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

Freddy Asenjo, Alejandro Olmos, Patricia Henríquez-Piskulich, Patricia Aldea, Juan A Ugalde, Annette N Trombert

**Background**. 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 CDD or Colony Collapsed Disorder, characterized by the seemingly spontaneous abandonment of their hives. One of the characteristics of CDD in honey bees is the alteration of the bacterial communities in their gastrointestinal track, mainly due to the decrease of *Firmicutes* populations, such as the *Lactobacilli*. At present, 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 prophage sequences using PHAST. Comparisons between *L. kunkeei* and other *Lactobacillus* species were performed using GET_HOMOLOGUES. **Results.** The complete genome of *L. kunkeei* MP2 comprises one circular chromosome of 1,614,522 nt. with a GC content of 36,9%. Comparisons of *L. kunkeei* MP2 with two other *L. kunkeei* sequences (*L. kunkeei* EFB6 and *L. kunkeei* AR114) showed 95 unique genes, most of them related to phage insertions. In fact, a large and unique region of *L. kunkeei* MP2 genome contains several phage genes that encode for phage structural proteins, including replication components. In comparisons of MP2 with other *Lactobacillus* species, we identified several unique genes of *L. kunkeei* MP2 related with metabolism, biofilm generation, survival under stress conditions, and mobile genetic elements (MGEs).

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

Freddy Asenjo[1], Alejandro Olmos[1], Patricia Henríquez-Piskulich[2], Patricia Aldea[2], Juan A. Ugalde[1], Annette N. Trombert[1]

[1]Centro de Genómica y Bioinformática, Facultad de Ciencias, Universidad Mayor, Santiago, Chile.

[2]CEAPI Mayor, Facultad de Ciencias, Universidad Mayor, Santiago, Chile.

Corresponding Authors:

Annette N. Trombert[1]

Camino a La Piramide 5750, Santiago, Chile

Email address: annette.trombert@umayor.cl

Juan A. Ugalde[1]

Camino a La Piramide 5750, Santiago, Chile

Email address: juan@ecogenomica.cl

## Abstract

**Background**. 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 CDD or Colony Collapsed Disorder, characterized by the seemingly spontaneous abandonment of their hives. One of the characteristics of CDD in honey bees is the alteration of the bacterial communities in their gastrointestinal track, mainly due to the decrease of *Firmicutes* populations, such as the *Lactobacilli*. At present, 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* and other *Lactobacillus* species were performed using GET_HOMOLOGUES.

**Results.** The complete genome of *L. kunkeei* MP2 comprises one circular chromosome of 1,614,522 nt. with a GC content of 36,9%. Comparisons of *L. kunkeei* MP2 with two

other *L. kunkeei* sequences (*L. kunkeei* EFB6 and *L. kunkeei* AR114) showed 95 unique genes, most of them related to phage insertions. In fact, a large and unique region of *L. kunkeei* MP2 genome contains several phage genes that encode for phage structural proteins, including replication components. In comparisons of MP2 with other *Lactobacillus* species, we 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) could given by its genetic capacity to change and adapt to different environments. This is particularly interesting because a vast diversity of phenotypes could be observed depending of strain of *L. kunkeei.* In fact, the three strains analyzed in this study were genetically different. Thus, *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.

## 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 present, an unusual decreased

in honey bee colonies has been taking place, known as Colony Collapse Disorder (CCD). CCD describes the seemingly spontaneous abandonment of their hives by honeybees, where queens often survive accompanied by a small group of younger worker bees. The specific causes of CCD are unknown, but several factors can impact the health of honeybees, 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 (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 honeybee 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 with 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 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). Also 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* corresponds to the most frequent lactobacilli isolate from the honey bee gut (Corby-Harris, Maes & Anderson, 2014; McFrederick et al., 2014).

In previous work, we isolated an *L. kunkeei* strain (denominated 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 to completely resolve the genome without gaps. In addition, multiple repetitions, including the presence of multiple copies of the ribosomal operon,

115 could not be resolved in this draft genome. To overcome these limitations, we
116 performed a re-sequencing of the *L. kunkeei* MP2 genome, using single molecule
117 sequencing in the Pacific Biosciences platform.
118 In this work, we report the first complete genome sequence of *L. kunkeei* MP2, its
119 characterization, and comparison with other *Lactobacillus* genomes.
120

121 **Methods**

122 *DNA isolation and genome sequencing*

123 Honey bees (*Apis mellifera)* were collected from commune hives located in the María
124 Pinto area, at the Melipilla Province in the Central zone of Chile. *L. kunkeei* MP2 was
125 isolated from the guts of these bees, as previously described (Olmos et al., 2014).
126 For DNA extraction, *L. kunkeei* MP2 colonies were cultured in MRS broth (37°C, 5%
127 CO2) and genomic DNA was obtained using a silica-based protocol (Boom et al., 1999).
128 Briefly, bacterial pellet was lysed using a solution composed of SDS 10%, proteinase K
129 (10 mg/mL, Thermo Scientific) and lysozyme (5 mg/mL, Pierce) at 37°C for 60 min. The
130 lysate was mixed with guanidine chloride 6M and a silica suspension (50% w/v) and
131 incubated for 10 min. The silica was centrifuged, and DNA was recovered, after washes
132 with 70% alcohol, into sterile, free nuclease water. Approximately 13.5 µg of DNA were
133 used to construct sequencing libraries with an average insert size of 20kb, and
134 sequenced using one SMRT cell (P6-C4 Chemistry) on a PacBio RSII sequencer
135 (Pacific Biosciences) at the UCSD IGM Genomics Center.
136 *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, 9Kb) 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 (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 0.67 (Krzywinski et al., 2009).

*Pan-genome analysis*

Comparative genomic analysis was performed from a list of selected *Lactobacillus* genomes from Genbank (April 2015). Supplementary Table 1 shows the species name and accession numbers. To avoid possible differences due to different annotation procedures, all of the genomes were re-annotated using Prokka version 1.11 (Seemann, 2014). Comparisons between the selected *L. kunkeei* strains and all the *Lactobacillus* genomes, were done using GET_HOMOLOGUES version 20150306 (Contreras-Moreira & Vinuesa, 2013). All the re-annotated genomes were compared

using the OMCL and COG algorithms, and a pan-genome matrix was generated for each of the different comparisons performed.

## Results and Discussion

### *Assembly description and comparison*

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 1 (gene annotation on Supplementary Table 2). 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 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 on Figure 1. All 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.

When this work was performed (April 2015), there were two other sequenced strains of *L. kunkeei,* EFB6 (Djukic et al., 2015) and AR114 (Porcellato et al., 2015). Comparison

of these two genomes with MP2, shows a region of high sequence identity (Figure 1), as well as the presence of variable regions present in only some of the strains. In the case of MP2, we found unique regions not present in the other draft genomes, which will be discussed in more detail later in the text.

*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 include carotenoids, sterols, prenyl side-chains of chlorophylls, and plastoquinone, exhibiting

206 many biological functions (Daum et al., 2009). In whiteflies (*Bemisia tabaci*), the
207 genome of its endosymbiotic bacteria, *Candidatus Portiera aleyrodidarum*, encodes for
208 key enzymes in carotenoids synthesis, suggesting that whitefly not only can acquire
209 carotenoids from the diet, but also from their microbiota (Sloan & Moran, 2012).
210 Therefore, if *L. kunkeei* produces key enzymes involved in isoprenoid synthesis, it is
211 possible that it could be playing an important role in honey bee nutrition.
212 *Nucleotide and amino acid metabolism.* The pathways for purine biosynthesis are
213 complete. However, in the de novo pirimidine pathway, *L. kunkeei* lacks the gene *pyrB*,
214 which codifies for the aspartate carbamoyl transferase, and *ndk*, codifying for the
215 nucleoside diphosphate kinase. The analysis of the metabolic pathways in MP2
216 revealed a minimal amino acid auxotrophy (methionine or cysteine), with the presence
217 of the genes that encode for a D-methionine transport system, suggesting the ability of
218 *L. kunkeei* MP2 to acquire methionine/cysteine from the environment. These results are
219 in line with previous reports of the lactobacilli being auxotrophic for both methionine and
220 cysteine (Seefeldt & Weimer, 2000), and where the supplement of culture media with
221 these amino acids improved bacterial growth (Lozo et al., 2008). A gene that encodes
222 for serine hydroxymethyltransferase (SHMT) was found in the genome of MP2. This
223 enzyme catalyzes the addition of formaldehyde to glycine, a key step for the production
224 of serine (Jiang et al., 2014), and appears to be absent in the other *Lactobacillus*
225 genomes analyzed in this study. Its presence in *L. kunkeei* MP2 could be part of specific
226 adaptation mechanisms of this species to its environment.
227
228 *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; 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 (Supplementary Table 3). One of them, located in the region between 594,506 to 613,136, was found to be present in all the 23 *Lactobacillus* genomes used in this work. The second region, located around 32,973 to 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 with two other publicly available *L. kunkeei* strains (EFB6 and AR114). This analysis can provide us with a snapshot of the unique adaptive features that are present in this strain. While recent work focused on the pan-genomic study of *L. kunkeei* (Tamarit et al., 2015), we focused our efforts on

identifying unique elements that could help us to understand the evolutionary history of *L. kunkeei* MP2.

To identify genes that are unique to MP2, we performed a clustering analysis using the GET_HOMOLOGUEs software (Contreras-Moreira & Vinuesa, 2013), with the OMCL and COG algorithms. The comparison was done against the EFB6 and AR114 genomes, and against a set of 22 other *Lactobacillus* genomes (Supplementary Table 1). In the case of the *L. kunkeei* comparison (MP2 vs EFB6 and AR114), we identified 95 genes (6.12% of the total genes) as unique (Supplementary Table 3).

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 (Figure 1). It is a large region, which includes several phage-related proteins, including structural and replications components. Analysis of the sequences shows that these proteins are related to phages that 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) (Supplementary Table 4).

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

Several unique genes were found in *L. kunkeei* MP2, compared with the genome of other *Lactobacillus* species. 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

274 considered a virulence factor in *S. mutants*, promoting bacterial adhesion to smooth
275 surfaces and cells (Tsumori & Kuramitsu, 1997). Also, GtfC is part of the synthesis route
276 of a mixture of insoluble and soluble glucans, which are important components of
277 cariogenic biofilms (Yousefi et al., 2012). Considering the rich carbohydrate
278 environment where *L. kunkeei* can survive, the presence of unique glucosyltransferase
279 genes, such as GftC, could facilitate bacterial colonization of flowers, as well as the
280 honey bee gut.
281 Another unique gene found in *L. kunkeei*, encodes for the adapter protein MecA, a
282 pleiotropic regulator of bacterial development. This protein has been shown to affect
283 competence, protein degradation and sporulation in Bacteria, such as *Bacillus subtilis*
284 (Schlothauer et al., 2003). MecA interacts with the chaperone ClpC, and with the
285 transcription factor ComK, promoting the degradation of this protein during the
286 logarithmic growth phase. The degradation of ComK stops when bacteria enters to
287 stationary growth phase, where the quorum-sensing pheromone ComX promotes the
288 synthesis of ComS, which binds to MecA and prevents the interaction of MecA-ComK
289 (Persuh, Mandic-Mulec & Dubnau, 2002; Prepiak et al., 2011; Wahl et al., 2014). This
290 could have an effect on the biofilm generation capabilities of *L. kunkeei* MP2, but this
291 needs to be explored experimentally.
292 At least seven different *Lactobacillus* species have been characterized in the gut
293 microbiota of *A. mellifera*, where it has been suggested that they play different roles in
294 the stability of the host functions (Engel & Moran, 2013). *L. kunkeei* MP2 appears to
295 have a unique set of genes when compared to other strains of *L. kunkeei*, as well as

296 with other species of *Lactobacillus* (Supplementary Table 3), which suggest unique
297 adaptation strategies of *L. kunkeei* MP2 to the gut of *A. mellifera*.
298 We also identified a hypothetical protein with similarities to a low-molecular-weight
299 protein-tyrosine phosphatase (LMPTP), unique to the *L. kunkeei* MP2 genome,
300 compared to other *L.kunkeei* strains and also other *Lactobacillus* species. This LMPTP
301 is similar to the *YfkJ* protein from *Bacillus subtilis*, which has been involved in the
302 response to ethanol stress (Musumeci et al., 2005). Ethanol, and other organic
303 compunds, are commonly present in the environment, and accumulate in the bacterial
304 membrane affecting its physical-chemical properties, and in consequence, their
305 functions (Weber & de Bont, 1996). This could suggest a better tolerance to organic
306 compounds, such as ethanol, for *L. kunkeei* MP2, which could help this organism to
307 tolerate unfavorable conditions, and have a unique competitive advantage compared to
308 other *Lactobacillus* species (de Guchte et al., 2002).
309 The diversity of *Firmicutes* species in *A. mellifera* could imply a metabolic diversity that
310 could be crucial for honey bee fitness (Engel & Moran, 2013). Comparative genomics of
311 *Lactobacillus* genomes, have shown that close to 45% of its accessory genome encode
312 for proteins involved in carbohydrate metabolism and transport functions (Ellegaard et
313 al., 2015). With this metabolic diversity found in the accessory genome, is no surprising
314 to find unique genes in the accessory genome of *L. kunkeei* MP2, when compared to
315 other strains of *L. kunkeei*, as well as other *Lactobacillus* species (Supplementary Table
316 3). These genes encode for proteins that take part of the degradation of carbohydrates,
317 transport of molecules, transcription, as well as membrane proteins. It is very likely that
318 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* as well, something that has been observed in the adaptation of strains of *Gilliamela 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 (Figure 1, Supplementary Table 3). 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) (Supplementary Table 3). 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

342 mechanisms, suggesting that the incorporation and stability of these unique genes in
343 the genome of *L. kunkeei* MP2, is providing an increase in the fitness of this bacterial
344 strain in the gut of *A. mellifera*. Among the predicted phage-like sequences, we found
345 one coding for a *mef(A)/msr(D)* resistance protein, with similarity to a sequence from
346 *Streptococcus pyogenes*, involved in the resistance to macrolides (Iannelli et al., 2014).
347 In the European Union, the usage of antibiotic and antibiotic-containing compounds is
348 not permitted. However, macrolides (such as tylosin and streptomycin), are still used as
349 a preventive treatment against *Paenibacillus larvae*, the causal agent of American
350 foulbrood, in many countries (Reynaldi et al., 2010; Gaudin, Hedou & Verdon, 2012).
351 Thus, if bees were exposed to antibiotics in their diet, is of not surprise that their
352 microbiota may have acquired the necessary molecular mechanisms to adapt and
353 survive in an exposed environment, such as the insect gut.
354 Most of ICEs coding genes are usually present within genomic islands in the host
355 genome (Hacker & Carniel, 2001; Boyd, Almagro-Moreno & Parent, 2009), but in the
356 case of *L. kunkeei* MP2, none of the predicted ICEs genes were found in the context of
357 genomic islands according to the predictions performed with IslandViewer 3 (Dhillon et
358 al., 2015). This could suggest the presence of previously uncharacterized ICEs, or also
359 our current limitation in the detection of ICEs from *Lactobacillus* species.
360
361 *Prediction of horizontal gene transfer events*
362 To predict horizontally transferred genes we used Darkhorse (Podell & Gaasterland,
363 2007) to analyze the complete genome of *L. kunkeei* MP2. We did not consider hits to
364 organisms within the same Phylum, to avoid false predictions, although this could lead

365 to ignore real transfer events between more closely related organisms. A total of 19
366 genes were predicted to have been acquired via horizontal gene transfer
367 (Supplementary Table 3), with a normalized LPI score cutoff of 0.546. Seven of these
368 genes had matches with the genome of *A. mellifera*, and when we looked in more detail,
369 this suggest a contamination of the genome of *A. mellifera* with sequences of
370 *Lactobacillus*, something that has been previously reported for other genome projects
371 (Merchant, Wood & Salzberg, 2014). Only one of the genes predicted to be acquired via
372 HGT was unique to *L. kunkeei* MP2 when compared to other *L. kunkeei* strains and
373 other *Lactobacillus* genomes, which codifies for a hypothetical protein, with a best hit as
374 a phage protein from *Halomonas* sp. HAL1. None of the predicted genes was found
375 associated with an ICE or a genomic island. This could be due to a lack of reference
376 genomes of isolates obtained from the gut of *A. mellifera*, limiting our ability to predict
377 events of horizontal gene transfer between members of the honey gut microbiota.
378

379 **Conclusions**

380 In this study we present the results of the sequencing and analysis of the first closed
381 genome for a *L. kunkeei* strain, isolated from gut of *A. mellifera*. The comparison of the
382 genome sequence against other *Lactobacilus* species, showed a percentage of genes
383 that are unique to the MP2 strain, including metabolic key enzymes that could play an
384 important role in the honey bee nutrition and fitness. The genome of *L. kunkeei* MP2
385 also has genes encoding for proteins involved in important roles such as adhesion,
386 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 sequence that are 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, this suggest a future goal of investigation, to elucidate the role of these proteins in *L. kunkeei* MP2.

## References

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.

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.

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*:1–10.

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.

Baugher JL, Durmaz E, Klaenhammer TR 2014. Spontaneously induced prophages in Lactobacillus gasseri contribute to horizontal gene transfer. *Applied and Environmental Microbiology* 80:3508–3517.

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.

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:D621–6.
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.
Burrus V, Pavlovic G, Decaris B, Guédon G 2002. Conjugative transposons: the tip of the iceberg. *Molecular Microbiology* 46:601–610.
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.
Contreras-Moreira B, Vinuesa P 2013. GET_HOMOLOGUES, a versatile software package for scalable and robust microbial pangenome analysis. *Applied and Environmental Microbiology* 79:7696–7701.
Corby-Harris V, Maes P, Anderson KE 2014. The bacterial communities associated with honey bee (Apis mellifera) foragers. *PLoS ONE* 9:e95056.
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.
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.
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.
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:W104–8.
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.
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* 110:18466–18471.
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* 10:16.
Dorscht J, Klumpp J, Bielmann 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.
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.
Ellegaard KM, Tamarit D, Javelind E, Olofsson TC, Andersson SGE, Vásquez A 2015. Extensive intra-phylotype diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* 16:284.
Endo A 2012. Fructophilic lactic acid bacteria inhabit fructose-rich niches in nature. *Microbial ecology in health and disease* 23:647.
Endo A, Salminen S 2013. Honeybees and beehives are rich sources for fructophilic lactic acid bacteria. *Systematic and Applied Microbiology* 36:444–448.
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.
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.
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.
Evans JD, Schwarz RS 2011. Bees brought to their knees: microbes affecting honey bee health. *Trends in Microbiology* 19:614–620.
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.
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.
Hastings WJ, Ritter MA, Chamakura KR, Kuty Everett GF 2013. Complete Genome of Bacillus megaterium Siphophage Staley. *Genome announcements* 1:e00864–13–e00864–13.
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.
Iannelli F, Santagati M, Santoro F, Oggioni MR, Stefani S, Pozzi G 2014. Nucleotide sequence of conjugative prophage Φ1207.3 (formerly Tn1207.3) carrying the mef(A)/msr(D) genes for efflux resistance to macrolides in Streptococcus pyogenes. *Frontiers in Microbiology* 5:687.
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.
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*

512 *of Sciences* 106:17193–17198.
513 Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra
514 MA 2009. Circos: An information aesthetic for comparative genomics. *Genome*
515 *Research* 19:1639–1645.
516 Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL
517 2004. Versatile and open software for comparing large genomes. *Genome Biology*
518 5:R12.
519 Kwong WK, Engel P, Koch H, Moran NA 2014. Genomics and host specialization of
520 honey bee and bumble bee gut symbionts. *Proceedings of the National Academy of*
521 *Sciences* 111:11509–11514.
522 Lozo J, Begovic J, Jovcic B, Golic N, Topisirovic L 2008. Effect of Methionine and
523 Cysteine Deprivation on Growth of Different Natural Isolates of Lactobacillus Spp. in
524 Chemically Defined Media. *Archives of Biological Sciences* 60:509–517.
525 Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA 2011. A
526 simple and distinctive microbiota associated with honey bees and bumble bees.
527 *Molecular Ecology* 20:619–628.
528 McFrederick QS, Wcislo WT, Hout MC, Mueller UG 2014. Host species and
529 developmental stage, but not host social structure, affects bacterial community
530 structure in socially polymorphic bees. *FEMS Microbiology Ecology* 88:398–406.
531 Merchant S, Wood DE, Salzberg SL 2014. Unexpected cross-species contamination in
532 genome sequencing projects. *PeerJ* 2:e675.
533 Musumeci L, Bongiorni C, Tautz L, Edwards RA, Osterman A, Perego M, Mustelin T,
534 Bottini N 2005. Low-molecular-weight protein tyrosine phosphatases of Bacillus
535 subtilis. *Journal Of Bacteriology* 187:4945–4956.
536 Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, Flanigan MJ, Kravitz SA,
537 Mobarry CM, Reinert KH, Remington KA, Anson EL, Bolanos RA, Chou HH, Jordan
538 CM, Halpern AL, Lonardi S, Beasley EM, Brandon RC, Chen L, Dunn PJ, Lai Z,
539 Liang Y, Nusskern DR, Zhan M, Zhang Q, Zheng X, Rubin GM, Adams MD, Venter
540 JC 2000. A whole-genome assembly of Drosophila. *Science* 287:2196–2204.
541 Neveling DP, Endo A, Dicks LMT 2012. Fructophilic Lactobacillus kunkeei and
542 Lactobacillus brevis isolated from fresh flowers, bees and bee-hives. *Current*
543 *microbiology* 65:507–515.
544 Olmos A, Henríquez-Piskulich P, Sanchez C, Rojas-Herrera M, Moreno-Pino M, Gómez
545 M, Rodríguez Da Silva R, Maracaja-Coutinho V, Aldea P, Trombert AN 2014. Draft
546 Genome of Chilean Honeybee (Apis mellifera) Gut Strain Lactobacillus kunkeei
547 MP2. *Genome announcements* 2:e01013–14–e01013–14.
548 Persuh M, Mandic-Mulec I, Dubnau D 2002. A MecA paralog, YpbH, binds ClpC,
549 affecting both competence and sporulation. *Journal Of Bacteriology* 184:2310–2313.
550 Podell S, Gaasterland T 2007. DarkHorse: a method for genome-wide prediction of
551 horizontal gene transfer. *Genome Biology* 8:R16.
552 Porcellato D, Frantzen C, Rangberg A, Umu OC, Gabrielsen C, Nes IF, Amdam GV,
553 Diep DB 2015. Draft Genome Sequence of Lactobacillus kunkeei AR114 Isolated
554 from Honey Bee Gut. *Genome announcements* 3:e00144–15.
555 Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, Gabaldón T,
556 Rattei T, Creevey C, Kuhn M, Jensen LJ, Mering von C, Bork P 2014. eggNOG
557 v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Research*

42:D231–D239.
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.
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* 3:231.
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.
Seefeldt KE, Weimer BC 2000. Diversity of sulfur compound production in lactic acid bacteria. *Journal of dairy science* 83:2740–2746.
Seemann T 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England)* 30:2068–2069.
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.
Shterzer N, Mizrahi I 2015. The animal gut as a melting pot for horizontal gene transfer. *Canadian Journal of Microbiology*:1–3.
Sloan DB, Moran NA 2012. Endosymbiotic bacteria as a source of carotenoids in whiteflies. *Biology Letters* 8:986–989.
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.
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.
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.
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.
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.

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 σX by the MecA-ClpCP protease complex. *Journal Of Bacteriology* 196:2807–2816.

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.

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.

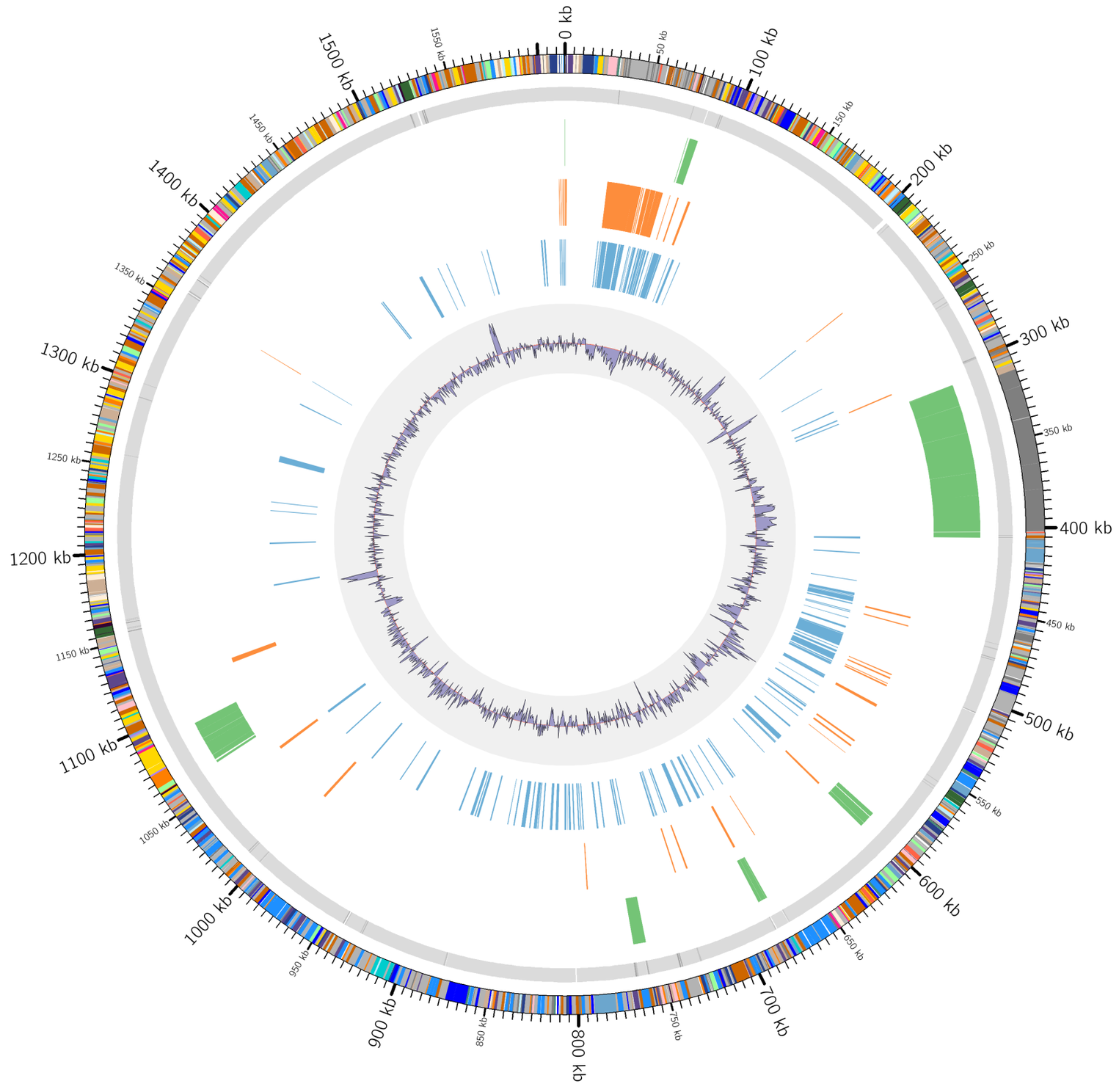
Yousefi B, Ghaderi S, Rezapoor-Lactooyi A, Amiri N, Verdi J, Shoae-Hassani A 2012. Hydroxy decenoic acid down regulates gtfB and gtfC expression and prevents Streptococcus mutans adherence to the cell surfaces. *Annals of clinical microbiology and antimicrobials* 11:21.

Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS 2011. PHAST: A Fast Phage Search Tool. *Nucleic Acids Research* 39:W347–W352.

## Figure 1(on next page)

Genome organisation of *L. kunkeei* MP2

Circular overview of the complete genome of *L. kunkee*i MP2, highlighthing some of the features. Starting from the outside ring towards the interior: EggNOG annotation of the predicted CDS; Contig recruitment of the previous *L. kunkeei* MP2 genome sequencing (Olmos et al., 2014); Phage island predictions using Island Viewer 3; Unique genes of *L. kunkeei* MP2, compared to strains EFB6 and AR114 of *L. kunkeei*; Unique genes of *L. kunkeei* MP2 compared with 22 genomes of *Lactobacillus* species; %GC contento of the *L. kunkeei* MP2 genome




## Information Storage and Processing

- Translation, ribosomal structure and biogenesis
- RNA processing and modification
- Transcription

## Cellular Processes and Signaling

- Cell cycle control, cell division, chromosome partitioning
- Defense mechanisms
- Signal transduction mechanisms
- Cell wall/membrane/envelope biogenesis
- Cell motility
- Intracellular trafficking, secretion, and vesicular transport
- Post-translational modification, protein turnover, and chaperones

## Metabolism

- Energy production and conversion
- Carbohydrate transport and metabolism
- Amino acid transport and metabolism
- Nucleotide transport and metabolism
- Coenzyme transport and metabolism
- Lipid transport and metabolism
- Inorganic ion transport and metabolism
- Secondary metabolites biosynthesis, transport and catabolism

## Poorly Characterized

- Function unknown

## Table 1(on next page)

Summary of the functional assignation for the predicted genes of *L. kunkeei* MP2 based on EggNOG categories

| | |
|---|---|
| **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 |

1