during co-cultivation with Bifidobacterium bifidum PRL2010 in a mucin-based medium. BMC Microbiol 14:282.

219. Ferrario C, Duranti S, Milani C, Mancabelli L, Lugli GA, Turroni F, Mangifesta M, Viappiani A, Ossiprandi MC, van Sinderen D, Ventura M. 2015. Exploring Amino Acid Auxotrophy in Bifidobacterium bifidum PRL2010. Front Microbiol 6:1331.
220. Serafini F, Strati F, Ruas-Madiedo P, Turroni F, Foroni E, Duranti S, Milano F, Perotti A, Viappiani A, Guglielmetti S, Buschini A, Margolles A, van Sinderen D, Ventura M. 2013. Evaluation of adhesion properties and antibacterial activities of the infant gut commensal Bifidobacterium bifidum PRL2010. Anaerobe 21:9-17.
221. Serafini F, Turroni F, Ruas-Madiedo P, Lugli GA, Milani C, Duranti S, Zamboni N, Bottacini F, van Sinderen D, Margolles A, Ventura M. 2014. Kefir fermented milk and kefiran promote growth of Bifidobacterium bifidum PRL2010 and modulate its gene expression. Int J Food Microbiol 178:50-9.
222. Turroni F, Taverniti V, Ruas-Madiedo P, Duranti S, Guglielmetti S, Lugli GA, Gioiosa L, Palanza P, Margolles A, van Sinderen D, Ventura M. 2014. Bifidobacterium bifidum PRL2010 modulates the host innate immune response. Appl Environ Microbiol 80:730-40.
223. Turroni F, Foroni E, Montanini B, Viappiani A, Strati F, Duranti S, Ferrarini A, Delledonne M, van Sinderen D, Ventura M. 2011. Global genome transcription profiling of Bifidobacterium bifidum PRL2010 under in vitro conditions and identification of reference genes for quantitative real-time PCR. Appl Environ Microbiol 77:8578-87.
224. Turroni F, Serafini F, Mangifesta M, Arioli S, Mora D, van Sinderen D, Ventura M. 2014. Expression of sortase-dependent pili of Bifidobacterium bifidum PRL2010 in response to environmental gut conditions. FEMS Microbiol Lett 357:23-33.
225. Jimenez E, Fernandez L, Marin ML, Martin R, Odriozola JM, Nueno-Palop C, Narbad A, Olivares M, Xaus J, Rodriguez JM. 2005. Isolation of commensal bacteria from umbilical cord blood of healthy neonates born by cesarean section. Curr Microbiol 51:270-4.
226. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M. 2006. ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res 34:D32-6.
227. Lakin SM, Dean C, Noyes NR, Dettenwanger A, Ross AS, Doster E, Rovira P, Abdo Z, Jones KL, Ruiz J, Belk KE, Morley PS, Boucher C. 2017. MEGARes: an antimicrobial resistance database for high throughput sequencing. Nucleic Acids Res 45:D574-D580.
228. Gupta SK, Padmanabhan BR, Diene SM, Lopez-Rojas R, Kempf M, Landraud L, Rolain JM. 2014. ARG-ANNOT, a new bioinformatic tool to discover antibiotic resistance genes in bacterial genomes. Antimicrob Agents Chemother 58:212-20.
229. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. 2012. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother 67:2640-4.
230. McArthur AG, Waglechner N, Nizam F, Yan A, Azad MA, Baylay AJ, Bhullar K, Canova MJ, De Pascale G, Ejim L, Kalan L, King AM, Koteva K, Morar M, Mulvey MR, O'Brien JS, Pawlowski AC, Piddock LJ, Spanogiannopoulos P, Sutherland AD, Tang I, Taylor PL, Thaker M, Wang W, Yan M, Yu T, Wright GD. 2013. The comprehensive antibiotic resistance database. Antimicrob Agents Chemother 57:3348-57.
231. Bush K, Jacoby GA. 2010. Updated functional classification of beta-lactamases. Antimicrob Agents Chemother 54:969-76.
232. Duranti S, Lugli GA, Mancabelli L, Turroni F, Milani C, Mangifesta M, Ferrario C, Anzalone R, Viappiani A, van Sinderen D, Ventura M. 2017. Prevalence of Antibiotic Resistance Genes among Human Gut-Derived Bifidobacteria. Appl Environ Microbiol 83.
233. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. 2011. Adaptive seeds tame genomic sequence comparison. Genome Res 21:487-93.
234. Authority EFS. 2012. Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. EFSA Journal 10:2740.
235. Arthur M, Depardieu F, Molinas C, Reynolds P, Courvalin P. 1995. The vanZ gene of Tn1546 from Enterococcus faecium BM4147 confers resistance to teicoplanin. Gene 154:87-92.

236. Bugg TD, Wright GD, Dutka-Malen S, Arthur M, Courvalin P, Walsh CT. 1991. Molecular basis for vancomycin resistance in Enterococcus faecium BM4147: biosynthesis of a depsipeptide peptidoglycan precursor by vancomycin resistance proteins VanH and VanA. Biochemistry 30:10408-15.
237. Evers S, Courvalin P. 1996. Regulation of VanB-type vancomycin resistance gene expression by the VanS(B)-VanR (B) two-component regulatory system in Enterococcus faecalis V583. J Bacteriol 178:1302-9.
238. Giovanetti E, Brenciani A, Lupidi R, Roberts MC, Varaldo PE. 2003. Presence of the tet(O) gene in erythromycin- and tetracycline-resistant strains of Streptococcus pyogenes and linkage with either the mef(A) or the erm(A) gene. Antimicrob Agents Chemother 47:2844-9.
239. Roberts MC. 1996. Tetracycline resistance determinants: mechanisms of action, regulation of expression, genetic mobility, and distribution. FEMS Microbiol Rev 19:1-24.
240. Hedayatianfard K, Akhlaghi M, Sharifiyazdi H. 2014. Detection of tetracycline resistance genes in bacteria isolated from fish farms using polymerase chain reaction. Vet Res Forum 5:269-75.
241. Wang N, Hang X, Zhang M, Liu X, Yang H. 2017. Analysis of newly detected tetracycline resistance genes and their flanking sequences in human intestinal bifidobacteria. Sci Rep 7:6267.
242. Zou Y, Xue W, Luo G, Deng Z, Qin P, Guo R, Sun H, Xia Y, Liang S, Dai Y, Wan D, Jiang R, Su L, Feng Q, Jie Z, Guo T, Xia Z, Liu C, Yu J, Lin Y, Tang S, Huo G, Xu X, Hou Y, Liu X, Wang J, Yang H, Kristiansen K, Li J, Jia H, Xiao L. 2019. 1,520 reference genomes from cultivated human gut bacteria enable functional microbiome analyses. Nat Biotechnol 37:179-185.
243. Martinez N, Luque R, Milani C, Ventura M, Banuelos O, Margolles A. 2018. A Gene Homologous to rRNA Methylase Genes Confers Erythromycin and Clindamycin Resistance in Bifidobacterium breve. Appl Environ Microbiol 84.
244. Skold O. 2001. Resistance to trimethoprim and sulfonamides. Vet Res 32:261-73.
245. Phuong Hoa PT, Nonaka L, Hung Viet P, Suzuki S. 2008. Detection of the sul1, sul2, and sul3 genes in sulfonamide-resistant bacteria from wastewater and shrimp ponds of north Vietnam. Sci Total Environ 405:377-84.
246. Vetting MW, Hegde SS, Fajardo JE, Fiser A, Roderick SL, Takiff HE, Blanchard JS. 2006. Pentapeptide repeat proteins. Biochemistry 45:1-10.
247. Merens A, Matrat S, Aubry A, Lascols C, Jarlier V, Soussy CJ, Cavallo JD, Cambau E. 2009. The pentapeptide repeat proteins MfpAMt and QnrB4 exhibit opposite effects on DNA gyrase catalytic reactions and on the ternary gyrase-DNA-quinolone complex. J Bacteriol 191:1587-94.
248. Park KS, Lee JH, Jeong DU, Lee JJ, Wu X, Jeong BC, Kang CM, Lee SH. 2011. Determination of pentapeptide repeat units in Qnr proteins by the structure-based alignment approach. Antimicrob Agents Chemother 55:4475-8.
249. Gueimonde M, Sanchez B, C GdLR-G, Margolles A. 2013. Antibiotic resistance in probiotic bacteria. Front Microbiol 4:202.
250. Delcour J, Ferain T, Deghorain M, Palumbo E, Hols P. 1999. The biosynthesis and functionality of the cell-wall of lactic acid bacteria. Antonie Van Leeuwenhoek 76:159-84.
251. Florez AB, Ladero V, Alvarez-Martin P, Ammor MS, Alvarez MA, Mayo B. 2007. Acquired macrolide resistance in the human intestinal strain Lactobacillus rhamnosus E41 associated with a transition mutation in 23S rRNA genes. Int J Antimicrob Agents 30:341-4.
252. Hummel AS, Hertel C, Holzapfel WH, Franz CM. 2007. Antibiotic resistances of starter and probiotic strains of lactic acid bacteria. Appl Environ Microbiol 73:730-9.
253. Ammor MS, Gueimonde M, Danielsen M, Zagorec M, van Hoek AH, de Los Reyes-Gavilan CG, Mayo B, Margolles A. 2008. Two different tetracycline resistance mechanisms, plasmid-

carried tet(L) and chromosomally located transposon-associated tet(M), coexist in Lactobacillus sakei Rits 9. Appl Environ Microbiol 74:1394-401.
254. Rojo-Bezares B, Saenz Y, Poeta P, Zarazaga M, Ruiz-Larrea F, Torres C. 2006. Assessment of antibiotic susceptibility within lactic acid bacteria strains isolated from wine. Int J Food Microbiol 111:234-40.
255. Johnning A, Karami N, Tang Hallback E, Muller V, Nyberg L, Buongermino Pereira M, Stewart C, Ambjornsson T, Westerlund F, Adlerberth I, Kristiansson E. 2018. The resistomes of six carbapenem-resistant pathogens - a critical genotype-phenotype analysis. Microb Genom 4.
256. Dagher C, Salloum T, Alousi S, Arabaghian H, Araj GF, Tokajian S. 2018. Molecular characterization of Carbapenem resistant Escherichia coli recovered from a tertiary hospital in Lebanon. PLoS One 13:e0203323.
257. Khan I, Yasir M, Farman M, Kumosani T, AlBasri SF, Bajouh OS, Azhar EI. 2019. Evaluation of gut bacterial community composition and antimicrobial resistome in pregnant and non-pregnant women from Saudi population. Infect Drug Resist 12:1749-1761.
258. Hong HA, Duc le H, Cutting SM. 2005. The use of bacterial spore formers as probiotics. FEMS Microbiol Rev 29:813-35.
259. Monod M, Denoya C, Dubnau D. 1986. Sequence and properties of pIM13, a macrolide-lincosamide-streptogramin B resistance plasmid from Bacillus subtilis. J Bacteriol 167:138-47.
260. Phelan RW, Clarke C, Morrissey JP, Dobson AD, O'Gara F, Barbosa TM. 2011. Tetracycline resistance-encoding plasmid from Bacillus sp. strain #24, isolated from the marine sponge Haliclona simulans. Appl Environ Microbiol 77:327-9.
261. Dai L, Wu CM, Wang MG, Wang Y, Wang Y, Huang SY, Xia LN, Li BB, Shen JZ. 2010. First report of the multidrug resistance gene cfr and the phenicol resistance gene fexA in a Bacillus strain from swine feces. Antimicrob Agents Chemother 54:3953-5.
262. Siefert JL. 2009. Defining the mobilome. Methods Mol Biol 532:13-27.
263. Smets BF, Barkay T. 2005. Horizontal gene transfer: perspectives at a crossroads of scientific disciplines. Nat Rev Microbiol 3:675-8.
264. Guglielmetti S, Mayo B, Alvarez-Martin P. 2013. Mobilome and genetic modification of bifidobacteria. Benef Microbes 4:143-66.
265. Mahony J, Lugli GA, van Sinderen D, Ventura M. 2018. Impact of gut-associated bifidobacteria and their phages on health: two sides of the same coin? Appl Microbiol Biotechnol 102:2091-2099.
266. Siguier P, Varani A, Perochon J, Chandler M. 2012. Exploring bacterial insertion sequences with ISfinder: objectives, uses, and future developments. Methods Mol Biol 859:91-103.
267. Mavrich TN, Casey E, Oliveira J, Bottacini F, James K, Franz C, Lugli GA, Neve H, Ventura M, Hatfull GF, Mahony J, van Sinderen D. 2018. Characterization and induction of prophages in human gut-associated Bifidobacterium hosts. Sci Rep 8:12772.
268. Bottacini F, Medini D, Pavesi A, Turroni F, Foroni E, Riley D, Giubellini V, Tettelin H, van Sinderen D, Ventura M. 2010. Comparative genomics of the genus Bifidobacterium. Microbiology 156:3243-54.
269. Menouni R, Hutinet G, Petit MA, Ansaldi M. 2015. Bacterial genome remodeling through bacteriophage recombination. FEMS Microbiol Lett 362:1-10.
270. Bondy-Denomy J, Davidson AR. 2014. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. J Microbiol 52:235-42.
271. Florez AB, Ammor MS, Alvarez-Martin P, Margolles A, Mayo B. 2006. Molecular analysis of tet(W) gene-mediated tetracycline resistance in dominant intestinal Bifidobacterium species from healthy humans. Appl Environ Microbiol 72:7377-9.

272. Gueimonde M, Florez AB, van Hoek AH, Stuer-Lauridsen B, Stroman P, de los Reyes-Gavilan CG, Margolles A. 2010. Genetic basis of tetracycline resistance in Bifidobacterium animalis subsp. lactis. Appl Environ Microbiol 76:3364-9.
273. Scott KP, Barbosa TM, Forbes KJ, Flint HJ. 1997. High-frequency transfer of a naturally occurring chromosomal tetracycline resistance element in the ruminal anaerobe Butyrivibrio fibrisolvens. Appl Environ Microbiol 63:3405-11.
274. Cain BD, Norton PJ, Eubanks W, Nick HS, Allen CM. 1993. Amplification of the bacA gene confers bacitracin resistance to Escherichia coli. J Bacteriol 175:3784-9.
275. El Ghachi M, Bouhss A, Blanot D, Mengin-Lecreulx D. 2004. The bacA gene of Escherichia coli encodes an undecaprenyl pyrophosphate phosphatase activity. J Biol Chem 279:30106-13.
276. Zhurina D, Dudnik A, Waidmann MS, Grimm V, Westermann C, Breitinger KJ, Yuan J, van Sinderen D, Riedel CU. 2013. High-Quality Draft Genome Sequence of Bifidobacterium longum E18, Isolated from a Healthy Adult. Genome Announc 1.
277. Wright LD, Grossman AD. 2016. Autonomous Replication of the Conjugative Transposon Tn916. J Bacteriol 198:3355-3366.
278. Cury J, Touchon M, Rocha EPC. 2017. Integrative and conjugative elements and their hosts: composition, distribution and organization. Nucleic Acids Res 45:8943-8956.
279. Moradigaravand D, Palm M, Farewell A, Mustonen V, Warringer J, Parts L. 2018. Prediction of antibiotic resistance in Escherichia coli from large-scale pan-genome data. PLoS Comput Biol 14:e1006258.
280. Ventura M, Turroni F, Lugli GA, van Sinderen D. 2014. Bifidobacteria and humans: our special friends, from ecological to genomics perspectives. J Sci Food Agric 94:163-8.
281. Mayrhofer S, van Hoek AH, Mair C, Huys G, Aarts HJ, Kneifel W, Domig KJ. 2010. Antibiotic susceptibility of members of the Lactobacillus acidophilus group using broth microdilution and molecular identification of their resistance determinants. Int J Food Microbiol 144:81-7.
282. Lin CF, Fung ZF, Wu CL, Chung TC. 1996. Molecular characterization of a plasmid-borne (pTC82) chloramphenicol resistance determinant (cat-TC) from Lactobacillus reuteri G4. Plasmid 36:116-24.
283. Parnanen K, Karkman A, Hultman J, Lyra C, Bengtsson-Palme J, Larsson DGJ, Rautava S, Isolauri E, Salminen S, Kumar H, Satokari R, Virta M. 2018. Maternal gut and breast milk microbiota affect infant gut antibiotic resistome and mobile genetic elements. Nat Commun 9:3891.
284. Milani C, Hevia A, Foroni E, Duranti S, Turroni F, Lugli GA, Sanchez B, Martin R, Gueimonde M, van Sinderen D, Margolles A, Ventura M. 2013. Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. PLoS One 8:e68739.
285. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods 13:581-3.
286. Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. Appl Environ Microbiol 71:8228-35.
287. Vandeputte D, Kathagen G, D'Hoe K, Vieira-Silva S, Valles-Colomer M, Sabino J, Wang J, Tito RY, De Commer L, Darzi Y, Vermeire S, Falony G, Raes J. 2017. Quantitative microbiome profiling links gut community variation to microbial load. Nature 551:507-511.
288. Mancino W, Lugli GA, Sinderen DV, Ventura M, Turroni F. 2019. Mobilome and Resistome Reconstruction from Genomes Belonging to Members of the Bifidobacterium Genus. Microorganisms 7.
289. Milani C, Duranti S, Napoli S, Alessandri G, Mancabelli L, Anzalone R, Longhi G, Viappiani A, Mangifesta M, Lugli GA, Bernasconi S, Ossiprandi MC, van Sinderen D, Ventura M,

Turroni F. 2019. Colonization of the human gut by bovine bacteria present in Parmesan cheese. Nat Commun 10:1286.
290. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357-9.
291. Carver T, Harris SR, Berriman M, Parkhill J, McQuillan JA. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28:464-9.
292. Macfarlane GT, Macfarlane S, Gibson GR. 1998. Validation of a Three-Stage Compound Continuous Culture System for Investigating the Effect of Retention Time on the Ecology and Metabolism of Bacteria in the Human Colon. Microb Ecol 35:180-7.
293. Maier L, Pruteanu M, Kuhn M, Zeller G, Telzerow A, Anderson EE, Brochado AR, Fernandez KC, Dose H, Mori H, Patil KR, Bork P, Typas A. 2018. Extensive impact of non-antibiotic drugs on human gut bacteria. Nature 555:623-628.
294. Young VB, Schmidt TM. 2004. Antibiotic-associated diarrhea accompanied by large-scale alterations in the composition of the fecal microbiota. J Clin Microbiol 42:1203-6.
295. Espinosa-Gongora C, Jessen LR, Kieler IN, Damborg P, Bjornvad CR, Gudeta DD, Pires Dos Santos T, Sablier-Gallis F, Sayah-Jeanne S, Corbel T, Neviere A, Hugon P, Saint-Lu N, de Gunzburg J, Guardabassi L. 2020. Impact of oral amoxicillin and amoxicillin/clavulanic acid treatment on bacterial diversity and beta-lactam resistance in the canine faecal microbiota. J Antimicrob Chemother 75:351-361.
296. MacPherson CW, Mathieu O, Tremblay J, Champagne J, Nantel A, Girard SA, Tompkins TA. 2018. Gut Bacterial Microbiota and its Resistome Rapidly Recover to Basal State Levels after Short-term Amoxicillin-Clavulanic Acid Treatment in Healthy Adults. Sci Rep 8:11192.
297. Mangin I, Leveque C, Magne F, Suau A, Pochart P. 2012. Long-term changes in human colonic Bifidobacterium populations induced by a 5-day oral amoxicillin-clavulanic acid treatment. PLoS One 7:e50257.
298. Turroni F, Milani C, Duranti S, Mahony J, van Sinderen D, Ventura M. 2018. Glycan Utilization and Cross-Feeding Activities by Bifidobacteria. Trends Microbiol 26:339-350.
299. Hidalgo-Cantabrana C, Delgado S, Ruiz L, Ruas-Madiedo P, Sanchez B, Margolles A. 2017. Bifidobacteria and Their Health-Promoting Effects. Microbiol Spectr 5.
300. Francino MP. 2015. Antibiotics and the Human Gut Microbiome: Dysbioses and Accumulation of Resistances. Front Microbiol 6:1543.
301. Tanaka S, Kobayashi T, Songjinda P, Tateyama A, Tsubouchi M, Kiyohara C, Shirakawa T, Sonomoto K, Nakayama J. 2009. Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota. FEMS Immunol Med Microbiol 56:80-7.
302. Duranti S, Milani C, Lugli GA, Mancabelli L, Turroni F, Ferrario C, Mangifesta M, Viappiani A, Sanchez B, Margolles A, van Sinderen D, Ventura M. 2016. Evaluation of genetic diversity among strains of the human gut commensal Bifidobacterium adolescentis. Sci Rep 6:23971.
303. Bottacini F, O'Connell Motherway M, Kuczynski J, O'Connell KJ, Serafini F, Duranti S, Milani C, Turroni F, Lugli GA, Zomer A, Zhurina D, Riedel C, Ventura M, van Sinderen D. 2014. Comparative genomics of the Bifidobacterium breve taxon. BMC Genomics 15:170.
304. Duytschaever G, Huys G, Boulanger L, De Boeck K, Vandamme P. 2013. Amoxicillin-clavulanic acid resistance in fecal Enterobacteriaceae from patients with cystic fibrosis and healthy siblings. J Cyst Fibros 12:780-3.
305. Nakano V, Nascimento e Silva A, Merino VR, Wexler HM, Avila-Campos MJ. 2011. Antimicrobial resistance and prevalence of resistance genes in intestinal Bacteroidales strains. Clinics (Sao Paulo) 66:543-7.
306. Huddleston JR. 2014. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. Infect Drug Resist 7:167-76.

307. Gudeta DD, Moodley A, Bortolaia V, Guardabassi L. 2014. vanO, a new glycopeptide resistance operon in environmental Rhodococcus equi isolates. Antimicrob Agents Chemother 58:1768-70.
308. Gunn JS, Lim KB, Krueger J, Kim K, Guo L, Hackett M, Miller SI. 1998. PmrA-PmrB-regulated genes necessary for 4-aminoarabinose lipid A modification and polymyxin resistance. Mol Microbiol 27:1171-82.
309. Fouhy F, Guinane CM, Hussey S, Wall R, Ryan CA, Dempsey EM, Murphy B, Ross RP, Fitzgerald GF, Stanton C, Cotter PD. 2012. High-throughput sequencing reveals the incomplete, short-term recovery of infant gut microbiota following parenteral antibiotic treatment with ampicillin and gentamicin. Antimicrob Agents Chemother 56:5811-20.
310. Milani C, Casey E, Lugli GA, Moore R, Kaczorowska J, Feehily C, Mangifesta M, Mancabelli L, Duranti S, Turroni F, Bottacini F, Mahony J, Cotter PD, McAuliffe FM, van Sinderen D, Ventura M. 2018. Tracing mother-infant transmission of bacteriophages by means of a novel analytical tool for shotgun metagenomic datasets: METAnnotatorX. Microbiome 6:145.
311. Song Y, Kato N, Liu C, Matsumiya Y, Kato H, Watanabe K. 2000. Rapid identification of 11 human intestinal Lactobacillus species by multiplex PCR assays using group- and species-specific primers derived from the 16S-23S rRNA intergenic spacer region and its flanking 23S rRNA. FEMS Microbiol Lett 187:167-73.
312. Brolazo EM, Leite DS, Tiba MR, Villarroel M, Marconi C, Simoes JA. 2011. Correlation between api 50 ch and multiplex polymerase chain reaction for the identification of vaginal lactobacilli in isolates. Braz J Microbiol 42:225-32.
313. Zhang R, Daroczy K, Xiao B, Yu L, Chen R, Liao Q. 2012. Qualitative and semiquantitative analysis of Lactobacillus species in the vaginas of healthy fertile and postmenopausal Chinese women. J Med Microbiol 61:729-739.
314. Yeruva T, Rajkumar H, Donugama V. 2017. Vaginal lactobacilli profile in pregnant women with normal & abnormal vaginal flora. Indian J Med Res 146:534-540.
315. Rabe LK, Hillier SL. 2003. Optimization of media for detection of hydrogen peroxide production by Lactobacillus species. J Clin Microbiol 41:3260-4.
316. Hutt P, Lapp E, Stsepetova J, Smidt I, Taelma H, Borovkova N, Oopkaup H, Ahelik A, Roop T, Hoidmets D, Samuel K, Salumets A, Mandar R. 2016. Characterisation of probiotic properties in human vaginal lactobacilli strains. Microb Ecol Health Dis 27:30484.
317. Jain C, Rodriguez RL, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nat Commun 9:5114.
318. van Heel AJ, de Jong A, Montalban-Lopez M, Kok J, Kuipers OP. 2013. BAGEL3: Automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acids Res 41:W448-53.
319. Pan M, Hidalgo-Cantabrana C, Goh YJ, Sanozky-Dawes R, Barrangou R. 2019. Comparative Analysis of Lactobacillus gasseri and Lactobacillus crispatus Isolated From Human Urogenital and Gastrointestinal Tracts. Front Microbiol 10:3146.
320. Mancino W, Duranti S, Mancabelli L, Longhi G, Anzalone R, Milani C, Lugli GA, Carnevali L, Statello R, Sgoifo A, van Sinderen D, Ventura M, Turroni F. 2019. Bifidobacterial Transfer from Mother to Child as Examined by an Animal Model. Microorganisms 7.
321. Gajer P, Brotman RM, Bai G, Sakamoto J, Schutte UM, Zhong X, Koenig SS, Fu L, Ma ZS, Zhou X, Abdo Z, Forney LJ, Ravel J. 2012. Temporal dynamics of the human vaginal microbiota. Sci Transl Med 4:132ra52.
322. DiGiulio DB, Callahan BJ, McMurdie PJ, Costello EK, Lyell DJ, Robaczewska A, Sun CL, Goltsman DS, Wong RJ, Shaw G, Stevenson DK, Holmes SP, Relman DA. 2015. Temporal and spatial variation of the human microbiota during pregnancy. Proc Natl Acad Sci U S A 112:11060-5.

323. Ma ZS, Li L. 2017. Quantifying the human vaginal community state types (CSTs) with the species specificity index. PeerJ 5:e3366.
324. Ma B, Forney LJ, Ravel J. 2012. Vaginal microbiome: rethinking health and disease. Annu Rev Microbiol 66:371-89.
325. De Seta F, Campisciano G, Zanotta N, Ricci G, Comar M. 2019. The Vaginal Community State Types Microbiome-Immune Network as Key Factor for Bacterial Vaginosis and Aerobic Vaginitis. Front Microbiol 10:2451.
326. Hillmann B, Al-Ghalith GA, Shields-Cutler RR, Zhu Q, Gohl DM, Beckman KB, Knight R, Knights D. 2018. Evaluating the Information Content of Shallow Shotgun Metagenomics. mSystems 3.
327. Mendling W. 2016. Vaginal Microbiota. Adv Exp Med Biol 902:83-93.
328. Alessandri G, Milani C, Mancabelli L, Mangifesta M, Lugli GA, Viappiani A, Duranti S, Turroni F, Ossiprandi MC, van Sinderen D, Ventura M. 2019. Metagenomic dissection of the canine gut microbiota: insights into taxonomic, metabolic and nutritional features. Environ Microbiol 21:1331-1343.
329. Alessandri G, Milani C, Mancabelli L, Longhi G, Anzalone R, Lugli GA, Duranti S, Turroni F, Ossiprandi MC, van Sinderen D, Ventura M. 2020. Deciphering the Bifidobacterial Populations within the Canine and Feline Gut Microbiota. Appl Environ Microbiol 86.
330. Mancabelli L, Milani C, Lugli GA, Turroni F, Ferrario C, van Sinderen D, Ventura M. 2017. Meta-analysis of the human gut microbiome from urbanized and pre-agricultural populations. Environ Microbiol 19:1379-1390.
331. Hidalgo-Cantabrana C, Goh YJ, Pan M, Sanozky-Dawes R, Barrangou R. 2019. Genome editing using the endogenous type I CRISPR-Cas system in Lactobacillus crispatus. Proc Natl Acad Sci U S A 116:15774-15783.
332. Adhikari B, Kwon YM. 2017. Characterization of the Culturable Subpopulations of Lactobacillus in the Chicken Intestinal Tract as a Resource for Probiotic Development. Front Microbiol 8:1389.
333. Nami Y, Haghshenas B, Yari Khosroushahi A. 2018. Molecular Identification and Probiotic Potential Characterization of Lactic Acid Bacteria Isolated from Human Vaginal Microbiota. Adv Pharm Bull 8:683-695.
334. Petrova MI, van den Broek M, Balzarini J, Vanderleyden J, Lebeer S. 2013. Vaginal microbiota and its role in HIV transmission and infection. FEMS Microbiol Rev 37:762-92.
335. Rastogi R, Su J, Mahalingam A, Clark J, Sung S, Hope T, Kiser PF. 2016. Engineering and characterization of simplified vaginal and seminal fluid simulants. Contraception 93:337-346.
336. Amabebe E, Anumba DOC. 2018. The Vaginal Microenvironment: The Physiologic Role of Lactobacilli. Front Med (Lausanne) 5:181.
337. Mirmonsef P, Hotton AL, Gilbert D, Burgad D, Landay A, Weber KM, Cohen M, Ravel J, Spear GT. 2014. Free glycogen in vaginal fluids is associated with Lactobacillus colonization and low vaginal pH. PLoS One 9:e102467.
338. van der Veer C, Hertzberger RY, Bruisten SM, Tytgat HLP, Swanenburg J, de Kat Angelino-Bart A, Schuren F, Molenaar D, Reid G, de Vries H, Kort R. 2019. Comparative genomics of human Lactobacillus crispatus isolates reveals genes for glycosylation and glycogen degradation: implications for in vivo dominance of the vaginal microbiota. Microbiome 7:49.
339. Vallor AC, Antonio MA, Hawes SE, Hillier SL. 2001. Factors associated with acquisition of, or persistent colonization by, vaginal lactobacilli: role of hydrogen peroxide production. J Infect Dis 184:1431-6.
340. Chetwin E, Manhanzva MT, Abrahams AG, Froissart R, Gamieldien H, Jaspan H, Jaumdally SZ, Barnabas SL, Dabee S, Happel AU, Bowers D, Davids L, Passmore JS, Masson L. 2019. Antimicrobial and inflammatory properties of South African clinical Lactobacillus isolates and vaginal probiotics. Sci Rep 9:1917.

341. Klebanoff SJ, Hillier SL, Eschenbach DA, Waltersdorph AM. 1991. Control of the microbial flora of the vagina by H2O2-generating lactobacilli. J Infect Dis 164:94-100.
342. Hawes SE, Hillier SL, Benedetti J, Stevens CE, Koutsky LA, Wolner-Hanssen P, Holmes KK. 1996. Hydrogen peroxide-producing lactobacilli and acquisition of vaginal infections. J Infect Dis 174:1058-63.
343. Ocana VS, Pesce de Ruiz Holgado AA, Nader-Macias ME. 1999. Selection of vaginal H2O2-generating Lactobacillus species for probiotic use. Curr Microbiol 38:279-84.
344. Zheng J, Wittouck S, Salvetti E, Franz C, Harris HMB, Mattarelli P, O'Toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Ganzle MG, Lebeer S. 2020. A taxonomic note on the genus Lactobacillus: Description of 23 novel genera, emended description of the genus Lactobacillus Beijerinck 1901, and union of Lactobacillaceae and Leuconostocaceae. Int J Syst Evol Microbiol 70:2782-2858.
345. Koryszewska-Baginska A, Gawor J, Nowak A, Grynberg M, Aleksandrzak-Piekarczyk T. 2019. Comparative genomics and functional analysis of a highly adhesive dairy Lactobacillus paracasei subsp. paracasei IBB3423 strain. Appl Microbiol Biotechnol 103:7617-7634.
346. Petrova MI, Lievens E, Malik S, Imholz N, Lebeer S. 2015. Lactobacillus species as biomarkers and agents that can promote various aspects of vaginal health. Front Physiol 6:81.
347. Prabhurajeshwar C, Chandrakanth RK. 2017. Probiotic potential of Lactobacilli with antagonistic activity against pathogenic strains: An in vitro validation for the production of inhibitory substances. Biomed J 40:270-283.
348. Joerger MC, Klaenhammer TR. 1986. Characterization and purification of helveticin J and evidence for a chromosomally determined bacteriocin produced by Lactobacillus helveticus 481. J Bacteriol 167:439-46.
349. Yoo D, Bagon BB, Valeriano VDV, Oh JK, Kim H, Cho S, Kang DK. 2017. Complete genome analysis of Lactobacillus fermentum SK152 from kimchi reveals genes associated with its antimicrobial activity. FEMS Microbiol Lett 364.
350. Busarcevic M, Dalgalarrondo M. 2012. Purification and genetic characterisation of the novel bacteriocin LS2 produced by the human oral strain Lactobacillus salivarius BGHO1. Int J Antimicrob Agents 40:127-34.
351. Tosukhowong A, Zendo T, Visessanguan W, Roytrakul S, Pumpuang L, Jaresitthikunchai J, Sonomoto K. 2012. Garvieacin Q, a novel class II bacteriocin from Lactococcus garvieae BCC 43578. Appl Environ Microbiol 78:1619-23.
352. Tymoszewska A, Diep DB, Aleksandrzak-Piekarczyk T. 2018. The extracellular loop of Man-PTS subunit IID is responsible for the sensitivity of Lactococcus garvieae to garvicins A, B and C. Sci Rep 8:15790.
353. Tymoszewska A, Diep DB, Wirtek P, Aleksandrzak-Piekarczyk T. 2017. The Non-Lantibiotic Bacteriocin Garvicin Q Targets Man-PTS in a Broad Spectrum of Sensitive Bacterial Genera. Sci Rep 7:8359.
354. Fontana A, Falasconi I, Molinari P, Treu L, Basile A, Vezzi A, Campanaro S, Morelli L. 2019. Genomic Comparison of Lactobacillus helveticus Strains Highlights Probiotic Potential. Front Microbiol 10:1380.
355. Zhou X, Yang B, Stanton C, Ross RP, Zhao J, Zhang H, Chen W. 2020. Comparative analysis of Lactobacillus gasseri from Chinese subjects reveals a new species-level taxa. BMC Genomics 21:119.
356. Sun Z, Harris HM, McCann A, Guo C, Argimon S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O'Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, de Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui Y, Zhang H, O'Toole PW. 2015. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. Nat Commun 6:8322.

357. Price TK, Shaheen M, Kalesinskas L, Malki K, Hilt EE, Putonti C, Wolfe AJ. 2016. Draft Genome Sequence of a Urinary Isolate of Lactobacillus crispatus. Genome Announc 4.
358. Clabaut M, Boukerb AM, Racine PJ, Pichon C, Kremser C, Picot JP, Karsybayeva M, Redziniak G, Chevalier S, Feuilloley MGJ. 2020. Draft Genome Sequence of Lactobacillus crispatus CIP 104459, Isolated from a Vaginal Swab. Microbiol Resour Announc 9.
359. Thomas-White KJ, Kliethermes S, Rickey L, Lukacz ES, Richter HE, Moalli P, Zimmern P, Norton P, Kusek JW, Wolfe AJ, Brubaker L, National Institute of D, Digestive, Kidney Diseases Urinary Incontinence Treatment N. 2017. Evaluation of the urinary microbiota of women with uncomplicated stress urinary incontinence. Am J Obstet Gynecol 216:55 e1-55 e16.
360. Falsen E, Pascual C, Sjoden B, Ohlen M, Collins MD. 1999. Phenotypic and phylogenetic characterization of a novel Lactobacillus species from human sources: description of Lactobacillus iners sp. nov. Int J Syst Bacteriol 49 Pt 1:217-21.
361. Antonio MA, Rabe LK, Hillier SL. 2005. Colonization of the rectum by Lactobacillus species and decreased risk of bacterial vaginosis. J Infect Dis 192:394-8.
362. Brotman RM, Melendez JH, Ghanem KG. 2011. A case control study of anovaginal distance and bacterial vaginosis. Int J STD AIDS 22:231-3.
363. Hilton E, Isenberg HD, Alperstein P, France K, Borenstein MT. 1992. Ingestion of yogurt containing Lactobacillus acidophilus as prophylaxis for candidal vaginitis. Ann Intern Med 116:353-7.
364. Reid G, Bruce AW, Fraser N, Heinemann C, Owen J, Henning B. 2001. Oral probiotics can resolve urogenital infections. FEMS Immunol Med Microbiol 30:49-52.
365. Invernici MM, Salvador SL, Silva PHF, Soares MSM, Casarin R, Palioto DB, Souza SLS, Taba M, Jr., Novaes AB, Jr., Furlaneto FAC, Messora MR. 2018. Effects of Bifidobacterium probiotic on the treatment of chronic periodontitis: A randomized clinical trial. J Clin Periodontol 45:1198-1210.
366. Solano-Aguilar G, Shea-Donohue T, Madden KB, Quinones A, Beshah E, Lakshman S, Xie Y, Dawson H, Urban JF. 2018. Bifidobacterium animalis subspecies lactis modulates the local immune response and glucose uptake in the small intestine of juvenile pigs infected with the parasitic nematode Ascaris suum. Gut Microbes 9:422-436.
367. Waitzberg DL, Quilici FA, Michzputen S, Friche Passos Mdo C. 2015. The Effect of Probiotic Fermented Milk That Includes Bifidobacterium Lactis Cncm I-2494 on the Reduction of Gastrointestinal Discomfort and Symptoms in Adults: A Narrative Review. Nutr Hosp 32:501-9.
368. Lee A, Lee YJ, Yoo HJ, Kim M, Chang Y, Lee DS, Lee JH. 2017. Consumption of Dairy Yogurt Containing Lactobacillus paracasei ssp. paracasei, Bifidobacterium animalis ssp. lactis and Heat-Treated Lactobacillus plantarum Improves Immune Function Including Natural Killer Cell Activity. Nutrients 9.
369. Vaishampayan PA, Kuehl JV, Froula JL, Morgan JL, Ochman H, Francino MP. 2010. Comparative metagenomics and population dynamics of the gut microbiota in mother and infant. Genome Biol Evol 2:53-66.
370. Makino H, Kushiro A, Ishikawa E, Kubota H, Gawad A, Sakai T, Oishi K, Martin R, Ben-Amor K, Knol J, Tanaka R. 2013. Mother-to-infant transmission of intestinal bifidobacterial strains has an impact on the early development of vaginally delivered infant's microbiota. PLoS One 8:e78331.
371. Makino H, Kushiro A, Ishikawa E, Muylaert D, Kubota H, Sakai T, Oishi K, Martin R, Ben Amor K, Oozeer R, Knol J, Tanaka R. 2011. Transmission of intestinal Bifidobacterium longum subsp. longum strains from mother to infant, determined by multilocus sequencing typing and amplified fragment length polymorphism. Appl Environ Microbiol 77:6788-93.