Assessment of the microbial diversity of Brazilian kefir grains by PCR-DGGE and pyrosequencing analysis

Leite, A.M.O.[1,2]; Mayo, B.[2], Rachid, C.T.C.C.[3]; Peixoto, R.S.[3]; Silva, J.T.[1]; Paschoalin, V.M.F.[1]; Delgado, S.[2*]

Universidade Federal do Rio de Janeiro, Instituto de Química, Departamento de Bioquímica, Avenida Athos da Silveira Ramos, 149-Bloco A, Sala 545. 21941-909-Cidade Universitária, Rio de Janeiro, RJ, Brazil[1], Instituto de Productos Lácteos de Asturias (IPLA-CSIC), Departamento de Microbiología y Bioquímica, Carretera de Infiesto s/n, 33300-Villaviciosa, Asturias, Spain[2], and Universidade Federal do Rio de Janeiro, Instituto de Microbiologia, Departamento de Microbiologia Geral, Avenida Carlos Chagas Filho, 373, 21941904-Cidade Universitária, Rio de Janeiro, RJ, Brazil[3].

Keywords: kefir grains, culture-independent microbiology, PCR-DGGE, pyrosequencing, *Lactobacillus*

Short title: Culture-independent microbiology of kefir grains

[*]Correspondence to: Susana Delgado, Instituto de Productos Lácteos de Asturias (IPLA-CSIC), Departamento de Microbiología y Bioquímica, Carretera de Infiesto s/n, 33300-Villaviciosa, Asturias, Spain, Phone/fax number: +34 985 892131

e-mail: sdelgado@ipla.csic.es

Abstract

This study evaluated the microbial diversity and community structure of three different kefir grains collected in different regions of Brazil, by combining two culture-independent methods: PCR-DGGE and barcode pyrosequencing. The DGGE analysis showed that the dominant bacterial populations in all three grains were similar and composed of two *Lactobacillus* species: *Lactobacillus kefiranofaciens* and *Lactobacillus kefiri*. The yeast community was dominated by *Saccharomyces cerevisiae,* which was present in all three samples. A total of 14,314 partial 16S rDNA sequence reads were obtained from the three grains by pyrosequencing. Sequence analysis grouped the reads into three phyla, of which *Firmicutes* was the most abundant. Members of the genus *Lactobacillus* were predominant operational taxonomic units (OTUs) in all samples, comprising up to 96% of the sequences. At low levels, OTUs belonging to other lactic-acid bacteria species and members of different phyla were also found. Two of the grains showed identical DGGE profiles and a similar number of OTUs, while the third sample showed the highest diversity by both techniques. The pyrosequencing approach allowed the identification of bacteria that were present in low numbers and are rarely associated with the microbial community of this complex ecosystem.

1 Introduction

Kefir is a viscous, acidic, and mildly alcoholic milk beverage produced by fermentation of milk with a kefir grain as the starter culture (FAO/WHO, 2003). Thought to be native to the Caucasus and Middle East regions, production and consumption of kefir has now spread throughout the world, led by a long history of beneficial health effects (Farnworth, 2005). Kefir grains are cauliflower-like florets of white to yellowish-white color, composed of an inert polysaccharide/protein matrix in which a relatively stable and specific microbial community composed of different lactic acid bacteria (LAB), acetic acid bacteria (AAB) and yeast species coexists in a complex symbiotic relationship (Farnworth, 2005). Kefir grains are supposed to have developed spontaneously in milk stored in animal-based containers made from skins, intestines or bladders. Kefir grains may have arisen independently at different locations, giving rise to grain-specific microbial populations, which produce beverages with distinctive sensory properties (Rea et al., 1996). Therefore, analysis of different kefir grains is of key importance to characterize the microbes of the grain ecosystem and to correlate the populations with sensory profiles.

The microbial diversity of kefir has traditionally been assessed by culture methods, by which different LAB species have been identified. A wide variety of *Lactobacillus* species have been isolated from both the beverage and the grains, including *Lactobacillus kefiri*, *Lactobacillus kefiranofaciens*, *Lactobacillus kefirgranum,* and *Lactobacillus parakefiri*, which constitute dominant populations (Rea et al., 1996; Kuo and Lin, 1999; Garrote et al., 2001; Simova et al., 2002). Often reported are *Lactococcus lactis* subsp. *lactis* and *Lactococcus lactis* subsp. *cremoris*, which are thought to be loosely associated with the grains and responsible for acidification. Both culturing and culture-independent techniques have identified *Lc. lactis* as dominant in

the fermented product (Simova et al., 2002; Chen et al., 2008; Dobson et al., 2011). *Leuconostoc* and other *Lactobacillus* species have been isolated in low numbers (Simova et al., 2003; Mainville et al., 2006). AAB have received less attention, although they are presumed to be essential in both the microbial consortium and the organoleptic characteristics of the final product (Rea et al., 1996). Among the yeasts, *Kluyveromyces marxianus, Torulaspora delbrueckii, Saccharomyces cerevisiae, Candida kefir, Saccharomyces unisporus, Pichia fermentans* and *Yarrowia lipolytica* have all been detected (Simova et al., 2002; Wang et al., 2008).

Culturing methods have proved to be unreliable for a complete microbial characterization of different ecosystems, including those of food fermentation (Giraffa and Neviani, 2001; Jany and Barbier, 2008). Some culture-independent microbial techniques, such as denaturing gradient gel electrophoresis (DGGE) (Wang et al., 2006; Chen et al., 2008) and construction and analysis of libraries of conserved genes such as the 16S rRNA gene (Ninane et al., 2007), have been applied to the microbial study of kefir grains. By means of these techniques, most cultured species have been detected, together with previously undetected microorganisms. However, in spite of this extensive knowledge, the inventory of the microbial species associated with the kefir grains is thought to be far from complete.

Pyrosequencing, an automated high-throughput parallel sequencing technique, which involves the synthesis of single-stranded deoxyribonucleic acid and the detection of the light generated by the pyrophosphate released through a coupled reaction with luciferase (Margulies et al., 2005), has recently begun to be applied to the study of food fermentation (Humblot and Guyot, 2009; Roh et al., 2010; Jung et al., 2011). This technique enables a rapid and accurate analysis of nucleotide sequences, which can be used to analyze the population structure, gene content, and metabolic potential of the

microbial communities in an ecosystem. Pyrosequencing has recently been applied to study the diversity and dynamics of the bacterial populations of an Irish kefir grain and its corresponding fermented product (Dobson et al., 2011).

This study characterized the microbial diversity of three different kefir grains collected in different regions of Brazil, by two culture-independent microbial methods: PCR-DGGE and barcode pyrosequencing. Here we report on the catalog of the microbial species identified by these two techniques, and compare them to those reported in the literature.

2 Material and Methods

2.1 Kefir grain samples

The three kefir grains utilized in this study were collected from different cities in southeastern Brazil (AR, Niterói, Rio de Janeiro; AV, Viçosa, Minas Gerais; and AD, Lavras, Minas Gerais). Grains were activated in sterile reconstituted skim milk (10% w/v) at 25ºC for 24 h, filtered to remove the clotted milk, and rinsed with sterile water. This activation step was repeated three times.

2.2 Isolation of total microbial DNA

For microbial genomic DNA extraction, activated kefir grains were homogenized in 2% sodium citrate, and 2 ml of each homogenate was centrifuged for 10 min at 10,000 *g*. Total DNA from the pellets was extracted and purified using a FastDNA Spin kit (QBIOgene, Carlsbad, CA, USA) according to the manufacturer's instructions. The DNA obtained was quantified using a Qubit flourometer apparatus (Invitrogen Detection Technologies, Eugene, OR, USA).

2.3 DGGE analysis of kefir grains

2.3.1 PCR amplification of 16S and 26S rDNA sequences

Genomic DNA was used as a template in PCR amplifications of the V3 region of the bacterial 16S rRNA gene, using the universal primers F357-GC (5'–TACGGGAGGCAGCAG–3' and R518 (5'–ATTACCGCGGCTGCTGG–3'), as reported by Muyzer *et al*. (1993). Group-specific primers for the detection of LAB were also used. These were the primer pair Lac1 (5'–AGCAGTAGGGAATCTTCCA–3') and Lac2-GC (5'–GATTYCACCGCTACACATG–3') to detect members of the genera *Lactobacillus*, *Pediococcus*, *Leuconostoc* and *Weissella* (Walter et al., 2001), and primers Lac3 (5'–AGCAGTAGGGAATCTTCGG–3') and Lac2-GC to detect members of the genera *Lactococcus*, *Streptococcus*, *Enterococcus*, *Tetragenococcus* and *Vagococcus* (Endo and Okada, 2005). The D1 domain of the 26S rRNA gene of fungi was amplified using the primers NL1-GC (5'–GCCATATCAATAAGCGGAGGAAAG–3') and LS2 (5'–ATTCCCAAACAACTCGACTC–3'), as reported by Cocolin *et al*. (2002). All GC primers contained a 39 bp GC clamp sequence at their 5' end to prevent complete denaturation of amplicons. PCR was performed in 50 µl reaction volumes using a Taq-DNA polymerase master mix (Ampliqon, Skovlunde, Denmark) with ~100 ng of each DNA sample as a template and 0.2 mM of each primer.

2.3.2 Electrophoretic conditions and identification of bands

DGGE was performed by using a DCode apparatus (Bio-Rad, Richmond, CA, USA) at 60°C and employing 8% polyacrylamide gels with a denaturing range of 40-60% for total bacteria, 40-50% for group-specific LAB and 30-50% for fungi. Electrophoresis was performed at 75 V for 16 h and 130 V for 4.5 h for bacterial and fungal amplifications, respectively. Bands were visualized under UV light after staining with ethidium bromide (0.5 µg ml^{-1}) and photographed.

In addition, all bands in the gels were identified by sequencing. For this purpose, bands were excised from the acrylamide gels and DNA was eluted overnight in 50 µl of sterile water at 4ºC. The DNA was re-amplified with the same primer pair without the GC-clamp, and sequenced by cycle extension in an ABI 373 DNA sequencer (Applied Biosystems, Foster City, CA, USA). The identity of the sequences was determined by the BLASTN algorithm in the GenBank database (http://www.ncbi.nlm.nih.gov/BLAST/).

2.4 Pyrosequencing analysis of kefir grains

2.4.1 Primers and 16S rRNA gene amplification conditions

Two universal primers, Y1 (5'–TGGCTCAGGACGAACGCTGGCGGC–3') (position 20-43 on 16S rRNA gene, *Escherichia coli* numbering) and Y2 (5'–CCTACTGCTGCCTCCCGTAGGAGT–3') (positions 361-338) (Young et al., 1991), were used to amplify by PCR a 348-bp stretch of DNA embracing the V1 and V2 variable regions of the prokaryotic 16S rDNA. 454-adaptors were included in both forward (5'–CGTATCGCCTCCCTCGCGCCATCAG–3') and reverse (5'–CTATGCGCCTTGCCAGCCCGCTCAG–3') primers, followed by a 10-bp sample-specific barcode sequence.

Amplifications were carried out as described above, using the following PCR conditions: 95°C for 5 min, 25 cycles of 94°C for 30 s, 52°C for 40 s and 72°C for 30 s, and a final extension step at 72°C for 10 min.

Amplicons were purified through GenElute^{TM} PCR Clean-Up columns (Sigma-Aldrich, St. Louis, MO, USA), and DNA concentration and quality was measured using an Epoch micro-volume spectrophotometer system (BioTek Instruments, Winooski, VT, USA). Equal amounts of the three samples were pooled, for a total amount of 100 ng. Pooled DNA was subsequently amplified in PCR-mixture-oil emulsions and sequenced

167 in different lanes of a PicoTiterPlate on a 454 Genome Sequencer 20 system (Roche,
168 Basel, Switzerland). The sequences obtained were uploaded and are available at the
169 NCBI Sequence Read Archive (SRA) under accession numbers SRA045648.2,
170 SRR340042.2, SRR340043.1 and SRR340041.1.

171 **2.4.2 Sequence treatment and bioinformatics analysis**

172 Raw sequences were processed through the Ribosomal Database Project (RDP)
173 pyrosequencing pipeline (http://wildpigeon.cme.msu.edu/pyro/index.jsp). Sequences
174 were excluded from the analysis if they had low quality, if the read length was less than
175 300 bp, or if one of the primer sequences was missing. The high-quality partial 16S
176 rDNA sequences were submitted to the RDP-II classifier using an 80% confidence
177 threshold, to obtain the taxonomic assignment and the relative abundance of the
178 different bacterial groups, as reported elsewhere (Wang et al., 2007). Multiple sequence
179 alignments for each sample were made by the Aligner tool in the RDP website (with the
180 default parameters). These alignments served as inputs for MOTHUR v. 1.14.0 software
181 (Schloss et al., 2009) to construct the distance matrix and for clustering the sequences
182 into operational taxonomic units (OTUs). The clusters were constructed at a 3%
183 dissimilarity cutoff and served as OTUs for generating predictive rarefaction models
184 and for making calculations with the richness indices Ace and Chao1 (Chao and Bunge,
185 2002) and the Shannon diversity index (Shannon and Weaver, 1949). The MOTHUR
186 program was also used to perform the Fast UniFrac test, which was used to compare the
187 phylogenetic structure of the libraries, and to generate the Venn diagrams. A neighbor-
188 joining tree was constructed with representative sequences of each OTU selected by
189 MOTHUR. These sequences were compared against the RDP database by using the
190 Seqmatch option to select for the nearest neighbors. All sequences were then aligned
191 using MEGA 5.0 sofware (Tamura et al., 2011) and the Jukes-Cantor model. The

192 equivalent sequence of the archaea *Halococcus saccharolyticus* (AB004876) was used
193 as an outgroup to root the tree.

194 **3 Results**

195 **3.1 PCR-DGGE analysis of bacterial and yeast communities**

196 PCR-DGGE analyses of 16S and 26S rRNA genes with universal primers were
197 conducted to obtain an overview of the community structure of the dominant bacterial
198 and fungal populations of the Brazilian kefir grains. Fingerprints of the microbial
199 communities were rather simple, as they contained one to five different bands (Fig. 1,
200 panels A through D). Most bands were shared among all three kefir samples. Individual
201 bands of both bacterial and fungal populations were sequenced and identified by
202 sequence comparison, and all of them showed 99-100% similarity with sequences in the
203 GenBank database. The species profile of the total bacteria as amplified with universal
204 primers was composed of up to five bands, but corresponded to only three different
205 species (Fig. 1, panel A). Bands corresponding to *Lb. kefiranofaciens* (bands 1, 2 and 5)
206 and to *Lb. kefiri* (band 4) were found in all samples. An additional band present in
207 sample AV (band 3) was identified as *Lc. lactis*. The same three species were also found
208 by using the group-specific primers for lactobacilli and lactococci (Figure 1, panels C
209 and D, respectively). The DGGE fingerprints of the yeast community were also narrow
210 and similar in the three kefir grains. A high-intensity band was present in all kefir
211 samples and was identified as *S. cerevisiae* (band 6, Fig. 1 panel B), while a low-
212 intensity band corresponding to *Kazachstania unispora* was revealed in kefir grain AD
213 (band 7, Fig. 1 panel B).

214 **3.2 Bacterial composition and community structure determined by pyrosequencing**

215 A total of 25,127 raw reads were obtained by pyrosequencing analysis, including
216 5,172 reads from sample AD, 4,651 from sample AR and 15,304 from sample AV. Of

these, a total of 14,314 corresponded to high-quality partial 16S rDNA sequences longer than 300 bp of samples AD (2,641 reads), AR (2,690 reads), and AV (8,983 reads). A comparative analysis was performed to assess whether the exclusion of low-quality fragments could influence the results. Comparison of the graphs and indexes of the classifier tool showed similar results, with no loss or difference in the proportion of phyla, families or genera (data not shown). Therefore, because much information could be obtained from the longer reads, all subsequent analyses were done with the selected, long reads.

Diversity richness, coverage, and evenness estimates calculated for each data set are presented in Table 1. Rarefaction curves showed similar patterns for all samples (Fig. 2), and suggested that the bacterial community was well represented, as they became flatter while the number of sequences analyzed increased. Additionally, when re-sampling analyses were performed, normalizing by sample size to that of the smallest one, the rarefaction curves proved to be saturated (Fig. 2 panel B). Moreover, the coverage at the 97% similarity level was above 0.99 for each of the kefir grains. According to Figure 2 and the OTU richness estimated by ACE and Chao 1 indexes at the 97% similarity level (Table 1), sample AV had higher species richness than the other two grains. Considering the microbial diversity estimated by the Shannon index at the 97% similarity level gave a similar result. Indeed, 14, 18, and 46 OTUs were associated with kefir samples AR, AD, and AV, respectively (Table 1).

The Unifrac test was used to compare the bacterial communities based on their phylogenetic information. This analysis also revealed that sample AV was significantly different from AD and AR ($p < 0.01$), when the relative proportion of sequences from each community was considered (Weighted Unifrac algorithm).

To evaluate the distribution of OTUs between the different kefir grains, a Venn diagram was constructed (Fig. 3). The diagram showed that 11 OTUs, embracing 95.8% of the sequences, were common to all three grains. Furthermore, despite the higher number of specific OTUs in the AV sample (24 OTUs), the occurrence of these grain-specific sequences (3.86%) was much lower than those shared by all samples (95.8%). Similarly, specific OTUs of the other two samples were represented by a low percentage of sequences.

The bacterial sequence reads were grouped into three different phyla: *Firmicutes,* *Actinobacteria*, and *Proteobacteria*. Of these, *Firmicutes* was the most abundant phylum, and was dominated by members of the class *Bacilli* belonging to the order *Lactobacillales.* Three families were found among the sequences belonging to this order: *Leuconostocaceae*, *Streptococcaceae,* and *Lactobacillaceae*. The family *Lactobacillaceae* was predominant in all three grains, and was represented by only one genus, *Lactobacillus,* which comprised 99.7, 93.9, and 99.6% of the reads for grains AR, AV, and AD, respectively (Fig. 4). In the family *Streptococcaceae*, the genus *Streptococcus* comprised only 0.01% and 0.04% of all sequences identified in grains AV and AD, respectively, whereas the genus *Lactococcus* was detected only in kefir grain AV (4.87% of the reads). At low levels, the genus *Leuconostoc* also occurred in samples AV (0.12%) and AD (0.23%). Few sequences were assigned to the phylum *Proteobacteria*, which comprised 0.3% of the total assigned sequences for grain AR, 1% for AV and 0.04% for AD. The sequences of this phylum belonged to the genus *Acetobacter* in sample AR (0.26%) and AD (0.04%), and to the genus *Pseudomonas* (0.99%) in sample AV. Phylum *Actinobacteria* was represented by reads belonging to the genus *Solirubrobacter* in grain AR (0.04%) and the genus *Bifidobacterium* in grain AV (0.02%).

Because of the low diversity found, unique representative sequences from each OTU were selected and used to construct a phylogenetic tree (Fig. 5). The different sequences were manually compared against the RDP database and further aligned with up to three of their nearest sequences in the database. The majority of the OTUs represented close phylogenetic lineages of *Lactobacillus* spp. commonly reported in kefir grains. These alignments and manual investigations further allowed the classification of the reads in a number of *Lactobacillus* species and subspecies, including among others *Lb. kefiranofaciens* subsp. *kefirgranum, Lb. kefiri, Lb. parabuchneri, Lb. parakefiri, Lb. amilovorus, Lb. crispatus, Lb. buchneri,* and *Lb. kefiranofaciens* subsp. *kefiranofaciens*. Sequences identified as *Lc. lactis* subsp. *cremoris* were revealed in kefir sample AV.

4 Discussion

The microbial diversity of kefir grains from different origins has been repeatedly analyzed by both culturing (Simova et al., 2002; Witthuhn et al., 2005; Mainville et al., 2006; Chen et al., 2008; Wang et al., 2008; Miguel et al., 2010) and culture-independent techniques (Garbers et al., 2004; Wang et al., 2006; Ninane et al., 2007; Chen et al., 2008; Wang et al., 2008; Miguel et al., 2010; Dobson et al., 2010). In this study, two independent techniques were used to evaluate the microbial diversity and community structure of three different kefir grains from different locations in Brazil. Dominant populations were tracked with the PCR-DGGE technique, while the next-generation sequencing technology allowed a more complete view of the overall community composition.

As in previous studies (Garbers et al., 2004; Chen et al., 2008; Jianzhong et al., 2009; Miguel et al., 2010), bacterial PCR-DGGE profiles were shown to be composed of a small number of bands. These corresponded to several *Lactobacillus* species that

have always been reported as prevalent, although the species dominating the grains seems to vary. *Lb. kefiranofaciens* (Chen et al., 2008; Jianzhong et al., 2009), *Lb. kefiri* (Miguel et al., 2010), and *Lb. casei* (Jianzhong et al., 2009) have all been described as accounting for the more intense DGGE bands. A small number of DGGE bands in the yeast profile has also been reported for many other kefir grains (Garbers et al., 2004; Wang et al., 2008; Jianzhong et al., 2009). The dominant yeasts belonged to a short list of species: *Saccharomyces* spp., *Kluyveromyces lactis, Kazachtania* spp., and *Candida* spp. (Garbers et al., 2004; Wang et al., 2008; Jianzhong et al., 2009). From the DGGE results, we concluded that the Brazilian kefir grains examined here were dominated by *Lb. kefiranofaciens*, followed by *Lb. kefiri*. These two bacterial species have also been reported as dominant by culturing in different kefir grains (Mainville et al., 2006; Chen et al., 2008; Miguel et al., 2010). *S. cerevisiae* was the main species among the yeasts. This and other related species have also been identified as a majority by culturing (Simova et al., 2002; Latorre-García et al., 2007).

Nowadays, pyrosequencing is becoming the state-of-the-art technique for the analysis of microbial populations from different ecosystems. It has been applied to study several types of food fermentation (Humblot and Guyot, 2009; Roh et al., 2010; Jung et al., 2011), including a single report of kefir in which the kefir grain and its fermented milk were analyzed by this technique (Dobson et al., 2011). The pyrosequencing analysis of the three Brazilian kefir grains revealed that the phylum *Firmicutes* was highly dominant, comprising more than 99% of the total sequences. This phylum is composed by a group of low-GC-content Gram-positive bacteria, which includes LAB. *Firmicutes* was also dominant in the study of Irish Kefir milk, which analyzed both the interior and exterior of the grain (Dobson et al., 2011). These authors also showed that all other phyla that they detected (*Actinobacteria*, *Proteobacteria*, and

316 *Bacteriodetes*) were minor components of the overall kefir community in the interior
317 part of the Irish kefir grain. Within the phylum *Proteobacteria*, *Pseudomonas* spp. was
318 identified in the grain AV, which has been suggested to be an environmental
319 contamination (Dobson et al., 2011). The genus *Acetobacter* (*Proteobacteria* subgroup)
320 was found in only two of the Brazilian grains (AR and AD). Although AAB have often
321 been mentioned (Rea et al., 1996; Garrote et al., 2001; Miguel et al., 2010) as one of the
322 main components that comprise the bacterial population of kefir grains, AAB have been
323 detected only occasionally (Garbers et al., 2004; Chen et al., 2008; Jianzhong et al.,
324 2009; Miguel et al., 2010; Dobson et al., 2011).

325 Phylogenetic and manual analysis showed that *Lb. kefiranofaciens* subsp.
326 *kefirgranum* was dominant among the reads. Reads assigned to *Lb. kefiri* ranked second,
327 although much lower than the number of those of *Lb. kefiranofaciens*. These results
328 completely agree with those obtained by the DGGE technique. The presence of reads
329 belonging to *Lc. lactis* subsp. *cremoris* in kefir sample AV further validates the DGGE
330 results. In general, the two techniques were consistent with respect to detection of the
331 predominant bacteria. However, some microorganisms identified by pyrosequencing
332 were not detected by DGGE analysis, probably because they were part of minority
333 populations in the grains. This limitation of the PCR-DGGE method was previously
334 noted by Ercolini (2004), who reported that minor bacterial groups in complex
335 communities may not be represented in the DGGE profiles. As seen in this study, the
336 use of pyrosequencing can allow the detection of rare microorganisms that are not part
337 of the dominant community.

338 As expected, the number of OTUs was lower than those found in other complex
339 ecosystems such as soil (Teixeira et al., 2010) and the human gastrointestinal tract
340 (Turnbaugh et al., 2009). The bacterial simplicity of the kefir grains is further revealed

by the Venn diagrams, where a few, highly prevalent species are shared by all grains, together with a small number of minor bacteria that are specific for each grain. As already discussed, traditional culturing and molecular techniques indicated that a few specific microbial genera and species may be constantly present in the kefir grain, whereas others may or may not occur (Simova et al., 2002; Witthuhn et al., 2005; Mainville et al., 2006; Wang et al., 2006; Ninane et al., 2007; Chen et al., 2008; Wang et al., 2008; Miguel et al., 2010; Dobson et al., 2011). Furthermore, as Farnworth and Mainville (2008) have recently noted, the list of bacteria and yeasts of kefir grains should not vary significantly from one part of the world to another if good care, similar growth conditions, and proper sanitary conditions are maintained. However, these small microbial differences may produce distinctive, grain-specific sensory profiles (Pintado et al., 1996; Rea et al., 1996; Simova et al., 2002).

5 Conclusions

Two culture-independent methods were used to evaluate the microbial diversity of three Brazilian kefir grains: PCR-DGGE and pyrosequencing. Both techniques showed that *Lb. kefiranofaciens* was dominant, while DGGE showed that *S. cerevisiae* constituted the main eukaryotic microorganism. The pyrosequencing analysis also allowed the identification of minor bacterial components. For a complete description of the microbial communities of the kefir grains, a pyrosequencing approach using specific primers for eukaryotic and archaea organisms should also be performed.

Acknowledgements

The study was supported by the CAPES Foundation (process number PDEE 5019109) and a project from the Spanish Ministry of Science and Innovation (MCINN) (reference AGL2007-61869-ALI). S. Delgado was supported by a research contract from MICINN under the Juan de la Cierva Program (reference JCI-2008-02391).

References

Chao, A., & Bunge, J. (2002). Estimating the number of species in a stochastic abundance model. *Biometrics, 58,* 531-539.

Chen, H.C., Wang, S.Y., & Chen, M.-J. (2008). Microbiological study of lactic acid bacteria in kefir grains by culture-dependent and culture-independent methods. *Food Microbiology, 25,* 492-501.

Cocolin, L., Aggio, D., Manzano, M., Cantoni, C., & Comi, G. (2002). An application of PCR-DGGE analysis to profile the yeast populations in raw milk. *International Dairy Journal, 12*, 407-411.

Dobson, A., O'sullivan, O., Cotter, P.D., Ross, P., & Hill, C. (2011). High-throughput sequence-based analysis of the bacterial composition of kefir and an associated kefir grain. *FEMS Microbiology Letters, 320,* 56-62.

Endo, A., & Okada, S. (2005). Monitoring the lactic acid bacterial diversity during shochu fermentation by PCR-denaturing gradient gel electrophoresis. *Journal of Bioscience and Bioengineering, 99,* 216-221.

Ercolini, D. (2004). PCR-DGGE fingerprinting: novel strategies for detection of microbes in food. *Journal of Microbiological Methods, 56,* 297-314.

FAO/WHO. (2003). CODEX Standard for Fermented Milks. *Codex Stan 243-2003*. Reviewed 2010, 2nd ed. URL http://www.codexalimentarius.net/download/standards/400/CXS_243e.pdf.

Farnworth, E.R. (2005). Kefir –a complex probiotic. *Food Science and Technology Bulletin: Functional Foods, 2,* 1-17.

Farnworth, E.R., & Mainville, I. (2008). Kefir – A Fermented Milk Product. In E.R., Farnworth (Ed.), *Handbook of fermented functional foods* (2nd ed., pp. 89-127). Boca Raton, FL: CRC Press Taylor & Francis Group.

391 Garbers, I-M., Britz, T.J., & Witthuhn, R.C. (2004). PCR-based denaturing
392 gradient gel electrophoretic typification and identification of the microbial consortium
393 present in kefir grains. *World Journal of Microbiology and Biotechnology, 20,* 687-693.
394 Garrote, G.L., Abraham, A.G., & de Antoni, G.L. (2001). Chemical and
395 microbiological characterisation of kefir grains. *Journal of Dairy Research*, 68, 639-
396 652.
397 Giraffa, G., & Neviani, E. (2001). DNA-based, culture-independent strategies
398 for evaluating microbial communities in food-associated ecosystems. *International*
399 *Journal of Food Microbiology, 67,* 19-34.
400 Humblot, C., & Guyot, J.P. (2009). Pyrosequencing of tagged 16S rRNA gene
401 amplicons for rapid deciphering of the microbiomes of fermented foods such as pearl
402 millet slurries. *Applied and Environmental Microbiology, 75*, 4354-4361.
403 Jany, J.L., & Barbier, G. (2008). Culture-independent methods for identifying
404 microbial communities in cheese. *Food Microbiology, 25*, 839-848.
405 Jianzhong, Z., Xiaoli, L., Hanhu, J., & Mingsheng, D. (2009). Analysis of the
406 microflora in Tibetan kefir grains using denaturing gradient gel electrophoresis. *Food*
407 *Microbiology, 26*, 770-775.
408 Jung, J.Y., Lee, S.H., Kim, J.M., Park, M.S., Bae, J.W., Hahn, Y., Madsen, E.L.,
409 & Jeon, C.O. (2011). Metagenomic analysis of kimchi, a traditional korean fermented
410 food. *Applied and Environmental Microbiology*, *77*, 2264-2274.
411 Kuo, C-Y., & Lin, C.W. (1999). Taiwanese kefir grains: their growth, microbial
412 and chemical composition of fermented milk. *Australian Journal of Dairy Technology,*
413 *54*, 19-23.

Latorre-García, L., Castillo-Agudo, L., & Polaina, J. (2007). Taxonomical classification of yeasts isolated from kefir based on the sequence of their ribosomal RNA genes. *World Journal of Microbiology and Biotechnology, 23*, 785-791.

Mainville, I., Robert, N., Lee, B., & Farnworth, E.R. (2006). Polyphasic characterization of the lactic acid bacteria in kefir. *Systematic and Applied Microbiology, 29*, 59-68.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman. M.S., Chen, Y.J., Chen, Z., Dewell, S.B., Du, L., Fierro, J.M., Gomes, X.V., Godwin, B.C., He ,W., Helgesen, S., Ho, C.H., Irzyk, G.P., Jando, S.C., Alenquer, M.L., Jarvie, T.P., Jirage, K.B., Kim, J.B., Knight, J.R., Lanza, J.R., Leamon, J.H., Lefkowitz, S.M., Lei, M., Li, J., Lohman, K.L., Lu, H., Makhijani, V.B., McDade, K.E., McKenna, M.P., Myers, E.W., Nickerson, E., Nobile, J.R., Plant, R., Puc, B.P., Ronan, M.T., Roth, G.T., Sarkis, G.J., Simons, J.F., Simpson, J.W., Srinivasan, M., Tartaro, K.R., Tomasz, A., Vogt, K.A., Volkmer, G.A., Wang, S.H., Wang, Y., Weiner, M.P., Yu, P., Begley, R.F., & Rothberg, J.M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature, 437*, 376-380.

Miguel, M.G.C.P., Cardoso, P.G., Lago, L.A., & Schwan, R.F. (2010). Diversity of bacteria present in milk kefir grains using culture-dependent and culture-independent methods. *Food Research International, 42*, 1523-1528.

Muyzer, G., De Waal, E.C., & Uitterlinden, A.G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Applied and Environmental Microbiology, 59*, 695–700.

Ninane, V., Mukandayambaje, R., & Berben, G. (2007). Identification of lactic acid bacteria within the consortium of a kefir grain by sequencing of 16S rDNA variable regions. *Journal of AOAC International, 90,* 1111-1117.

Pintado, M.E., Lopes Da Silva, J.A., Fernandes, P.B., Malcata, F.X., & Hogg, T.A. (1996). Microbiological and rheological studies on Portuguese kefir grains. *International Journal of Food Science and Technology, 31,* 15–26.

Rea, M.C., Lennartsson, T., Dillon, P., Drinan, F.D., Reville, W.J., Heapes, M., & Cogan, T.M. (1996). Irish kefir-like grains: their structure, microbial composition and fermentation kinetics. *Journal of Applied Bacteriology, 81,* 83-94.

Roh, S.W., Kim, K.H., Nam, Y.D., Chang, H.W., Park, E.J., & Bae, J.W. (2010). Investigation of archaeal and bacterial diversity in fermented seafood using barcoded pyrosequencing. *The ISME Journal, 4,* 1-16.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., & Weber, C.F. (2009). Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology, 75,* 7537-7541.

Shannon, C.E., & Weaver, W. (1949). The mathematical theory of information. *AT&T Technical Journal, 27,* 359-423.

Simova, E., Beshkova, D., Angelov, A., Hristozova, T., Frengova, G., & Spasov, Z. (2002). Lactic acid bacteria and yeasts in kefir grains and kefir made from them. *Journal of Industrial Microbiology and Biotechnology, 28,* 1-6.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood,

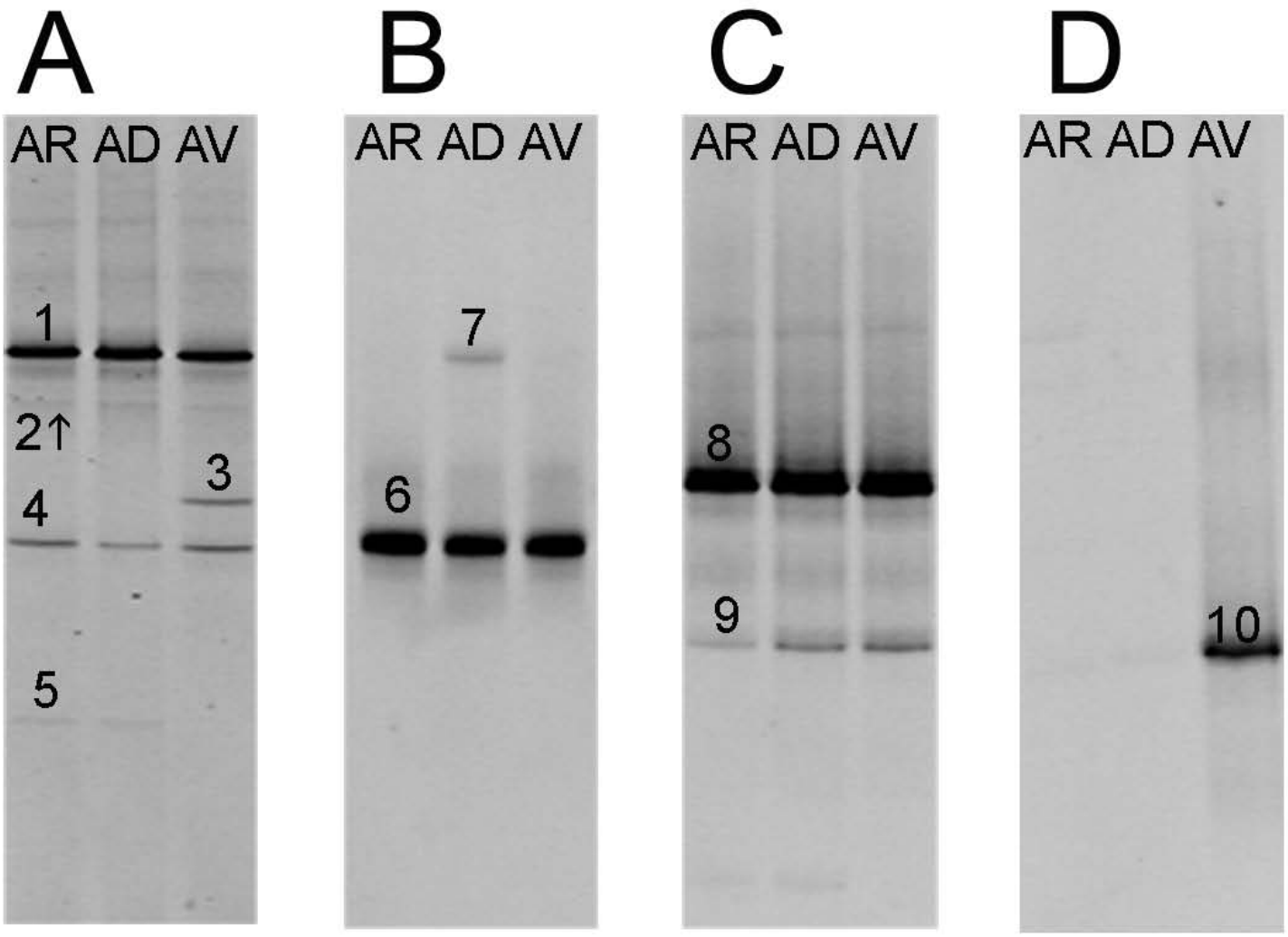
462 evolutionary distance, and maximum parsimony methods. *Molecular Biology and*
463 *Evolution*, doi:10.1093/molbev/msr121.
464 Teixeira, L.C.R.S., Peixoto, R.S., Cury, J.C., Sul, W.J., Pellizari, V.H., Tiedje, J.,
465 & Rosado, A. (2010). Bacterial diversity in rhizosphere soil from Antartic vascular
466 plants of Admiralty Bay, maritime Antarctica. *The ISME Journal*, *4*, 989-1001.
467 Turnbaugh, P.J., Hamady, M., Yatsunenko, T., Cantarel, B.L., Duncan, A., Ley,
468 R.E., Sogin, M.L., Jones, W.J., Roe, B.A., Affourtit, J.P., Egholm, M., Henrissat, B.,
469 Heath, A.C., Knight, R., & Gordon, J.I. (2009). A core gut microbiome in obese and
470 lean twins. *Nature*, *457*, 480-484.
471 Walter, J., Hertel, C., Tannock, G.W., Lis, C.M., Munro, K., & Hammes, W.P.
472 (2001). Detection of *Lactobacillus*, *Pediococcus*, *Leuconostoc*, and *Weissella* species in
473 human feces by using group-specific PCR primers and denaturing gradient gel
474 electrophoresis. *Applied and Enviromental Microbiology*, *67*, 2578-2585.
475 Wang, Y.Y., Li, H.R., Jia, S.F., Wu, Z.F., & Guo, B.H. (2006). Analysis of
476 bacterial diversity of kefir grains by denaturing gradient gel electrophoresis and 16S
477 rDNA sequencing. *Wei Sheng Wu Xue Bao*, *46*, 310-313.
478 Wang, Q., Garrity, G.M., Tiedje, J.M., & Cole, J.R. (2007). Naïve Bayesian
479 classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy.
480 *Applied and Environmental Microbiology, 73*, 5261-5267.
481 Wang, S.Y., Chen, H.C., Liu, J.R., Lin, Y.C., & Chen, M.J. (2008). Identification
482 of yeasts and evaluation of their distribution in Taiwanese kefir and viili starters.
483 *Journal of Dairy Science, 91*, 3798-805.
484 Witthuhn, R.C., Schoeman, T., & Britz, T.J. (2005). Characterisation of the
485 microbial population at different stages of kefir production and kefir grain mass
486 cultivation. *International Dairy Journal*, *15*, 383-389.

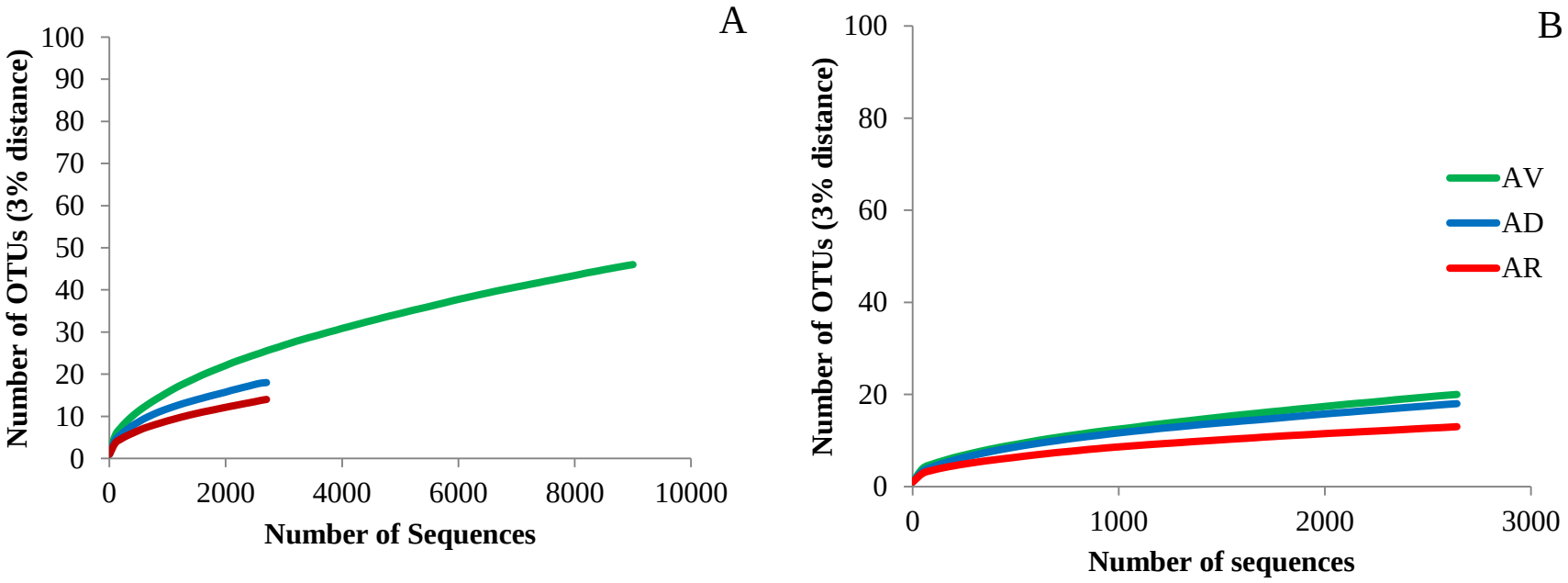
Young, J.P.W., Downer, H.L., & Eardly, B.D. (1991). Phylogeny of the phototrophic *Rhizobium* strain BTAil by polymerase chain reaction-based sequencing of a 16S rRNA gene segment. *Journal of Bacteriology, 173*, 2271-2277.

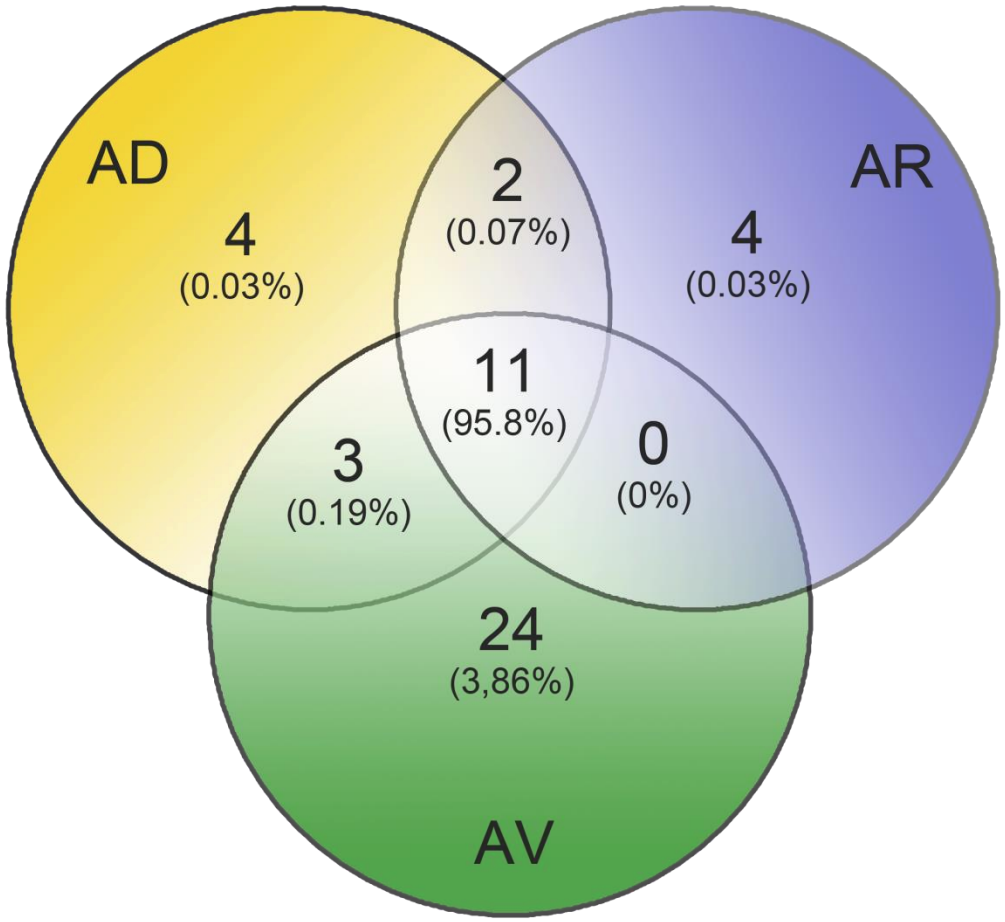
Table 1 - Estimated OTU richness, sample coverage and diversity index of 16S rDNA libraries of kefir grain samples.

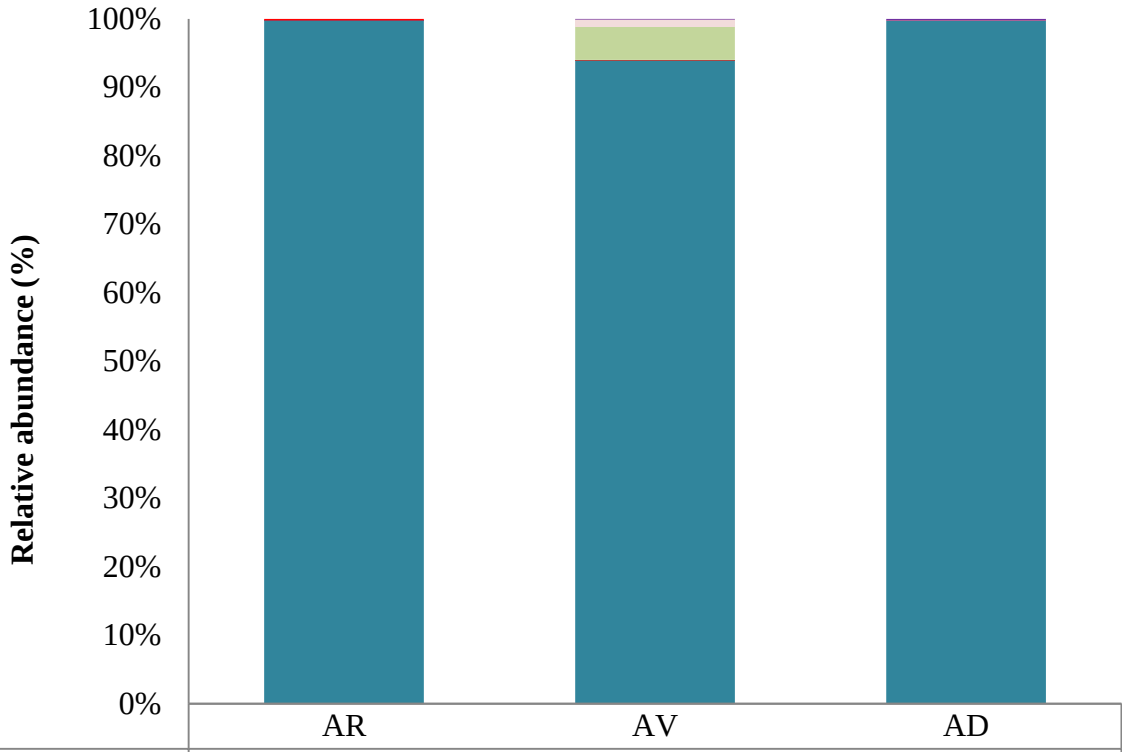
Library	NS	OTUs[a]	Estimated OTU richness		Shannon[b]	ESC[c]
			ACE	Chao1		
AD	2641	18	42.24 (28.17; 75.78)	54.00 (28.27; 144.19)	0.49 (0.45; 0.53)	0.99
AR	2690	14	38.57 (23.67; 76.40)	24.50 (16.03; 68.19)	0.33 (0.29; 0.37)	0.99
AV	8983	46	148.02 (109.96; 208.74)	82.14 (58.65; 149.23)	0.70 (0.67; 0.72)	0.99
Total	14314					

Abbreviations: ESC, estimated sample coverage; NS, number of sequences for each library; OTU, operational taxonomic unit.
[a] Calculated by MOTHUR at the 3% distance level.
[b] Shannon diversity index calculated using MOTHUR (3% distance).
[c] ESC: $C_x = 1 - (N_x/n)$, where N_x is the number of unique sequences and n is the total number of sequences.
Values in brackets are 95% confidence intervals as calculated by MOTHUR.


A
AR AD AV
1
2↑
3
4
5
B
AR AD AV
7
6
C
AR AD AV
8
9
D
AR AD AV
10


A
Number of OTUs (3% distance)
100
90
80
70
60
50
40
30
20
10
0
0
2000
4000
6000
8000
10000
Number of Sequences
B
Number of OTUs (3% distance)
100
80
60
40
20
0
0
1000
2000
3000
Number of sequences
AV
AD
AR


AD
4
(0.03%)
2
(0.07%)
AR
4
(0.03%)
11
(95.8%)
3
(0.19%)
0
(0%)
24
(3,86%)
AV



	AR	AV	AD
Leuconostocaceae	0,00%	0,12%	0,23%
Pseudomonadaceae	0,00%	1,00%	0,00%
Acetobacteraceae	0,26%	0,00%	0,04%
Streptococcaceae	0,00%	4,88%	0,04%
Bifidobacteriaceae	0,00%	0,02%	0,00%
Solirubrobacteraceae	0,04%	0,00%	0,00%
Lactobacillaceae	99,70%	93,93%	99,62%
Not assigned	0,00%	0,04%	0,08%

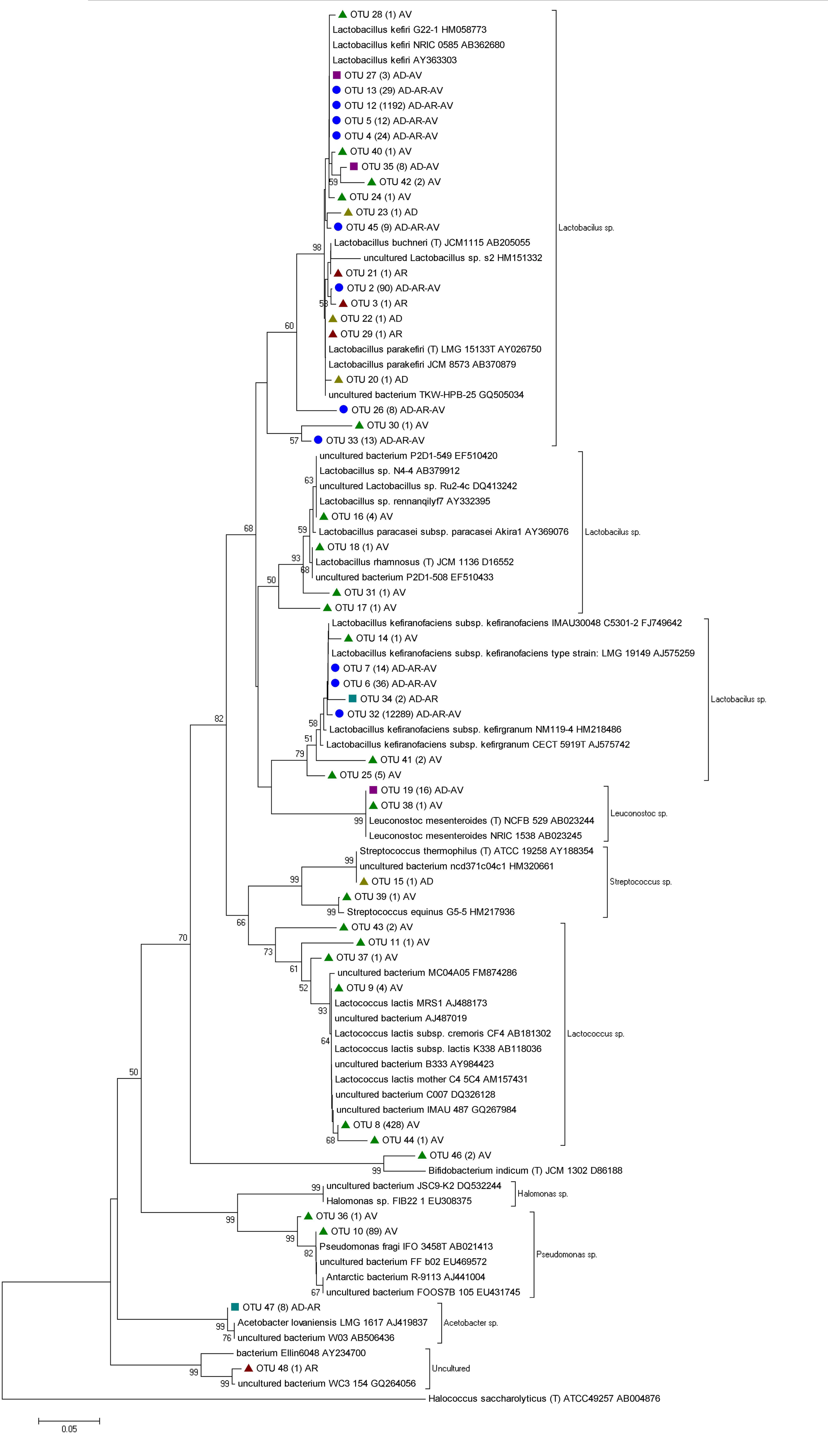

OTU 28 (1) AV
Lactobacillus kefiri G22-1 HM058773
Lactobacillus kefiri NRIC 0585 AB362680
Lactobacillus kefiri AY363303
OTU 27 (3) AD-AV
OTU 13 (29) AD-AR-AV
OTU 12 (1192) AD-AR-AV
OTU 5 (12) AD-AR-AV
OTU 4 (24) AD-AR-AV
OTU 40 (1) AV
OTU 35 (8) AD-AV
59
OTU 42 (2) AV
OTU 24 (1) AV
OTU 23 (1) AD
OTU 45 (9) AD-AR-AV
Lactobacilus sp.
98
Lactobacillus buchneri (T) JCM1115 AB205055
uncultured Lactobacillus sp. s2 HM151332
OTU 21 (1) AR
OTU 2 (90) AD-AR-AV
53
OTU 3 (1) AR
OTU 22 (1) AD
60
OTU 29 (1) AR
Lactobacillus parakefiri (T) LMG 15133T AY026750
Lactobacillus parakefiri JCM 8573 AB370879
OTU 20 (1) AD
uncultured bacterium TKW-HPB-25 GQ505034
OTU 26 (8) AD-AR-AV
OTU 30 (1) AV
57
OTU 33 (13) AD-AR-AV
uncultured bacterium P2D1-549 EF510420
Lactobacillus sp. N4-4 AB379912
63
uncultured Lactobacillus sp. Ru2-4c DQ413242
Lactobacillus sp. rennanqilyf7 AY332395
OTU 16 (4) AV
68
59
Lactobacillus paracasei subsp. paracasei Akira1 AY369076
Lactobacilus sp.
OTU 18 (1) AV
93
Lactobacillus rhamnosus (T) JCM 1136 D16552
68
uncultured bacterium P2D1-508 EF510433
50
OTU 31 (1) AV
OTU 17 (1) AV
Lactobacillus kefiranofaciens subsp. kefiranofaciens IMAU30048 C5301-2 FJ749642
OTU 14 (1) AV
Lactobacillus kefiranofaciens subsp. kefiranofaciens type strain: LMG 19149 AJ575259
OTU 7 (14) AD-AR-AV
OTU 6 (36) AD-AR-AV
OTU 34 (2) AD-AR
Lactobacilus sp.
82
OTU 32 (12289) AD-AR-AV
58
Lactobacillus kefiranofaciens subsp. kefirgranum NM119-4 HM218486
51
Lactobacillus kefiranofaciens subsp. kefirgranum CECT 5919T AJ575742
79
OTU 41 (2) AV
OTU 25 (5) AV
OTU 19 (16) AD-AV
OTU 38 (1) AV
99
Leuconostoc mesenteroides (T) NCFB 529 AB023244
Leuconostoc sp.
Leuconostoc mesenteroides NRIC 1538 AB023245
Streptococcus thermophilus (T) ATCC 19258 AY188354
99
uncultured bacterium ncd371c04c1 HM320661
99
OTU 15 (1) AD
OTU 39 (1) AV
99
Streptococcus equinus G5-5 HM217936
Streptococcus sp.
66
OTU 43 (2) AV
70
OTU 11 (1) AV
73
OTU 37 (1) AV
61
uncultured bacterium MC04A05 FM874286
52
OTU 9 (4) AV
Lactococcus lactis MRS1 AJ488173
93
uncultured bacterium AJ487019
Lactococcus lactis subsp. cremoris CF4 AB181302
64
Lactococcus sp.
Lactococcus lactis subsp. lactis K338 AB118036
uncultured bacterium B333 AY984423
50
Lactococcus lactis mother C4 5C4 AM157431
uncultured bacterium C007 DQ326128
uncultured bacterium IMAU 487 GQ267984
OTU 8 (428) AV
68
OTU 44 (1) AV
OTU 46 (2) AV
99
Bifidobacterium indicum (T) JCM 1302 D86188
99
uncultured bacterium JSC9-K2 DQ532244
Halomonas sp.
Halomonas sp. FIB22 1 EU308375
99
OTU 36 (1) AV
OTU 10 (89) AV
99
Pseudomonas fragi IFO 3458T AB021413
82
Pseudomonas sp.
uncultured bacterium FF b02 EU469572
Antarctic bacterium R-9113 AJ441004
67
uncultured bacterium FOOS7B 105 EU431745
OTU 47 (8) AD-AR
Acetobacter lovaniensis LMG 1617 AJ419837
Acetobacter sp.
99
76
uncultured bacterium W03 AB506436
bacterium Ellin6048 AY234700
Uncultured
OTU 48 (1) AR
99
99
uncultured bacterium WC3 154 GQ264056
Halococcus saccharolyticus (T) ATCC49257 AB004876
0.05

Figure legends

Fig. 1. DGGE profiles of the microbial community from three Brazilian kefir grains (samples AR, AD and AV). Panel A: DGGE profile of the eubacterial 16S rRNA gene obtained with universal primers (1) *Lactobacillus kefiranofaciens*; (2) *Lactobacillus kefiranofaciens*; (3) *Lactococcus lactis*; (4) *Lactobacillus kefiri*; (5) *Lactobacillus kefiranofaciens*. Panel B: DGGE profile of the eukaryotic domain D1 of 26S rRNA gene (6) *Saccharomyces cerevisiae*; (7) *Kazachstania unispora*. Panel C: DGGE profile of 16S rRNA gene obtained with specific primers for the lactobacilli group (8) *Lactobacillus kefiranofaciens*; (9) *Lactobacillus kefiri*. Panel D: DGGE profile of 16S rRNA gene obtained with specific primers for the lactococcus group (10) *Lactococcus lactis*.

Fig. 2. Rarefaction curves of partial sequences of the bacterial 16S rRNA gene from Brazilian kefir grains (AD, AR and AV) at a 97% similarity level (A) and rarefaction curves normalized with respect to sample size (B).

Fig. 3. Venn diagram showing specific and common OTUs in Brazilian kefir grains AD, AR and AV, and the percentage of occurrence of the total sequences (in parentheses).

Fig. 4. Relative abundances at family level, based on the classification of partial 16S rDNA sequences of bacteria from Brazilian kefir grains AD, AR and AV, using RDP-Classifier.

Fig. 5. Bacterial phylogenetic tree showing representative reads from the pyrosequencing analysis. The neighbor-joining tree was constructed with a representative sequence of each OTU selected by the MOTHUR program. Numbers at the nodes indicate bootstrap values (expressed as a percentage of 1000 replications). Values in brackets represent the number of sequences found for each OTU. Symbols and colors represent the sample group that contributed each OTU.