



Genome sequencing and analysis of the first complete genome of *Lactobacillus kunkeei* strain MP2, an *Apis mellifera* gut isolate

Freddy Asenjo¹, Alejandro Olmos², Patricia Henríquez-Piskulich³, Victor Polanco^{2,3}, Patricia Aldea³, Juan A. Ugalde¹ and Annette N. Trombert²

¹ Centro de Genética y Genómica, Facultad de Medicina, Clínica Alemana Universidad del Desarrollo, Santiago, Chile

² Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, Chile

³ Centro de Estudios Apícolas CEAPI Mayor, Facultad de Ciencias, Universidad Mayor, Santiago, Chile

ABSTRACT

Background. The honey bee (*Apis mellifera*) is the most important pollinator in agriculture worldwide. However, the number of honey bees has fallen significantly since 2006, becoming a huge ecological problem nowadays. The principal cause is CCD, or Colony Collapse Disorder, characterized by the seemingly spontaneous abandonment of hives by their workers. One of the characteristics of CCD in honey bees is the alteration of the bacterial communities in their gastrointestinal tract, mainly due to the decrease of *Firmicutes* populations, such as the *Lactobacilli*. At this time, the causes of these alterations remain unknown. We recently isolated a strain of *Lactobacillus kunkeei* (*L. kunkeei* strain MP2) from the gut of Chilean honey bees. *L. kunkeei*, is one of the most commonly isolated bacterium from the honey bee gut and is highly versatile in different ecological niches. In this study, we aimed to elucidate in detail, the *L. kunkeei* genetic background and perform a comparative genome analysis with other *Lactobacillus* species.

Methods. *L. kunkeei* MP2 was originally isolated from the guts of Chilean *A. mellifera* individuals. Genome sequencing was done using Pacific Biosciences single-molecule real-time sequencing technology. *De novo* assembly was performed using Celera assembler. The genome was annotated using Prokka, and functional information was added using the EggNOG 3.1 database. In addition, genomic islands were predicted using IslandViewer, and pro-phage sequences using PHAST. Comparisons between *L. kunkeei* MP2 with other *L. kunkeei*, and *Lactobacillus* strains were done using Roary.

Results. The complete genome of *L. kunkeei* MP2 comprises one circular chromosome of 1,614,522 nt. with a GC content of 36,9%. Pangenome analysis with 16 *L. kunkeei* strains, identified 113 unique genes, most of them related to phage insertions. A large and unique region of *L. kunkeei* MP2 genome contains several genes that encode for phage structural protein and replication components. Comparative analysis of MP2 with other *Lactobacillus* species, identified several unique genes of *L. kunkeei* MP2 related with metabolism, biofilm generation, survival under stress conditions, and mobile genetic elements (MGEs).

Discussion. The presence of multiple mobile genetic elements, including phage sequences, suggest a high degree of genetic variability in *L. kunkeei*. Its versatility and ability to survive in different ecological niches (bee guts, flowers, fruits among others)

Submitted 19 October 2015

Accepted 29 March 2016

Published 19 April 2016

Corresponding authors

Juan A. Ugalde, jugalde@udd.cl

Annette N. Trombert,

annette.trombert@umayor.cl

Academic editor

Blanca Landa

Additional Information and
Declarations can be found on
page 14

DOI 10.7717/peerj.1950

© Copyright

2016 Asenjo et al.

Distributed under

Creative Commons CC-BY 4.0

OPEN ACCESS

could be given by its genetic capacity to change and adapt to different environments. *L. kunkeei* could be a new source of *Lactobacillus* with beneficial properties. Indeed, *L. kunkeei* MP2 could play an important role in honey bee nutrition through the synthesis of components as isoprenoids.

Subjects Agricultural Science, Entomology, Food Science and Technology, Microbiology

Keywords *Lactobacillus*, Comparative genomics, *Apis mellifera*, Honeybee

INTRODUCTION

The honey bee (*Apis mellifera*) is the most important pollinator in agriculture worldwide, playing a key role in the human food supply by providing pollination services for diverse crops (Evans & Schwarz, 2011). However, from 2006 to this day, an unusual decrease in honey bee colonies has been taking place, known as Colony Collapse Disorder (CCD). CCD describes the seemingly spontaneous abandonment of the hives by honey bee workers, where queens often stay in the hive accompanied by a small group of nurse worker bees. The specific causes of CCD are unknown, but several factors can impact the health of honey bees, and contribute to this phenomenon: (1) pests and diseases (such as, American foulbrood, European foulbrood, chalkbrood nosema, small hive beetles, and tracheal mites); (2) the use of chemicals in bee colonies, and their surrounding environment; (3) beekeeping practices; (4) agricultural practices and (5) climate change (Henry et al., 2012; Di Pasquale et al., 2013; Di Prisco et al., 2013).

Multiple studies have suggested that CCD directly affects the microbial composition of the honey bee gut microbiota. Eight dominant groups can be found in the honey bee gut (Cox-Foster et al., 2007; Martinson et al., 2011): *Gammaproteobacteria* (*Enterobacteriaceae* and *Pasteurellaceae*), *Betaproteobacteria* (*Neisseriaceae*), *Alphaproteobacteria* (*Rhizobiales*, *Acetobacteraceae*), *Firmicutes* (*Lactobacillus* sp.), and *Actinobacteria* (*Bifidobacterium* sp.) groups (Cox-Foster et al., 2007; Martinson et al., 2011). Gut microbiome studies from individuals obtained from colonies affected and non-affected by CCD, indicated an increase in the *Gammaproteobacteria*, and a decrease of the *Firmicutes* in affected colonies, showing how the CCD condition affects commensal communities in the honey bee gut (Cox-Foster et al., 2007). *Firmicutes* includes Gram-positive and low-G + C bacteria, such as the *Lactobacillus* genus, where some of its members have been implicated in the fitness improvement of honey bees (Audisio & Benítez-Ahrendts, 2011; Audisio, Sabaté & Benítez-Ahrendts, 2015).

The study of lactobacilli members of the honey bee microbiota can give us information about beneficial species for honey bees. One of the most common lactobacilli species present in the honey bee gut microbiota is *Lactobacillus kunkeei*, described for the first time as a spoilage organism isolated from commercial grape wine (Edwards et al., 1998). Characterized as a Gram-positive fructophilic lactic acid bacterium (FLAB), *L. kunkeei* possesses a weak catalase activity and has the ability to ferment carbohydrates, such as glucose, fructose, sucrose, raffinose and mannitol but prefer fructose. The fermentation

products of these reactions are lactic acid and acetic acid ([Edwards et al., 1998](#); [Bae, Fleet & Heard, 2006](#); [Endo, Futagawa-Endo & Dicks, 2009](#); [Endo, 2012](#)). *L. kunkeei* can be found in fructose rich-niches, including honey, beebread, wine and flowers ([Vásquez et al., 2012](#); [Endo et al., 2012](#)). Furthermore it is present in the gastrointestinal tract of several insects found in flowers, such as tropical fruit flies, *Camponotus spp* (carpenter ants), bumblebees and honey bees ([Neveling, Endo & Dicks, 2012](#); [Anderson et al., 2013](#); [Endo & Salminen, 2013](#)). During the summer months, *L. kunkeei* is the most frequent lactobacilli isolate from the honey bee gut ([Corby-Harris, Maes & Anderson, 2014](#); [McFrederick et al., 2014](#)).

In previous work, we isolated a *L. kunkeei* strain (named as MP2) from the gut of Chilean honey bees from María Pinto, Melipilla ([Olmos et al., 2014](#)). The genome of this strain was sequenced using the Illumina MiSeq platform, which resulted in a draft genome of 44 contigs, for a total genome size of 1,581,395 bp, and 826 well-annotated protein coding-genes ([Olmos et al., 2014](#)). The nature of the short-reads used for this assembly did not allow for complete resolution of the genome without gaps. In addition, multiple repetitions, including the presence of multiple copies of the ribosomal operon, could not be resolved in this draft genome. To overcome these limitations, we performed a re-sequencing of the *L. kunkeei* MP2 genome, using single molecule sequencing in the Pacific Biosciences platform.

In this work, we report the first complete genome sequence of *L. kunkeei* MP2, its characterization, and comparison with other *Lactobacillus* genomes.

METHODS

DNA isolation and genome sequencing

The *L. kunkeei* strain MP2 reported in this study was isolated in a previous study from the gut of a honey bee (*Apis mellifera*), collected from commune hives in the Maria Pinto area, Melipilla Province in the Central zone of Chile ([Olmos et al., 2014](#)). Collection and use of honey bees for these studies was reviewed and approved by the Bioethics Committee of Universidad Mayor, which is governed by the regulations of the Animal Health Services of Chile.

For DNA extraction, *L. kunkeei* MP2 colonies were cultured in MRS broth (37 °C, 5% CO₂) and genomic DNA was obtained using a silica-based protocol ([Boom et al., 1999](#)). Briefly, bacterial pellet was lysed using a solution composed of SDS 10%, proteinase K (10 mg/mL, Thermo Scientific) and lysozyme (5 mg/mL, Pierce) at 37 °C for 60 min. The lysate was mixed with guanidine chloride 6M and a silica suspension (50% w/v) and incubated for 10 min. The silica was centrifuged, and DNA was recovered, after washes with 70% alcohol, into sterile, free nuclease water. Approximately 13.5 µg of DNA were used to construct sequencing libraries with an average insert size of 20 kb, and sequenced using one SMRT cell (P6-C4 Chemistry) on a PacBio RSII sequencer (Pacific Biosciences) at the UCSD IGM Genomics Center.

Genome assembly and annotation

Raw reads (~1 Gbps) were processed to remove SMRT bell adapters, short and low-quality reads (<80% accuracy) using SMRT Analysis version 2.3. A total of 154,044 filtered reads

(average length, 9 Kb) were used for *de novo* assembly using Celera Assembler version 8.3 (Myers *et al.*, 2000), with self-correction of the PacBio reads (Berlin *et al.*, 2015). Polishing was done using Quiver, using SMRT Analysis version 2.3. Comparisons between the previously sequenced *L. kunkeei* MP2 genome (Accession number PRJNA257367) (Olmos *et al.*, 2014), as well with the other two available genome sequences (Porcellato *et al.*, 2015; Djukic *et al.*, 2015) were performed using MUMMER (Kurtz *et al.*, 2004). Genome annotation was performed using Prokka version 1.11 (Seemann, 2014). The predicted CDS were classified into EggNOG categories using HMMER version 3.1 (<http://hmmer.org>) against the EggNOG 4.1 database (Powell *et al.*, 2014) with an *E*-value cutoff of 1E-05. Genomic islands were annotated using IslandViewer 3 (Dhillon *et al.*, 2015), and possible phage sequences were searched using PHAST (Zhou *et al.*, 2011). Genome visualization was done using Circos version 0.69 (Krzywinski *et al.*, 2009). The genome sequence an assembly is available at NCBI with the accession number PRJNA298292.

Pan-genome analysis

Comparative genomic analysis was performed from a list of selected genomes from *Lactobacillus* species, as well as other strains of *L. kunkeei* (Table 1). To avoid possible biases in the comparisons due to different annotation procedures, all of the genomes were re-annotated using Prokka version 1.11 (Seemann, 2014). Comparisons between all the selected genomes, were done using Roary (Page *et al.*, 2015), with a blast identity cutoff of 97% for the comparison between *L. kunkeei* strains, and a 40% for the comparison between *Lactobacillus* species. In addition at the genus level, paralog splitting was disabled. Visualization of the pan-genome data was done using Anvi'o (Eren *et al.*, 2015).

Phylogenetic analysis

16S rRNA gene sequences were obtained from the Silva database (Quast *et al.*, 2013), and aligned using the SINA webserver aligner (Pruesse, Peplies & Glöckner, 2012) with default parameters, and using the *Bacteria* variability profile. For the *L. kunkeei* strain tree, we used the core genome (genes shared by all *L. kunkeei* strains), aligning the genes using Prank (Löytynoja, 2014). For both cases, the phylogenetic trees were generated using FastTree v 2.1.7 (Price, Dehal & Arkin, 2010), with the *slow* option.

RESULTS AND DISCUSSION

Assembly description

The PacBio reads obtained for *L. kunkeei* MP2 were assembled using MHAP (Berlin *et al.*, 2015) implemented in the PBcR pipeline (Celera Assembler 8.3) (Myers *et al.*, 2000). This *de novo* assembly resulted in one contig, representing the complete genome of *L. kunkeei* MP2 in a single 1,614,522 nt chromosome. A total of 1,468 CDS were predicted in the genome, 67 tRNA and 5 copies of the ribosomal operon. Functional annotation was done using EggNOG V 4.1 (Powell *et al.*, 2014), and the summary of functional categories is shown on Table 2 (gene annotation on Table S1). The %GC content of the genome was 36.9%, and several features of interests, such as the presence of prophage regions, were found. With this assembly, we were able to differentiate the three ribosomal operons that

Table 1 Genome sequences used in this study. Species name and accession numbers of the genomes selected in this study.

Genomes	Bioproject	Assembly
<i>Lactobacillus kunkeei</i> DSM 12361	PRJNA222257	GCA_001433825.1
<i>Lactobacillus kunkeei</i> Fhon2	PRJNA270967	GCA_001281165.1
<i>Lactobacillus kunkeei</i> LAan	PRJNA270961	GCA_001281225.1
<i>Lactobacillus kunkeei</i> LAce	PRJNA270962	GCA_001421115.1
<i>Lactobacillus kunkeei</i> LAfl	PRJNA270964	GCA_001421135.1
<i>Lactobacillus kunkeei</i> LAko	PRJNA270965	GCA_001281205.1
<i>Lactobacillus kunkeei</i> LAla	PRJNA270966	GCA_001281215.1
<i>Lactobacillus kunkeei</i> LAni	PRJNA270969	GCA_001281285.1
<i>Lactobacillus kunkeei</i> LMbe	PRJNA270972	GCA_001308185.1
<i>Lactobacillus kunkeei</i> LMbo	PRJNA270973	GCA_001308195.1
<i>Lactobacillus kunkeei</i> LAdo	PRJNA270963	GCA_001308205.1
<i>Lactobacillus kunkeei</i> LAnu	PRJNA270970	GCA_001308215.1
<i>Lactobacillus kunkeei</i> EFB6	PRJNA227106	GCA_000687335.1
<i>Lactobacillus kunkeei</i> AR114	PRJNA253911	GCA_000830375.1
<i>Lactobacillus kunkeei</i> YH-15	PRJNA270974	GCA_001281265.1
<i>Lactobacillus kunkeei</i> MP2	PRJNA298292	GCA_001314945.1
<i>Lactobacillus acidophilus</i> 30SC	PRJNA63605	GCA_000191545.1
<i>Lactobacillus amylovorus</i> GRL1118	PRJNA160233	GCA_000194115.1
<i>Lactobacillus brevis</i> KB290	PRJNA195560	GCA_000359625.1
<i>Lactobacillus buchneri</i> NRRL B-30929	PRJNA66205	GCA_000211375.1
<i>Lactobacillus casei</i> BD-II	PRJNA162119	GCA_000194765.1
<i>Lactobacillus delbrueckii</i> subsp. bulgaricus ND02	PRJNA60621	GCA_000182835.1
<i>Lactobacillus fermentum</i> CECT 5716	PRJNA162003	GCA_000210515.1
<i>Lactobacillus gasseri</i> 130918	PRJNA224116	GCA_000814885.1
<i>Lactobacillus helveticus</i> H10	PRJNA162017	GCA_000189515.1
<i>Lactobacillus hokkaidonensis</i> JCM 18461	PRJNA224116	GCA_000829395.1
<i>Lactobacillus johnsonii</i> DPC 6026	PRJNA162057	GCA_000204985.1
<i>Lactobacillus kefranofaciens</i> ZW3	PRJNA67985	GCA_000214785.1
<i>Lactobacillus mucosae</i> LM1	PRJNA86029	GCA_000248095.3
<i>Lactobacillus paracasei</i> subsp. paracasei 8700:2	PRJNA55295	GCA_000155515.2
<i>Lactobacillus plantarum</i>	PRJNA224116	GCA_000931425.1
<i>Lactobacillus reuteri</i> SD2112	PRJNA55357	GCA_000159455.2
<i>Lactobacillus rhamnosus</i> LOCK908	PRJNA210958	GCA_000418495.1
<i>Lactobacillus ruminis</i> ATCC 27782	PRJNA73417	GCA_000224985.1
<i>Lactobacillus sakei</i> subsp. sakei 23K	PRJNA58281	GCA_000026065.1
<i>Lactobacillus salivarius</i>	PRJNA224116	GCA_000758365.1
<i>Lactobacillus sanfranciscensis</i> TMW 1.1304	PRJNA72937	GCA_000225325.1
<i>Lactobacillus</i> sp. wkB8	PRJNA224116	GCA_000761135.1

are present in the chromosome, something that was not possible in the previous sequenced genome of this strain (*Olmos et al., 2014*).

A comparison of the assembly of *L. kunkeei* MP2 obtained in this work, with the previously obtained using Illumina sequencing (*Olmos et al., 2014*), is shown in *Fig. 1*. All

Table 2 EggNOG functional categories for the predicted genes of *L. kunkeei* MP2.

Information Storage and Processing	
Translation, ribosomal structure and biogenesis	127
Transcription	73
Replication, recombination and repair	129
Cellular Processes and Signaling	
Cell cycle control, cell division, chromosome partitioning	22
Defense mechanisms	15
Signal transduction mechanisms	23
Cell wall/membrane/envelope biogenesis	77
Cell motility	4
Intracellular trafficking, secretion, and vesicular transport	19
Posttranslational modification, protein turnover, chaperones	45
Metabolism	
Energy production and conversion	41
Carbohydrate transport and metabolism	58
Amino acid transport and metabolism	107
Nucleotide transport and metabolism	69
Coenzyme transport and metabolism	25
Lipid transport and metabolism	32
Inorganic ion transport and metabolism	65
Secondary metabolites biosynthesis, transport and catabolism	8
Poorly Characterized	
Function unknown	414

of the previous assembled contigs mapped to the current assembly, and several gaps on the sequence were completed in this new version of the genome.

Central Metabolism of *L. kunkeei* MP2

Energy metabolism

MP2 has the complete route for acetate synthesis, with the presence of the gene codifying for phosphoglycerate kinase. No genes codifying for phosphoribulokinase (PRK) and ribulose-biphosphate carboxylase (RbcL), two of the enzymes involved in the synthesis of glyceraldehyde-3-phosphate synthesis, were found on the genome.

Carbohydrate metabolism

The genes that encode for the enzymes phosphofructokinase/glucokinase (PFK) and Fructose-biphosphate aldolase (FBA), were not found in the genome of *L. kunkeei* MP2. These enzymes are part of the Embden-Meyerhof pathway and are involved in the homofermentative metabolism of lactic acid. As a fructophilic bacterium, *L. kunkeei* MP2 can synthesize ribose-5-phosphate through pentose phosphate pathway from fructose and obtain PRPP (phosphoribose pyrophosphate), the precursor of purine, pyrimidine and histidine metabolism. For the synthesis of ribose-5-phosphate, *L. kunkeei* uses the route from B-D-fructose-6-phosphate through D-arabino-Hex-3-ulose-6-phosphate intermediate. *L. kunkeei* MP2 can synthesize UDP-glucose and has two isoprenoid biosynthesis pathways, the mevalonate and the non-mevalonate pathways. Isoprenoids

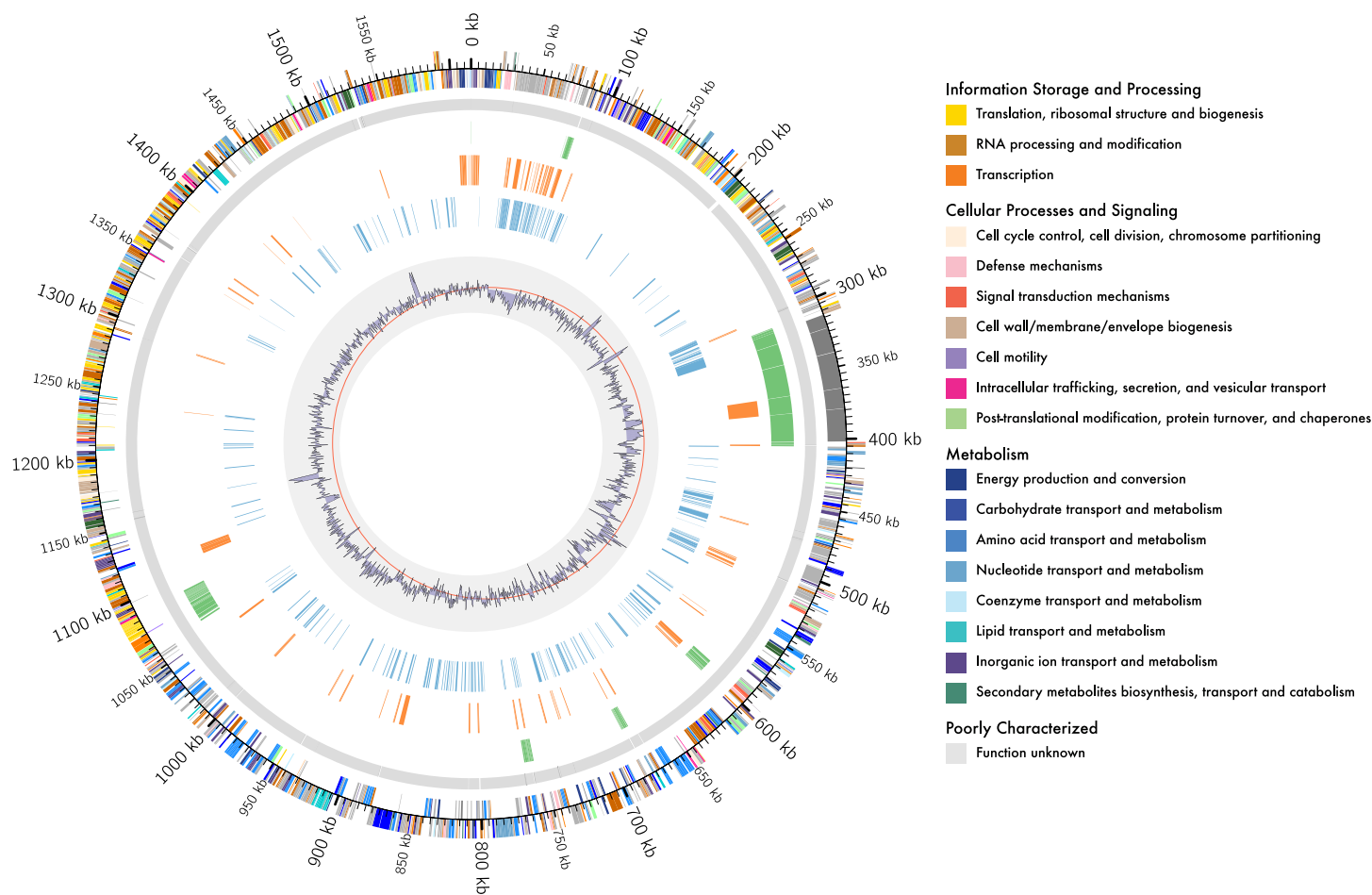


Figure 1 Genome organisation of *L. kunkei* MP2. Circular overview of the complete genome of *L. kunkei* MP2, highlighting some of the features. Starting from the outside ring towards the interior: EggNOG annotation of the predicted CDS; Contig recruitment of the previous *L. kunkei* MP2 genome sequencing (Olmos et al., 2014); Phage island predictions using Island Viewer 3; Unique genes of *L. kunkei* MP2, compared to 16 strains of *L. kunkei*; Unique genes of *L. kunkei* MP2 compared with 22 genomes of *Lactobacillus* species; %GC content of the *L. kunkei* MP2 genome.

include carotenoids, sterols, prenyl side-chains of chlorophylls, and plastoquinone, exhibiting many biological functions (Daum et al., 2009). In whiteflies (*Bemisia tabaci*), the genome of its endosymbiotic bacteria, *Candidatus Portiera aleyrodidarum*, encodes for key enzymes in carotenoids synthesis, suggesting that whitefly not only can acquire carotenoids from the diet, but also from their microbiota (Sloan & Moran, 2012). Therefore, if *L. kunkei* produces key enzymes involved in isoprenoid synthesis, it is possible that it could be playing an important role in honey bee nutrition.

Nucleotide and amino acid metabolism

The pathways for purine biosynthesis are complete. However, in the *de novo* pyrimidine pathway, *L. kunkei* lacks the gene *pyrB*, which codifies for the aspartate carbamoyl transferase, and *ndk*, codifying for the nucleoside diphosphate kinase. The analysis of the metabolic pathways in MP2 revealed a minimal amino acid auxotrophy (methionine

or cysteine), with the presence of the genes that encode for a D-methionine transport system, suggesting the ability of *L. kunkeei* MP2 to acquire methionine/cysteine from the environment. These results are in line with previous reports of the lactobacilli being auxotrophic for both methionine and cysteine (Seefeldt & Weimer, 2000), and where the supplement of culture media with these amino acids improved bacterial growth (Lozo et al., 2008). A gene that encodes for serine hydroxymethyltransferase (SHMT) was found in the genome of MP2. This enzyme catalyzes the addition of formaldehyde to glycine, a key step for the production of serine (Jiang et al., 2014), and appears to be absent in the other *Lactobacillus* genomes analyzed in this study. Its presence in *L. kunkeei* MP2 could be part of specific adaptation mechanisms of this species to its environment.

Prophage insertions

Previous work in other *Lactobacillus* species, reported the presence of regions with prophage genes in their genomes, including species such as *L. rhamnosus*, *L. gasseri*, *L. salivarius*, *L. casei*, *L. lactis*, and *L. johnsonii* (Ventura et al., 2004; Ventura et al., 2006; Kankainen et al., 2009; Savabi et al., 2014; Baugher, Durmaz & Klaenhammer, 2014). This shows the widespread abundance of prophages in the genomes of *Lactobacillus* species, a characteristic shared by *L. kunkeei* MP2. Two regions were identified by PHAST (Zhou et al., 2011), as putative prophage insertions (Table S2). One of them, located in the region between 594,506 and 613,136, was found to be present in all the 23 *Lactobacillus* genomes used in this work. The second region, located around 32,973–75,092, was found to be unique to *L. kunkeei* MP2, compared to other strains of *L. kunkeei*, as well as other *Lactobacillus* species. In at least one *Lactobacillus* species (*L. gasseri*), the presence of these inserted phages has been associated with the horizontal transfer of genes (Baugher, Durmaz & Klaenhammer, 2014), suggesting a possible role for these elements within the genome of *L. kunkeei* MP2. However, the detailed mechanisms, as well as the possible adaptive consequences of such events, need to be explored in more detail in the future.

Comparison of *L. kunkeei* MP2 with other *L. kunkeei* strains

We performed a comparative genomic analysis of MP2 against sixteen publicly available genome sequences of *L. kunkeei* strains (Table 1). This analysis can provide us with a snapshot of the unique features that are present in this strain, based on its gene content. Recent work, highlighted the important role that genome reduction played in the evolution of *L. kunkeei* (Tamarit et al., 2015), which suggests that we would expect to find only a few genes truly unique to *L. kunkeei* MP2, compared to other strains.

Pangenome analysis resulted in the identification of a set of 853 core genes (present in all the strains). For the shell genes, we found that 813 genes are found between 15% to 95% of the analyzed strains, while 1661 genes are present in less than 15% of the strains. Focusing on the MP2 strain, based on this analysis we identified 113 genes that are not present in any of the other strains, representing close to 7.7% of the total number of genes.

Using the pangenome matrix, we can visualize the relationships between the genomes based only on the clustering of the proteins (presence or absence of a gene in a protein group). Visualizations of the results (Fig. 2), show that the strains LAni, LAce, LAan and

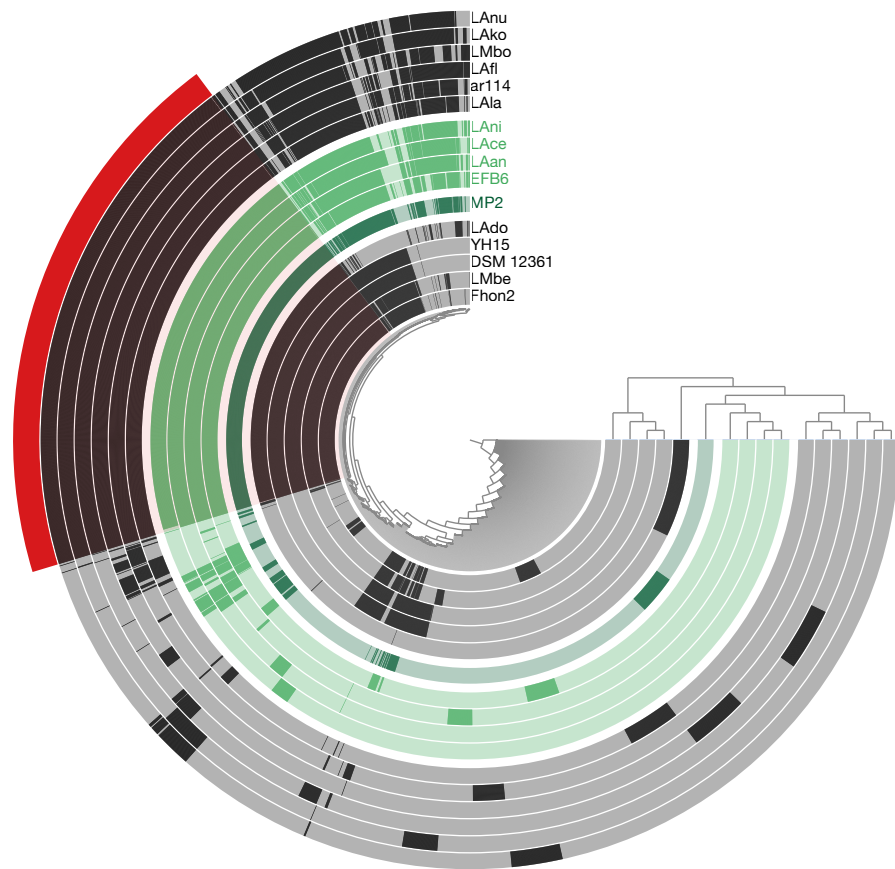


Figure 2 Anvi'o pangene visualization of 16 *L. kunkeei* genomes. The outer core in red, shows the core genome of *L. kunkeei*, protein clusters shared among all the strains (853 genes). The genomes are clustered based on the presence/absence pattern of protein clusters. MP2 is highlighted in green, while the more similar strains based on the clustering pattern, are highlighted in lighter green (LAAni, LAce, LAAn, and EFB6). Pan-genome visualization was generated using Anvi'o (Eren *et al.*, 2015).

EFB6, cluster in the same group as MP2, based on their gene content. This correlates with the phylogenetic placement of MP2 within the same group, based on a concatenated alignment of the core genome (853 genes; 807,585 nucleotides) (Fig. 3). The closest strain on the tree is EFB6, which was also isolated from the gut of *A. mellifera* (although from a larvae affected by European foulbrood) (Djukic *et al.*, 2015). Currently there is no evidence of association between particular strains and a particular species of *Apis* (Tamarit *et al.*, 2015). A better understanding of the interactions between the gut microbiota and the host needs to incorporate additional variables, such as the surrounding environment (including flowers and fruit), seasonal variations, among other possible elements. These factors could play a strong role in the association between an *Apis* species and a *L. kunkeei* strain.

One of the main differences of MP2 compared to the other *L. kunkeei* genomes, is the presence of multiple phage genes inserted in several parts of the genome. One of these unique phage regions can be found at coordinates 31,034–75,092 (Fig. 1). It is a large region, which includes several phage-related proteins, including structural and replications components. Sequence analysis using Blast, shows that these proteins are related to phages that

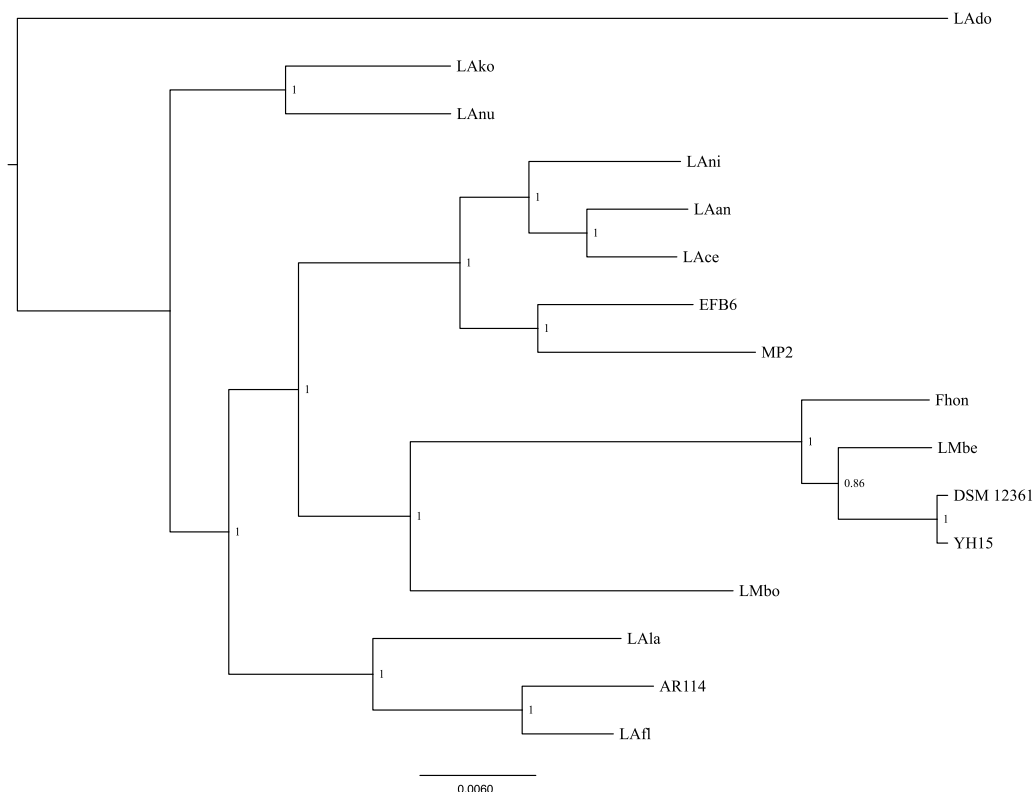


Figure 3 Phylogenetic tree of 16 *L. kunkeei* strains. The tree was constructed using all of the genes shared between all 16 strains (853 genes; 807,585 nucleotides).

infect Gram-positive Bacteria, such as *Bacillus* (Hastings *et al.*, 2013), *Listeria* (Dorscht *et al.*, 2009), *Enterococcus* (Yasmin *et al.*, 2010), and *Staphylococcus* (Chang *et al.*, 2013) (Table S3).

Comparison of *L. kunkeei* MP2 with other *Lactobacillus* strains

Phylogenetic reconstructions using the sequence that encodes for the 16S rRNA gene, shows that the closest species to *L. kunkeei* MP2 is *L. sanfranciscensis* (Fig. 4). This placement is in agreement with previous diversity analysis performed on acidophilic bacteria (McFrederick *et al.*, 2012), relating the taxonomy of both species.

Whole genome comparisons between *L. kunkeei* MP2 and other *Lactobacillus* species, discovered several unique genes. One example is *gtfC*, which encodes for a glucosyltransferase, which has been extensively studied in *Streptococcus mutans*, where is expressed in the presence of carbohydrates such as sucrose, D-glucose, D-fructose, among others (Shemesh *et al.*, 2006). GtfC (as well as GftB), is also considered a virulence factor in *S. mutans*, promoting bacterial adhesion to smooth surfaces and cells (Tsumori & Kuramitsu, 1997). Also, GtfC is part of the synthesis route of a mixture of insoluble and soluble glucans, which are important components of cariogenic biofilms (Yousefi *et al.*, 2012). Considering the rich carbohydrate environment where *L. kunkeei* can survive, the presence of unique glucosyltransferase genes, such as *gtfC*, could facilitate bacterial colonization of flowers, as well as the honey bee gut.

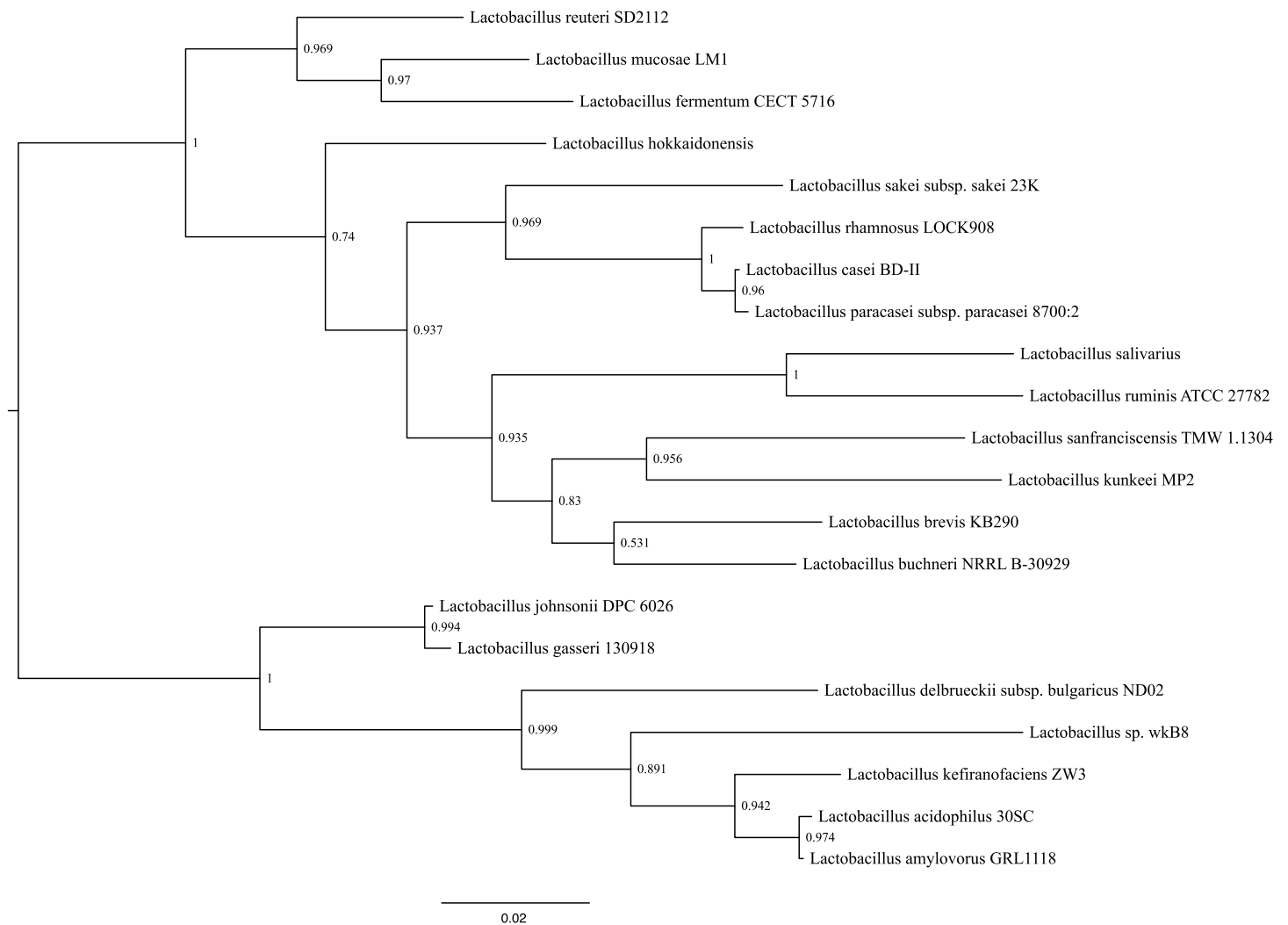


Figure 4 Phylogenetic tree of several *Lactobacillus* species, including *L. kunkeei* MP2. Phylogenetic reconstruction was done using the sequence of the 16S rRNA gene.

Another unique gene found in *L. kunkeei*, encodes for the adapter protein MecA, a pleiotropic regulator of bacterial development. This protein has been shown to affect competence, protein degradation and sporulation in *Bacteria*, such as *Bacillus subtilis* (Schlothauer et al., 2003). MecA interacts with the chaperone ClpC, and with the transcription factor ComK, promoting the degradation of this protein during the logarithmic growth phase. The degradation of ComK stops when bacteria enters to stationary growth phase, where the quorum-sensing pheromone ComX promotes the synthesis of ComS, which binds to MecA and prevents the interaction of MecA-ComK (Persuh, Mandic-Mulec & Dubnau, 2002; Prepiak et al., 2011; Wahl et al., 2014). This could have an effect on the biofilm generation capabilities of *L. kunkeei* MP2, but this needs to be explored experimentally.

At least seven different *Lactobacillus* species have been characterized in the gut microbiota of *A. mellifera*, where it has been suggested that they play different roles in the stability

of the host functions (Engel & Moran, 2013). *L. kunkeei* MP2 appears to have a unique set of genes when compared to other strains of *L. kunkeei*, as well as with other species of *Lactobacillus* (Table S3), which suggest unique adaptation strategies of *L. kunkeei* MP2 to the gut of *A. mellifera*.

We also identified a hypothetical protein with similarities to a low-molecular-weight protein-tyrosine phosphatase (LMPTP), unique to the *L. kunkeei* MP2 genome, compared to other *L. kunkeei* strains and other *Lactobacillus* species. This LMPTP is similar to the YfkJ protein from *Bacillus subtilis*, which has been involved in the response to ethanol stress (Musumeci et al., 2005). Ethanol, and other organic compounds, are commonly present in the environment, and accumulate in the bacterial membrane affecting its physical-chemical properties, and in consequence, their functions (Weber & De Bont, 1996). This could suggest a better tolerance to organic compounds, such as ethanol, for *L. kunkeei* MP2, which could help this organism to tolerate unfavorable conditions, and have a unique competitive advantage compared to other *Lactobacillus* species (De Guchte van et al., 2002).

The diversity of *Firmicutes* species in *A. mellifera* could imply a metabolic diversity that could be crucial for honey bee fitness (Engel & Moran, 2013). Comparative genomics of *Lactobacillus* genomes, have shown that close to 45% of its accessory genome encode for proteins involved in carbohydrate metabolism and transport functions (Ellegaard et al., 2015). With this metabolic diversity found in the accessory genome, is no surprising to find unique genes in the accessory genome of *L. kunkeei* MP2, when compared to other strains of *L. kunkeei*, as well as other *Lactobacillus* species (Table S3). These genes encode for proteins that take part of the degradation of carbohydrates, transport of molecules, transcription, as well as membrane proteins. It is very likely that some of these genes were acquired via horizontal gene transfer from a diverse group of organisms, including those that inhabit the gut of *A. mellifera*. This has been observed in the adaptation of strains of *Gilliamella apicola* and *Snodgrassella alvi* to the guts of the honey bee and the bumble bee (Kwong et al., 2014), as well as in other mammalian guts (Shterzer & Mizrahi, 2015).

Integrative and conjugative elements in MP2

Multiple mobile genetic elements (MGEs), were identified in the genome of *L. kunkeei* MP2, including prophages, transposons, and integrases. Several of these genes were unique to the MP2 genome, compared to the other draft genomes of *L. kunkeei* and other *Lactobacillus* strains. To explore a possible association between MGEs and the unique genes found in the genome of *L. kunkeei* MP2, we performed a prediction of genomic islands using Island Viewer 3 (Dhillon et al., 2015). With this approach, we found that most of the unique genes are found outside genomic islands (Fig. 1, Table S3). This could suggest either events of gene loss or ancestral transfer events in the genome of *L. kunkeei* MP2 (Tamarit et al., 2015).

Most of the MGEs found in the genome, had similarities to integrative and conjugative elements (ICEs), which are characterized by their prophage-like mode of maintenance (Burrus et al., 2002). To contrast this result, the uniquely identified genes in the genome of *L. kunkeei* MP2 were compared against the ICEberg database (Bi et al., 2012) (Table S3). ICEs commonly encode for genes that provide an increased fitness to the host, such as antibiotic resistance genes, phage resistance, and heavy metal transport (Burrus et

al., 2002). In the case of MP2 we found genes that have similarities to transmembrane proteins, phage-related proteins, and antibiotic resistance mechanisms, suggesting that the incorporation and stability of these unique genes in the genome of *L. kunkeei* MP2, is providing an increase in the fitness of this bacterial strain in the gut of *A. mellifera*. Among the predicted phage-like sequences, we found one coding for a *mef(A)/msr(D)* resistance protein, with similarity to a sequence from *Streptococcus pyogenes*, involved in the resistance to macrolides (Iannelli *et al.*, 2014). In the European Union, the usage of antibiotics, and antibiotic-containing compounds, is not permitted. However, macrolides (such as tylosin and streptomycin), are still used as a preventive treatment against *Paenibacillus larvae*, the causal agent of American foulbrood, in many countries (Reynaldi *et al.*, 2010; Gaudin, Hedou & Verdon, 2012). Thus, if bees were exposed to antibiotics in their diet, it is possible that the gut microbiota may have acquired the necessary molecular mechanisms to adapt and survive in an exposed environment (Tarapoulouzi *et al.*, 2013). This acquisition can be explained by horizontal gene transfer events from the surrounding natural environment (such as soil). Here we can find multiple bacterial and fungal species that commonly produce antimicrobial compounds, and could act as a source of these resistance genes (Alippi, León & López, 2014).

Most of ICEs coding genes are usually present within genomic islands in the host genome (Hacker & Carniel, 2001; Boyd, Almagro-Moreno & Parent, 2009), but in the case of *L. kunkeei* MP2, none of the predicted ICEs genes were found in the context of genomic islands according to the predictions performed with IslandViewer 3 (Dhillon *et al.*, 2015). This could suggest the presence of previously uncharacterized ICEs, or also our current limitation in the detection of ICEs from *Lactobacillus* species.

Prediction of horizontal gene transfer events

To predict horizontally transferred genes we used Darkhorse (Podell & Gaasterland, 2007) to analyze the complete genome of *L. kunkeei* MP2. We did not consider hits to organisms within the same Phylum, to avoid false predictions, although this could lead to ignore real transfer events between more closely related organisms. A total of 19 genes were predicted to have been acquired via horizontal gene transfer (Table S3), with a normalized LPI score cutoff of 0.546. Seven of these genes had matches with the genome of *A. mellifera*, which a detailed look suggested as a contamination of the genome of *A. mellifera* with sequences from *Lactobacillus* species. This has been previously reported for other genome projects (Merchant, Wood & Salzberg, 2014). Only one of the genes predicted to be acquired via HGT was unique to *L. kunkeei* MP2 when compared to other *L. kunkeei* strains and other *Lactobacillus* genomes, which codifies for a hypothetical protein, with a best hit as a phage protein from *Halomonas* sp. HAL1. None of the predicted genes was found associated with an ICE or a genomic island. Although the apparent lack of genes of acquired via HGT could be explained by the genome reduction that has been observed in *L. kunkeei* strains (Tamarit *et al.*, 2015), an alternative explanation is the lack of reference genomes of isolates obtained from the gut of *A. mellifera*. This could limit our ability to predict events of horizontal gene transfer between members of the honey gut microbiota.

CONCLUSIONS

Using long-read sequence technologies we were able to assemble the first closed genome for a *L. kunkeei* strain, resolving some of the assemble issues of a previous draft of this strain (Olmos *et al.*, 2014). The comparison of the genome sequence against other *Lactobacillus* species, showed a percentage of genes that are unique to the MP2 strain, including metabolic key enzymes that could play an important role in the honey bee nutrition and fitness. The genome of *L. kunkeei* MP2 also has genes encoding for proteins involved in important roles such as adhesion, biofilm synthesis, and stress tolerance, which in addition to the presence of antibiotic resistance related genes, highlights the versatility of this bacteria to adapt to different environments, such as flowers or insect guts.

One of the features highlighted in this study is the abundance of prophages in the *L. kunkeei* genome. The presence of prophages in *Lactobacillus* is common, but MP2 has sequences unique to this strain. This is the case of a large genomic region (located in the 31,034-75,092 region), with genes encoding for several phage-related proteins, including structural and replicative components. The presence of prophages could be associated with lateral transference events, and therefore, with the acquisition of genes related with bacterial fitness. Given the high percentage of hypothetical proteins encoded in this region, a future goal for research, is the elucidation of the role for these proteins in *L. kunkeei* MP2.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by the project Bees for Care (B4C 13CTI-21546) and a Conicyt grant (Fondecyt Iniciación No 11140666). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:

Bees for Care: B4C 13CTI-21546.

Conicyt Grant: Fondecyt Iniciación No 11140666.

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Freddy Asenjo, Juan A. Ugalde and Annette N. Trombert conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Alejandro Olmos conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools.
- Patricia Henríquez-Piskulich conceived and designed the experiments, performed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.

- Victor Polanco and Patricia Aldea conceived and designed the experiments, contributed reagents/materials/analysis tools, reviewed drafts of the paper.

DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

GenBank accession number [CP012920](#).

Data Availability

The following information was supplied regarding data availability:

Figshare: http://figshare.com/articles/Lactobacillus_kunkeei_MP2_Genome/1576406.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.1950#supplemental-information>.

REFERENCES

- Alippi AM, León IE, López AC. 2014.** Tetracycline-resistance encoding plasmids from *Paenibacillus larvae*, the causal agent of American foulbrood disease, isolated from commercial honeys. *International Microbiology* 17:49–61.
- Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, Schwan MR, Walton A, Jones BM, Corby-Harris V. 2013.** Microbial ecology of the hive and pollination landscape: bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). *PLoS ONE* 8:e83125 DOI [10.1371/journal.pone.0083125](https://doi.org/10.1371/journal.pone.0083125).
- Audisio MC, Benítez-Ahrendts MR. 2011.** *Lactobacillus johnsonii* CRL1647, isolated from *Apis mellifera* L. bee-gut, exhibited a beneficial effect on honeybee colonies. *Beneficial Microbes* 2:29–34 DOI [10.3920/BM2010.0024](https://doi.org/10.3920/BM2010.0024).
- Audisio MC, Sabaté DC, Benítez-Ahrendts MR. 2015.** Effect of *Lactobacillus johnsonii* CRL1647 on different parameters of honeybee colonies and bacterial populations of the bee gut. *Beneficial Microbes* 6(5):687–695 DOI [10.3920/BM2014.0155](https://doi.org/10.3920/BM2014.0155).
- Bae S, Fleet GH, Heard GM. 2006.** Lactic acid bacteria associated with wine grapes from several Australian vineyards. *Journal of Applied Microbiology* 100:712–727 DOI [10.1111/j.1365-2672.2006.02890.x](https://doi.org/10.1111/j.1365-2672.2006.02890.x).
- Baughner JL, Durmaz E, Klaenhammer TR. 2014.** Spontaneously induced prophages in *Lactobacillus gasseri* contribute to horizontal gene transfer. *Applied and Environmental Microbiology* 80:3508–3517 DOI [10.1128/AEM.04092-13](https://doi.org/10.1128/AEM.04092-13).
- Berlin K, Koren S, Chin C-S, Drake JP, Landolin JM, Phillippy AM. 2015.** Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nature Biotechnology* 33:623–630 DOI [10.1038/nbt.3238](https://doi.org/10.1038/nbt.3238).
- Bi D, Xu Z, Harrison EM, Tai C, Wei Y, He X, Jia S, Deng Z, Rajakumar K, Ou H-Y. 2012.** ICEberg: a web-based resource for integrative and conjugative elements found in Bacteria. *Nucleic Acids Research* 40:D621–D626 DOI [10.1093/nar/gkr846](https://doi.org/10.1093/nar/gkr846).
- Boom R, Sol C, Beld M, Weel J, Goudsmit J, Wertheim-van Dillen P. 1999.** Improved silica-guanidiniumthiocyanate DNA isolation procedure based on selective binding of bovine alpha-casein to silica particles. *Journal of Clinical Microbiology* 37:615–619.

- Boyd EF, Almagro-Moreno S, Parent MA. 2009.** Genomic islands are dynamic, ancient integrative elements in bacterial evolution. *Trends in Microbiology* 17:47–53 DOI 10.1016/j.tim.2008.11.003.
- Burrus V, Pavlovic G, Decaris B, Guédon G. 2002.** Conjugative transposons: the tip of the iceberg. *Molecular Microbiology* 46:601–610 DOI 10.1046/j.1365-2958.2002.03191.x.
- Chang Y, Lee J-H, Shin H, Heu S, Ryu S. 2013.** Characterization and complete genome sequence analysis of *Staphylococcus aureus* bacteriophage SA12. *Virus Genes* 47:389–393 DOI 10.1007/s11262-013-0938-7.
- Corby-Harris V, Maes P, Anderson KE. 2014.** The bacterial communities associated with honey bee (*Apis mellifera*) foragers. *PLoS ONE* 9:e95056 DOI 10.1371/journal.pone.0095056.
- Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, Quan P-L, Briese T, Hornig M, Geiser DM, Martinson V, VanEngelsdorp D, Kalkstein AL, Drysdale A, Hui J, Zhai J, Cui L, Hutchison SK, Simons JF, Egholm M, Pettis JS, Lipkin WI. 2007.** A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318:283–287 DOI 10.1126/science.1146498.
- Daum M, Herrmann S, Wilkinson B, Bechthold A. 2009.** Genes and enzymes involved in bacterial isoprenoid biosynthesis. *Current Opinion in Chemical Biology* 13:180–188 DOI 10.1016/j.cbpa.2009.02.029.
- De Guchte van M, Serror P, Chervaux C, Smokvina T, Ehrlich SD, Maguin E. 2002.** Stress responses in lactic acid bacteria. *Antonie van Leeuwenhoek* 82:187–216 DOI 10.1023/A:1020631532202.
- Dhillon BK, Laird MR, Shay JA, Winsor GL, Lo R, Nizam F, Pereira SK, Waglechner N, McArthur AG, Langille MGI, Brinkman FSL. 2015.** IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Research* 43:W104–W108 DOI 10.1093/nar/gkv401.
- Di Pasquale G, Salignon M, Le Conte Y, Belzunces LP, Decourtye A, Kretzschmar A, Suchail S, Brunet J-L, Alaux C. 2013.** Influence of pollen nutrition on honey bee health: do pollen quality and diversity matter? *PLoS ONE* 8:e72016 DOI 10.1371/journal.pone.0072016.
- Di Prisco G, Cavaliere V, Annoscia D, Varricchio P, Caprio E, Nazzi F, Gargiulo G, Pennacchio F. 2013.** Neonicotinoid clothianidin adversely affects insect immunity and promotes replication of a viral pathogen in honey bees. *Proceedings of the National Academy of Sciences of the United States of America* 110:18466–18471 DOI 10.1073/pnas.1314923110.
- Djukic M, Poehlein A, Strauß J, Tann FJ, Leimbach A, Hoppert M, Daniel R. 2015.** High quality draft genome of *Lactobacillus kunkeei* EFB6, isolated from a German European foulbrood outbreak of honeybees. *Standards in Genomic Sciences* 10:Article 16 DOI 10.1186/1944-3277-10-16.

- Dorscht J, Klumpp J, Biemann R, Schmelcher M, Born Y, Zimmer M, Calendar R, Loessner MJ. 2009.** Comparative genome analysis of *Listeria* bacteriophages reveals extensive mosaicism, programmed translational frameshifting, and a novel prophage insertion site. *Journal of Bacteriology* **191**:7206–7215 DOI [10.1128/JB.01041-09](https://doi.org/10.1128/JB.01041-09).
- Edwards CG, Haag KM, Collins MD, Hutson RA, Huang YC. 1998.** *Lactobacillus kunkeei* sp. nov.: a spoilage organism associated with grape juice fermentations. *Journal of Applied Microbiology* **84**:698–702 DOI [10.1046/j.1365-2672.1998.00399.x](https://doi.org/10.1046/j.1365-2672.1998.00399.x).
- Ellegaard KM, Tamarit D, Javelind E, Olofsson TC, Andersson SGE, Vásquez A. 2015.** Extensive intra-phylogroup diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* **16**:284 DOI [10.1186/s12864-015-1476-6](https://doi.org/10.1186/s12864-015-1476-6).
- Endo A. 2012.** Fructophilic lactic acid bacteria inhabit fructose-rich niches in nature. *Microbial Ecology in Health and Disease* **23** DOI [10.3402/mehd.v23i0.18563](https://doi.org/10.3402/mehd.v23i0.18563).
- Endo A, Futagawa-Endo Y, Dicks LMT. 2009.** Isolation and characterization of fructophilic lactic acid bacteria from fructose-rich niches. *Systematic and Applied Microbiology* **32**:593–600 DOI [10.1016/j.syapm.2009.08.002](https://doi.org/10.1016/j.syapm.2009.08.002).
- Endo A, Irisawa T, Futagawa-Endo Y, Takano K, Toit du M, Okada S, Dicks LMT. 2012.** Characterization and emended description of *Lactobacillus kunkeei* as a fructophilic lactic acid bacterium. *International Journal of Systematic And Evolutionary Microbiology* **62**:500–504 DOI [10.1099/ijs.0.031054-0](https://doi.org/10.1099/ijs.0.031054-0).
- Endo A, Salminen S. 2013.** Honeybees and beehives are rich sources for fructophilic lactic acid bacteria. *Systematic and Applied Microbiology* **36**:444–448 DOI [10.1016/j.syapm.2013.06.002](https://doi.org/10.1016/j.syapm.2013.06.002).
- Engel P, Moran NA. 2013.** Functional and evolutionary insights into the simple yet specific gut microbiota of the honey bee from metagenomic analysis. *Gut Microbes* **4**:60–65 DOI [10.4161/gmic.22517](https://doi.org/10.4161/gmic.22517).
- Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015.** Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* **3**:e1319 DOI [10.7717/peerj.1319](https://doi.org/10.7717/peerj.1319).
- Evans JD, Schwarz RS. 2011.** Bees brought to their knees: microbes affecting honey bee health. *Trends in Microbiology* **19**:614–620 DOI [10.1016/j.tim.2011.09.003](https://doi.org/10.1016/j.tim.2011.09.003).
- Gaudin V, Hedou C, Verdon E. 2012.** Validation of two ELISA kits for the screening of tylosin and streptomycin in honey according to the European decision 2002/657/EC. *Food Additives & Contaminants: Part A* **30**:93–109 DOI [10.1080/19440049.2012.722696](https://doi.org/10.1080/19440049.2012.722696).
- Hacker J, Carniel E. 2001.** Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Reports* **2**:376–381 DOI [10.1093/embo-reports/kve097](https://doi.org/10.1093/embo-reports/kve097).
- Hastings WJ, Ritter MA, Chamakura KR, Kutty Everett GF. 2013.** Complete Genome of *Bacillus megaterium* Siphophage Staley. *Genome Announcements* **1**: e00864–13–e00864–13 DOI [10.1128/genomeA.00864-13](https://doi.org/10.1128/genomeA.00864-13).
- Henry M, Béguin M, Requier F, Rollin O, Odoux J-F, Aupinel P, Aptel J, Tchamitchian S, Decourtye A. 2012.** A common pesticide decreases foraging success and survival in honey bees. *Science* **336**:348–350 DOI [10.1126/science.1215039](https://doi.org/10.1126/science.1215039).

- Iannelli F, Santagati M, Santoro F, Oggioni MR, Stefani S, Pozzi G. 2014.** Nucleotide sequence of conjugative prophage Φ 1207.3 (formerly Tn1207.3) carrying the *mef(A)/msr(D)* genes for efflux resistance to macrolides in *Streptococcus pyogenes*. *Frontiers in Microbiology* **5**:Article 687 DOI [10.3389/fmicb.2014.00687](https://doi.org/10.3389/fmicb.2014.00687).
- Jiang W, Chen L, Hu N, Yuan S, Li B, Liu Z. 2014.** A novel serine hydroxymethyltransferase from *Arthrobacter nicotianae*: characterization and improving catalytic efficiency by rational design. *BMC Biotechnology* **14**:93 DOI [10.1186/s12896-014-0093-9](https://doi.org/10.1186/s12896-014-0093-9).
- Kankainen M, Paulin L, Tynkkynen S, Ossowski von I, Reunanen J, Partanen P, Satokari R, Vesterlund S, Hendrickx APA, Lebeer S, De Keersmaecker SCJ, Vanderleyden J, Hämäläinen T, Laukkanen S, Salovuori N, Ritari J, Alatalo E, Korpela R, Mattila-Sandholm T, Lassig A, Hatakka K, Kinnunen KT, Karjalainen H, Saxelin M, Laakso K, Surakka A, Palva A, Salusjärvi T, Auvinen P, De Vos WM. 2009.** Comparative genomic analysis of *Lactobacillus rhamnosus* GG reveals pili containing a human- mucus binding protein. *Proceedings of the National Academy of Sciences of the United States of America* **106**:17193–17198 DOI [10.1073/pnas.0908876106](https://doi.org/10.1073/pnas.0908876106).
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA. 2009.** Circos: an information aesthetic for comparative genomics. *Genome Research* **19**:1639–1645 DOI [10.1101/gr.092759.109](https://doi.org/10.1101/gr.092759.109).
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004.** Versatile and open software for comparing large genomes. *Genome Biology* **5**: Article R12 DOI [10.1186/gb-2004-5-2-r12](https://doi.org/10.1186/gb-2004-5-2-r12).
- Kwong WK, Engel P, Koch H, Moran NA. 2014.** Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proceedings of the National Academy of Sciences of the United States of America* **111**:11509–11514 DOI [10.1073/pnas.1405838111](https://doi.org/10.1073/pnas.1405838111).
- Lozo J, Begovic J, Jovicic B, Golic N, Topisirovic L. 2008.** Effect of methionine and cysteine deprivation on growth of different natural Isolates of *Lactobacillus* spp. in chemically defined media. *Archives of Biological Sciences* **60**:509–517 DOI [10.2298/ABS0804509L](https://doi.org/10.2298/ABS0804509L).
- Löytynoja A. 2014.** Phylogeny-aware alignment with PRANK. *Methods in Molecular Biology* **1079**:155–170 DOI [10.1007/978-1-62703-646-7_10](https://doi.org/10.1007/978-1-62703-646-7_10).
- Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA. 2011.** A simple and distinctive microbiota associated with honey bees and bumble bees. *Molecular Ecology* **20**:619–628 DOI [10.1111/j.1365-294X.2010.04959.x](https://doi.org/10.1111/j.1365-294X.2010.04959.x).
- McFrederick QS, Wcislo WT, Hout MC, Mueller UG. 2014.** Host species and developmental stage, but not host social structure, affects bacterial community structure in socially polymorphic bees. *FEMS Microbiology Ecology* **88**:398–406 DOI [10.1111/1574-6941.12302](https://doi.org/10.1111/1574-6941.12302).
- McFrederick QS, Wcislo WT, Taylor DR, Ishak HD, Dowd SE, Mueller UG. 2012.** Environment or kin: whence do bees obtain acidophilic bacteria? *Molecular Ecology* **21**:1754–1768 DOI [10.1111/j.1365-294X.2012.05496.x](https://doi.org/10.1111/j.1365-294X.2012.05496.x).
- Merchant S, Wood DE, Salzberg SL. 2014.** Unexpected cross-species contamination in genome sequencing projects. *PeerJ* **2**:e675 DOI [10.7717/peerj.675](https://doi.org/10.7717/peerj.675).

- Musumeci L, Bongiorno C, Tautz L, Edwards RA, Osterman A, Perego M, Mustelin T, Bottini N. 2005. Low-molecular-weight protein tyrosine phosphatases of *Bacillus subtilis*. *Journal Of Bacteriology* **187**:4945–4956
DOI [10.1128/JB.187.14.4945-4956.2005](https://doi.org/10.1128/JB.187.14.4945-4956.2005).
- Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA, Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z, Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter JC. 2000. A whole-genome assembly of *Drosophila*. *Science* **287**:2196–2204
DOI [10.1126/science.287.5461.2196](https://doi.org/10.1126/science.287.5461.2196).
- Neveling DP, Endo A, Dicks LMT. 2012. Fructophilic *Lactobacillus kunkeei* and *Lactobacillus brevis* isolated from fresh flowers, bees and bee-hives. *Current Microbiology* **65**:507–515
DOI [10.1007/s00284-012-0186-4](https://doi.org/10.1007/s00284-012-0186-4).
- Olmos A, Henríquez-Piskulich P, Sanchez C, Rojas-Herrera M, Moreno-Pino M, Gómez M, Rodríguez Da Silva R, Maracaja-Coutinho V, Aldea P, Trombert AN. 2014. Draft Genome of Chilean Honeybee (*Apis mellifera*) Gut Strain *Lactobacillus kunkeei* MP2. *Genome Announcements* **2**: e01013–14–e01013–14
DOI [10.1128/genomeA.01013-14](https://doi.org/10.1128/genomeA.01013-14).
- Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**(22):3691–3693
DOI [10.1093/bioinformatics/btv421](https://doi.org/10.1093/bioinformatics/btv421).
- Persuh M, Mandic-Mulec I, Dubnau D. 2002. A MecA paralog, YpbH, binds ClpC, affecting both competence and sporulation. *Journal of Bacteriology* **184**:2310–2313
DOI [10.1128/JB.184.8.2310-2313.2002](https://doi.org/10.1128/JB.184.8.2310-2313.2002).
- Podell S, Gaasterland T. 2007. DarkHorse: a method for genome-wide prediction of horizontal gene transfer. *Genome Biology* **8**:Article R16
DOI [10.1186/gb-2007-8-2-r16](https://doi.org/10.1186/gb-2007-8-2-r16).
- Porcellato D, Frantzen C, Rangberg A, Umu OC, Gabrielsen C, Nes IF, Amdam GV, Diep DB. 2015. Draft genome sequence of *Lactobacillus kunkeei* AR114 isolated from honey bee gut. *Genome Announcements* **3**:e00144–e00115
DOI [10.1128/genomeA.00144-15](https://doi.org/10.1128/genomeA.00144-15).
- Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldón T, Rattei T, Creevey C, Kuhn M, Jensen LJ, Mering von C, Bork P. 2014. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Research* **42**:D231–D239
DOI [10.1093/nar/gkt1253](https://doi.org/10.1093/nar/gkt1253).
- Prepiak P, Defrancesco M, Spadavecchia S, Mirouze N, Albano M, Persuh M, Fujita M, Dubnau D. 2011. MecA dampens transitions to spore, biofilm exopolysaccharide and competence expression by two different mechanisms. *Molecular Microbiology* **80**:1014–1030
DOI [10.1111/j.1365-2958.2011.07627.x](https://doi.org/10.1111/j.1365-2958.2011.07627.x).
- Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**:e9490
DOI [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490).
- Pruesse E, Peplies J, Glöckner FO. 2012. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**:1823–1829
DOI [10.1093/bioinformatics/bts252](https://doi.org/10.1093/bioinformatics/bts252).

- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013.** The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**:D590–D596 DOI [10.1093/nar/gks1219](https://doi.org/10.1093/nar/gks1219).
- Reynaldi FJ, Lacunza J, Alippi AM, Rule R. 2010.** Unión de los antibióticos tilosina, tilmicosina y oxitetraciclina a proteínas presentes en abejas, larvas y productos de la colmena de *Apis mellifera* L. *Revista Argentina de Microbiología* **42**:279–283.
- Savabi O, Kazemi M, Kamali S, Salehi AR, Eslami G, Tahmourespour A, Salehi R. 2014.** Effects of biosurfactant produced by *Lactobacillus casei* on gtfB, gtfC, and ftf gene expression level in *S. mutans* by real-time RT-PCR. *Advanced Biomedical Research* **3**: Article 231 DOI [10.4103/2277-9175.145729](https://doi.org/10.4103/2277-9175.145729).
- Schlothauer T, Mogk A, Dougan DA, Bukau B, Turgay K. 2003.** MecA, an adaptor protein necessary for ClpC chaperone activity. *Proceedings of the National Academy of Sciences of the United States of America* **100**:2306–2311 DOI [10.1073/pnas.0535717100](https://doi.org/10.1073/pnas.0535717100).
- Seefeldt KE, Weimer BC. 2000.** Diversity of sulfur compound production in lactic acid bacteria. *Journal of Dairy Science* **83**:2740–2746 DOI [10.3168/jds.S0022-0302\(00\)75168-X](https://doi.org/10.3168/jds.S0022-0302(00)75168-X).
- Seemann T. 2014.** Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**:2068–2069 DOI [10.1093/bioinformatics/btu153](https://doi.org/10.1093/bioinformatics/btu153).
- Shemesh M, Tam A, Feldman M, Steinberg D. 2006.** Differential expression profiles of *Streptococcus mutans* ftf, gtf and vicR genes in the presence of dietary carbohydrates at early and late exponential growth phases. *Carbohydrate Research* **341**:2090–2097 DOI [10.1016/j.carres.2006.05.010](https://doi.org/10.1016/j.carres.2006.05.010).
- Shterzer N, Mizrahi I. 2015.** The animal gut as a melting pot for horizontal gene transfer. *Canadian Journal of Microbiology* **61**(9):603–605 DOI [10.1139/cjm-2015-0049](https://doi.org/10.1139/cjm-2015-0049).
- Sloan DB, Moran NA. 2012.** Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biology Letters* **8**:986–989 DOI [10.1098/rsbl.2012.0664](https://doi.org/10.1098/rsbl.2012.0664).
- Tamarit D, Ellegaard KM, Wikander J, Olofsson T, Vásquez A, Andersson SGE. 2015.** Functionally structured genomes in *Lactobacillus kunkeei* colonizing the honey crop and food products of honeybees and stingless bees. *Genome Biology and Evolution* **7**:1455–1473 DOI [10.1093/gbe/evv079](https://doi.org/10.1093/gbe/evv079).
- Tarapoulouzi M, Papachrysostomou C, Constantinou S, Kanari P, Hadjigeorgiou M. 2013.** Determinative and confirmatory method for residues of tetracyclines in honey by LC-MS/MS. *Food additives & contaminants. Part A, Chemistry, Analysis, Control, Exposure & Risk Assessment* **30**:1728–1732.
- Tsumori H, Kuramitsu H. 1997.** The role of the *Streptococcus mutans* glucosyltransferases in the sucrose-dependent attachment to smooth surfaces: essential role of the GtfC enzyme. *Oral Microbiology and Immunology* **12**:274–280 DOI [10.1111/j.1399-302X.1997.tb00391.x](https://doi.org/10.1111/j.1399-302X.1997.tb00391.x).
- Vásquez A, Forsgren E, Fries I, Paxton RJ, Flaberg E, Szekely L, Olofsson TC. 2012.** Symbionts as major modulators of insect health: lactic acid bacteria and honeybees. *PLoS ONE* **7**:e33188 DOI [10.1371/journal.pone.0033188](https://doi.org/10.1371/journal.pone.0033188).

- Ventura M, Canchaya C, Bernini V, Altermann E, Barrangou R, McGrath S, Claesson MJ, Li Y, Leahy S, Walker CD, Zink R, Neviani E, Steele J, Broadbent J, Klaenhammer TR, Fitzgerald GF, O'Toole PW, Van Sinderen D. 2006. Comparative genomics and transcriptional analysis of prophages identified in the genomes of *Lactobacillus gasseri*, *Lactobacillus salivarius*, and *Lactobacillus casei*. *Applied and Environmental Microbiology* 72:3130–3146 DOI [10.1128/AEM.72.5.3130-3146.2006](https://doi.org/10.1128/AEM.72.5.3130-3146.2006).
- Ventura M, Canchaya C, Pridmore RD, Brüssow H. 2004. The prophages of *Lactobacillus johnsonii* NCC 533: comparative genomics and transcription analysis. *Virology* 320:229–242 DOI [10.1016/j.virol.2003.11.034](https://doi.org/10.1016/j.virol.2003.11.034).
- Wahl A, Servais F, Drucbert A-S, Foulon C, Fontaine L, Hols P. 2014. Control of natural transformation in salivarius *Streptococci* through specific degradation of σ_X by the MecA-ClpCP protease complex. *Journal of Bacteriology* 196:2807–2816 DOI [10.1128/JB.01758-14](https://doi.org/10.1128/JB.01758-14).
- Weber FJ, De Bont JA. 1996. Adaptation mechanisms of microorganisms to the toxic effects of organic solvents on membranes. *Biochimica et Biophysica Acta* 1286:225–245 DOI [10.1016/S0304-4157\(96\)00010-X](https://doi.org/10.1016/S0304-4157(96)00010-X).
- Yasmin A, Kenny JG, Shankar J, Darby AC, Hall N, Edwards C, Horsburgh MJ. 2010. Comparative genomics and transduction potential of *Enterococcus faecalis* temperate bacteriophages. *Journal of Bacteriology* 192:1122–1130 DOI [10.1128/JB.01293-09](https://doi.org/10.1128/JB.01293-09).
- Yousefi B, Ghaderi S, Rezapoor-Lactooyi A, Amiri N, Verdi J, Shoaie-Hassani A. 2012. Hydroxy decenoic acid down regulates *gtfB* and *gtfC* expression and prevents *Streptococcus mutans* adherence to the cell surfaces. *Annals of Clinical Microbiology and Antimicrobials* 11:21 DOI [10.1186/1476-0711-11-21](https://doi.org/10.1186/1476-0711-11-21).
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Research* 39:W347–W352 DOI [10.1093/nar/gkr485](https://doi.org/10.1093/nar/gkr485).