Molecular investigation of bacterial communities during the manufacturing and ripening of semi-hard Iranian Liqvan cheese

Molecular Investigation of Bacterial Communities during the Manufacturing and Ripening of Semi-hard Iranian Liqvan Cheese

M. Ramezani [a], S.M. Hosseini [a], I. Ferrocino [b], M.A. Amoozegar [c, d], L. Cocolin [b, *]

[a] Department of Microbiology, Faculty of Biological Sciences & Technologies, Shahid Beheshti University, Tehran, Iran
[b] University of Torino, Department of Agricultural, Forest and Food Science, Largo Paolo Braccini 2, 10095, Grugliasco, Torino, Italy
[c] Microorganisms Bank, Iranian Biological Resource Centre (IBRC), ACECR, Tehran, Iran
[d] Extremophiles Laboratory, Department of Microbiology, Faculty of Biology and Center of Excellence in Phylogeny of Living Organisms, College of Science, University of Tehran, Tehran, Iran

Abstract

Liqvan (or Lighvan) is a traditional Iranian cheese from the East Azerbaijan province of Iran, which is made of raw ewe´s milk without the addition of a starter. The grazing pastures, environmental conditions and the ancient regional production methods allocate a distinctive microbial ecology to this type of cheese, and these factors are consequently associated with the quality of the product. In this study, the microbiota of the milk, curd and cheese has been investigated using culture independent approaches. Denaturing gradient gel electrophoresis (DGGE) of the bacteria, 16S rRNA based high-throughput sequencing and enumeration of the live bacterial community by means of quantitative PCR (qPCR) have been used for this purpose. The results showed that the main bacterial population in the milk belonged to both microbial contaminants and lactic acid bacteria (LAB). However, both of these populations were totally replaced by LAB during ripening. The present survey contributes by describing the microbiota of this ancient cheese in more detail during fermentation and ripening.

Keywords: Liqvan cheese, Illumina analysis, RNA-based methods

Introduction

Liqvan (or Lighvan) cheese is the most famous traditional, semi-hard, feta like, starter-free, Iranian cheese. It is manufactured from raw ewe's milk in the Liqvan village in the stratovolcano mountainous area of Sahand in the East Azerbaijan province of Iran. The production process is conducted within 2h from milking. Ewe's milk is cooled to 25°C and coagulated with homemade or industrial lamb rennet. After 2 hours of drainage, the coagulum is cut into 15x15 cm cubes and placed in 15-20% saline water, where it is left for 9 to 10 hours. The curd is consequently covered with dry salt and kept in a basin for 2-3 days, during which whey drainage continues. During this period, the cubes are turned (upside down) at least five times. Finally, the resulting curd is placed in 10–12% saline water in metal containers. Ripening takes place in deep natural or manufactured caves over a period of 6 months. This traditional manufacturing process, which has remained unchanged for centuries in this specific area, has attracted particular interest in the last few years (Edalatian et al., 2012; Kafili et al., 2009; Navidghasemizad et al., 2009; Barouei et al., 2008). Typical sensory properties of such traditional cheeses mostly affected by dairy animals' type, breeds and nutrition as well as their indigenous microbiota which drive biochemical process during cheese making. However, this traditional way of production has made also some problems during these days. First of all Standard Organization of Iran recommends strictly the pasteurization of milk to improve the safety of dairies and this process might eliminate the key functional microorganisms involved at cheese ripening and acidification. In spite of food safety, product standardization is necessary

for entrance at international marketing systems so identification and characterization of active starter and non-starter microorganisms during each step of production seems to be important.

Food microbiota is currently monitored by means of a variety of culture dependent and independent techniques. DNA and/or RNA based approaches are considered indispensable tools for the detection, identification and characterization of microorganisms in food, and they have led to a profound change in the microbial ecology study of fermented foods in recent decades. These techniques are faster and more reliable than conventional culturing methods, which fail to reproduce ecological niches and symbiotic relationships (Carraro et al., 2011, Cocolin et al., 2013; Ndoye et al., 2011). Moreover, Next Generation Sequencing (NGS) approaches have been used successfully to monitor microbial communities in foods, especially in milk-based foods lately (Aldrete-Tapia et al., 2014; Bassi et al., 2015; Dugat-Bony et al., 2015; Garofalo et al., 2015; Minervini et al., 2015; O'Sullivan et al., 2015; Alessandria et al., 2016; Dalmasso et al., 2016; De Pasquale et al., 2016; Escobar-Zepeda et al., 2016).

Previous studies on Liqvan diversity, based on culture-dependent and on DNA based DGGE method, led to the identification of some dominant bacteria (Edalatian et al., 2012; Kafili et al., 2009; Navidghasemizad et al., 2009; Barouei et al., 2008). However, to the best of the authors' knowledge, no RNA based studies on this cheese have been carried out so far. Moreover, in order to have a more complete picture of viable microbiota thriving in this cheese during ripening process, we used illumina-sequencing method as a new powerful tool for analyzing of milk-based foods. This thorough analysis of RNA-based microbiota, will help us to follow the trace of possible pathogenic microorganisms and their survival during manufacturing.

Therefore, in this study, the microbial population of the bacterial communities of Liqvan cheese have been studied during its manufacturing through reverse transcriptase PCR (RT-PCR)-DGGE. Furthermore, the total counts and diversity of the viable bacterial populations in the raw milk, curd and ripened cheese have been investigated using RT quantitative PCR (RT-qPCR) and 16S rRNA gene amplicon sequencing, respectively.

Material and methods

Sample collection

Samples of raw milk, curds and cheeses were collected aseptically on-site at five different local authentic production units (A to E) in the Liqvan village, which covers an area of 19.76 square kilometers in the south-western part of Tabriz province. All the milk samples collected at the same day after transferring to the units and cooled at 25 °C. While three days curds gathered after coagulation and drainage. The milk

85 and curds were transferred directly to the lab, albeit the cheeses were collected after 180 days of ripening
86 from the same source of milk. Three trials as replicates were taken from each dairy for every sampling.
87 The samples were transported at 4-8 °C and kept at -20 °C.
88 This cheese mainly characterized by a pH value and total fat and protein percentage around 5.02±0.5,
89 20.3±2.0 and 15.7±1.0, respectively.
90
91 **Reverse transcription and PCR amplification of the microbial community**
92 Aliquots of 2 milliliters of milk were centrifuged for 10 min at 14,000 × g under refrigerated conditions (4
93 °C). The samples were placed at -80 °C for 30 min to allow separation of the fat layer, which was then
94 removed by means of a sterile spatula. After thawing, the supernatant was discarded and cell pellets were
95 re-suspended in 1 ml ¼ strength Ringer solution. The mixture was centrifuged at 14,000 × g for 10 min at
96 4 °C, and the supernatant was again discarded. The pellets were re-suspended in a 50 µl lysozyme
97 solution (50 mg/ml, Sigma, Milan, Italy) and incubated at 37 °C for one hour. The lysate was then
98 subjected to nucleic acid extraction, using the Master-Pure complete DNA and RNA Purification Kit
99 (Epicentre, Madison, WI, USA), according to the manufacturer´s instructions.
100 In the case of the curds and cheeses, five grams of samples was homogenized with 20 ml of ¼ strength
101 Ringer solution. One ml of this solution was transferred to a 1.5 ml micro-tube, and the same procedure as
102 the one described for the milk was followed. The resulting RNA was then treated with 3 µl of TURBO
103 DNase (Ambion, Milan, Italy) and incubated for 3 h at 37 °C in order to eliminate the DNA. Complete
104 DNA digestion was confirmed using 1 µl of extracted RNA in PCR reactions with primers 338f and 518r
105 (Muyzer et al., 1993). If a PCR product was obtained, the DNase treatment was repeated. The resulting
106 RNA was quantified using the NanoDrop 1000 spectrophotometer (Thermo Scientific, Milan, Italy) and
107 standardized at 500 ng/µL.
108 RT reaction was carried out using M-MLV reverse transcriptase (Promega, Milan, Italy). An aliquot of
109 500 ng of RNA was mixed with 1 µl of random primer (Promega Milan, Italy) for each sample. A volume
110 of 10 µl was reached by adding DNase- and RNase-free sterile water (Sigma), and the solution was
111 incubated at 75 °C for 10 min. The mixture was then immediately placed on ice for 5 min, and then added
112 to a microtube vial containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl2, 10 mM DTT, 2 mM
113 of each dNTP, 1 µl of 200 U/µl M-MLV and 0.96 U of RNasin ribonuclease inhibitor (Promega). Reverse
114 transcription was carried out at 42 °C for 1 h (Alessandria et al., 2010). The PCR reaction was performed
115 in triplicate in a MyCycler (BioRad, Hercules, CA, USA) on each sample.
116 The V3 region of the 16S rRNA gene in the bacteria was amplified with primers 338f-GC/518r (Muyzer
117 et al., 1993). PCR amplification was carried out adopting the same procedure used in previous studies on
118 the ripening of hard cheese (Bautista-Gallego et al., 2014).

119 **DGGE analysis and band sequencing**

120 The RT-PCR products were analyzed by means of DGGE, using the DCode system apparatus (BioRad,
121 Hercules, CA, USA). Polyacrylamide gels (8% W/V, Acrylamide-Bisacrylamide 40% solution 37.5:1, 0.8
122 mm thickness) were prepared using 30-50% urea-formamide denaturing gradients (100% corresponded to
123 8 M urea and 40% (wt/vol) formamide). The gels were run for 240 min at 130 V, stained with SYBR®
124 Gold Nucleic Acid Gel Stain (Life Technologies, Milano), visualized under UVtransillumination and
125 photographed by using UVIpro Platinum 1.1 Gel Software (Eppendorf).

126 Identification of the microbial populations was carried out by cutting selected bands from the DGGE gels.
127 DNA was eluted in 50 µl water overnight at 4 °C, and again run in DGGE after re-amplification in order
128 to check the electrophoretic mobility with respect to the excised bands. After the check, DNA was
129 amplified again with non GC-clamp primers and sent for sequencing to GATC-Biotech (Cologne,
130 Germany). A fingerprint database was created using Bionumerics software, version 4.6 (Applied Maths,
131 Sint Marten Latem, Belgium).

132

133 **Construction of a/the LAB standard curve for viable bacterial enumeration**

134 The viable bacterial populations in the milk, curd and cheese were determined according to Alessandria et
135 al. (2016). Milk, curd and cheese were sterilized. At the same time, overnight pure cultures of
136 *Lactococcus lactis* subsp. *lactis, Lactobacillus plantarum, Lactobacillus pentosus, Enterococcus faecium*
137 and *Lactobacillus helveticus* was prepared. The count of each culture is reached to approximately 10^9
138 (CFU/g or ml, which checked by culturing methods) and mixed. Serial dilutions of this overnight cultures
139 were prepared using ¼ strength Ringer and 10 ml of each dilution (which contains the final
140 concentrations of 10^8 to 10 CFU/g or ml), inoculated to 10 g or ml of food samples.

141 The samples were then homogenized in a Stomacher (Interscience Rockland, MA, USA) for 1 min, and 1
142 ml of each mixture was subjected to RNA extraction and reverse transcription, as described above. QPCR
143 amplifications, targeting the V3 region of the 16S rRNA gene, were performed in a final volume of 25 µl,
144 using SSo Advanced Sybr Green Supermix (Bio-Rad, Italy). One µl of cDNA was amplified with 338f
145 and 518r primers at a final concentration of 400 nM in a Chromo 4 real-time PCR Detection System
146 (Biorad, Milan, Italy) (Alessandria et al., 2010). In all cases and at each step the bacterial
147 counts was checked by triple plate counting routine methods and Standard curves were constructed by
148 plotting the threshold cycle (C_T) values against log CFU/g or CFU/ml on MRS agar. MRS agar plates
149 were incubated at 37 °C for 48 h in microaerophilic conditions. The correlation coefficients ($R2$) and
150 efficiency of amplification were calculated as described in Higuchi et al. (1993).

151 Quantification of the viable bacteria in the milk, curd and cheese was carried out by amplifying 1 µl of
152 cDNA, and using the standard curves for the calculation of the counts.

153 **Bacterial RNA analysis by means of Illumina high-throughput sequencing**

154 A high-throughput sequencing approach was applied to a total of 15 samples (5 samples from milk, curd
155 and cheese). cDNA, obtained as explained above, was used to study the microbial diversity of the active
156 populations. The V3–V4 region of the 16S rRNA gene was amplified using the following primers: the
157 16S Amplicon Forward Primer (5'-
158 TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG-3') and the 16S
159 Amplicon Reverse Primer (5'-
160 GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC-3'), as
161 reported in Klindworth et al. (2013). Twenty-five µl PCR reactions were prepared using 12.5 mL of the
162 2X KAPA HiFi HotStart ReadyMix (Kapa Biosystems, Wilmington, MA), 1 µM of each primer, 2.5 µL
163 of cDNA template and PCR grade water. Twenty five cycles of denaturation (95 °C/ 30 sec), primer
164 annealing (55 °C/30 sec) and primer elongation (72 °C/30 sec), followed by a final elongation step (72
165 °C/ 5 min), were carried out.

166 The PCR products were purified by means of an Agencourt AMPure kit (Beckman Coulter,
167 Milan, Italy), and the resulting products were tagged by using the Nextera XT Index Kit
168 (Illumina Inc. San Diego, CA), according to the manufacturer's instructions. The sequencing
169 reaction was performed by Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d'Aosta,
170 S.C. Controllo Alimenti e Igiene delle Produzioni (Turin, Italy) with a MiSeq Illumina instrument
171 (Illumina) with V3 chemistry and generated 250 bp paired-end reads according to the manufacturer's
172 instruction. The software used for the base-calling and Illumina barcode demultiplexing processes, were
173 the MiSeq Control Soft. V2.3.0.3, the RTA v1.18.42.0 and the CASAVA v1.8.2.

174

175 **Bioinformatics analysis**

176 Paired-end reads were first merged using FLASH software (Magoc and Salzberg, 2011) with default
177 parameters. Join reads were further quality filtered (at Phred > Q20) using QIIME 1.9.0 software
178 (Caporaso et al. 2010) through multiple_extract_barcode.py and multiple_split_libraries_fastq.py script.
179 Usearch v8.1 software (Edgar et al., 2011) was adopted for chimera filtering, using the 16S reference
180 databases v9 (RDP classifier training database). OTUs were picked at 99% of similarity by means of
181 UCLUST clustering methods (Edgar, 2010), and representative sequences of each cluster were used to
182 assign taxonomy using the Greengenes 16S rRNA gene database, version 2013, by means of the RDP
183 classifier (Wang et al. 2007). Representative sequences belonging to clusters identified as *Lactobacillus*
184 spp. were double-checked using the BLAST (BLASTN) search program (http://
185 www.ncbi.nlm.nih.gov/blast/) to get the species level as best hit. Statistics and plotting were carried out in
186 the R environment (www.r-project.org). Alpha diversity indices were calculated using the diversity

187 function of the vegan package (Dixon, 2003). The Shannon-Wiener diversity index H' was further
188 analyzed using the t-test to assess any differences between the milk, curd and cheese samples. Weighted
189 UniFrac distance matrices and OTU tables were used to perform Adonis and Anosim statistical tests in R
190 environment. A filtered OTU table was generated at 0.5% abundance in at least two samples through
191 QIIME. PICRUSt (Langille et al., 2013) was used to predict the abundance of the gene families, on the
192 basis of 16S rRNA sequence data, as recently described (Ferrocino et al., 2016, Bokulich et al., 2015).
193 Nearest Sequenced Taxon Indexes (NSTI) were calculated in order to evaluate the accuracy of the
194 metagenome predictions (Langille et al., 2013). KEGG orthologs were then collapsed at a hierarchy level
195 of 3, and the table was imported into R. The made4 package was adopted, and hierarchical Ward-linkage
196 clustering, based on the Spearman correlation coefficients of the proportion of the activities belonging to
197 lipid carbohydrates and amino acid metabolism pathways, was used to produce a heatplot.
198 All the sequencing data were deposited at the Sequence Read Archive of the National Center for
199 Biotechnology Information (SRP074051).
200
201 **Results**
202 **Identification and total bacterial counts during ripening**
203 Eleven different bands, corresponding to the bacterial V3 region, were observed in the milk, curd and
204 cheese samples (Tab. 1). *Lactococcus lactis* and *Leuconostoc mesenteroides* were the predominant
205 species that recovered from all of the different types of samples, while *Streptococcus thermophilus* was
206 only found in the curd. On the other hand, bands related to *Enterobacter aerogenes* and *Pseudomonas*
207 *fluorescens* were mostly abundant in the milk and curd, but there was no trace of them in the cheese
208 samples. Other LAB were randomly distributed between curd and cheese except *Lactococcus*
209 *raffinolactis* that was not found in cheese.
210 The results of the quantification RT-PCR showed 8.5± 0.2 log CFU/ml for milk and 5.8± 0.2 and 3.5± 0.1
211 log CFU/g for curd and cheese, respectively. These data, which all approved by plate counting methods
212 (data was not shown) underline that, the total bacterial populations decreased during the ripening process.
213
214 **16S rRNA gene sequencing**
215 A total of 1,216,136 raw reads (2x250bp) were obtained after sequencing. After joint, a total of 431,012
216 reads passed the filters applied through QIIME, with an average value of 28,734 reads/sample, and a
217 sequence length of 457 bp. The estimated sample coverage for milk, curd and cheese are 94.69± 1.24,
218 95.21± 0.77 and 95.43± 1.38, respectively which indicates that there was a satisfactory coverage for all of
219 the samples (ESC > 94%). Moreover, The average OUT numbers between different producers are
220 2333.8± 515.63 in milk, 1545.6± 489.01 in curd and 1594.4± 298.14 in cheese while Chaol values that

221 reflect the operational taxonomic unit richness are ranged from 3933.67± 440.16, 2758.24± 904.50 and
222 2419.38± 306.59 between milk, curd and cheese . In addition, overall sample diversity index (Shannon)
223 are calculated as 6.47± 0.38, 5.17± 0.65 and 5.63±0.42 in milk, curd and cheese, respectively. Besides,
224 the alpha-diversity showed that there was a higher level of complexity ($P < 0.05$) in the milk samples than
225 in the curds and cheeses and between the cheeses and the curds ($P < 0.05$). It was possible to show,
226 through a principal coordinate analysis (PCoA), with a Weighted UniFrac distance matrix as well as from
227 Adonis and Anosim statistical tests, that the samples were grouped together according to their type ($P <$
228 0.001). In addition, no differences were found between the milk, cheese and curd samples for the different
229 repeats, according to the Adonis and Anosim statistical tests ($P > 0.001$). Adonis and Anosim statistical
230 tests of Weighted UniFrac distance matrix showed. significant difference among sample type ($P < 0.001$).
231 Only OTUs with a relative abundance of 0.5% in at least two samples are shown in Figure 1. The
232 abundance of the OTUs from the milks, curds and cheeses was averaged. The raw milk showed a varied
233 microbiota, characterized by a predominance of *Moraxellaceae* (20% of the relative abundance),
234 *Leuconostoc mesenteroides* (16%), *Staphylococcaceae* (13%), *Lactococcus raffinolactis* (7%),
235 *Pseudomonas* sp. (5%), *Enterobacteriaceae* (6%), and *Lactococcus lactis* (6%).
236 The curd samples showed a predominance of *Leuconostoc mesenteroides* (43%), *Lactococcus lactis*
237 (11%), *Enterobacteriaceae* (18%) and *Streptococcus* sp. (8%). The cheese samples were characterized by
238 a predominance of *Lactococcus lactis* (30%), *Leuconostoc mesenteroides* (20%), *Lactobacillus*
239 *fuchuensis* (10%), *Lactobacillus* sp. (10%) and *Pediococcus* sp. (10%).
240 The main differences between the three types of analyzed samples were due to the presence of several
241 contaminant OTUs, such as *Pseudomonas* and *Moraxellaceae,* which were found to be significantly more
242 abundant (g-test, $P < 0.001$) in the milk samples, while *Enterobacteriaceae* discriminated the curd
243 samples. On the other hand, the cheese samples were characterized by the presence of several lactic acid
244 bacteria (LAB), such as *Lactobacillus curvatus, Lactobacillus zeae*, *Lactobacillus fuchunensis*,
245 *Lactococcus lactis* and *Lactobacillus pentosus* (g-test, $P < 0.001$).
246 The OTU co-occurrence/exclusion pattern is shown in Figure 2, where only significant correlations are
247 reported (False Discovery Rate - FDR < 0.05). The characteristic OTU of the cheese samples, such as
248 *Lactobacillus curvatus, Lactobacillus zeae*, *Lactobacillus fuchunensis*, *Lactococcus lactis* , *Lactobacillus*
249 *pentosus* and *Lactobacillus kefiri,* showed the highest number of negative correlations, including a strong
250 exclusion of *Lactococcus* sp., *Moraxellaceae*, *Enterobacteriaceae*, *Enterococcus* sp., *Staphylococcaceae*
251 and other contaminants. *Moraxellaceae* showed a positive correlation with *Pseudomonas* sp., *Rothia* sp.
252 and *Staphylococacceae*.
253 As far as the predicted metagenomes are concerned, the weighted nearest sequenced taxon index (NSTI)
254 of the samples, expressed as the mean ± SD, was 0.028 ± 0.003. This index is the average branch length

255 that separates each OTU from a reference bacterial genome, weighted by the abundance of that OTU in
256 the sample. Thus, an NSTI score of 0.028 indicates a satisfactory accuracy for all of the samples (98%). It
257 was possible to differentiate the three different sample types (cheese, curd and milk) on the basis of the
258 predicted gene repertoires, associated with their microbiota (Fig. 3). A sub-cluster was mainly identified
259 for samples derived from the milk samples, and another cluster was identified for most of the cheeses and
260 curd samples. The dominant spoilage-related microbiota from the milk samples could considered, as a
261 microbial proteolitic consortium, due to presence of a presumptive abundance of genes related to lysine,
262 valine and the tryptophan metabolism. Besides, metabolic pathways related to a lipid metabolism, such as
263 a steroid metabolism, and to biosynthesis of the unsaturated fatty acids are also include. The other group
264 (cheese and curd samples) showed a higher presumptive abundance of genes related to a carbohydrate
265 metabolism, such as starch and sucrose metabolisms, amino-sugar and fructose metabolisms as well as
266 amino acid catabolism related genes.
267
268 **Discussion**
269 Liqvan is a traditional Iranian cheese that has been found to represent an interesting case for the study of
270 fermented food as it originates from a particular region in Iran, and it has long been prepared the same
271 way using an ancient methodology. Moreover, as this cheese is made from raw milk without the addition
272 of starter cultures, it can be speculated that its microbial ecology mainly depends on indigenous
273 microbiota. In this context, the first step towards protecting the microbial diversity of this traditional
274 regional food is to elucidate the evolution of the active microbial populations during its manufacturing
275 and ripening, in order to guarantee both the safety and the quality of the final products (Alegría et al.,
276 2012). Previous studies on Liqvan cheese (Edalatian et al., 2012; Kafili et al., 2009; Navidghasemizad et
277 al., 2009; Barouei et al., 2008), which were mainly based on the DNA PCR-DGGE method, showed that
278 *Lactococcus lactis, Lactococcus raffinolactis, Lactobacillus plantarum, Lactococcus garvieae,*
279 *Lactobacillus sakei, Lactobacillus casei* and *paracaei, Leuconostoc mesenteroides, Streptococcus* sp. and
280 *Enterobacter* sp. are the predominant groups of bacteria in this cheese.
281 To the best of our knowledge, no data have been reported on RNA based DGGE or Illumina analysis of
282 the microbiota of this traditional cheese, yet. In our research, the microbial ecology and active bacterial
283 communities of Liqvan cheese have been investigated using different culture-independent methods, at an
284 RNA level, during the ripening process.
285 Fifteen samples from raw milk, curd and cheese were collected from five different producers in different
286 households in the Liqvan village, East Azerbaijan province of Iran, the Mydanchay district of the central
287 part of the city of Tabriz, and were examined. The microbial investigation focused on bacteria population
288 targeting the V3 region of the 16S rRNA gene.

289 RNA-seq DGGE analysis of our samples indicates mainly that lactic acid bacteria mostly includes
290 *Lactococcus* sp., *Leuconostoc* sp., *Lactobacillus* sp. and *Streptococcus* sp. are detected in all stages of
291 production while spoilage bacteria like S*taphylococcus* sp., *Enterobacter* sp. and *Pseudomonas* sp. are
292 detected only at milk samples and not in curds and cheese.
293 Phylogenetic assignment of NGS sequence data analysis revealed that microbial α-diversity was more
294 remarkable at milk level and corresponds to *Firmicutes, Proteobacteria* and *Actinobacteria* however, this
295 diversity of population decreased during ripening dominated by *Firmicutes*.
296 At the family level, *Moraxcellaceae, Pseudomonadaceae, Corynebacterineae*, *Flavobacteriaceae*,
297 *Comamonadaceae, Staphylococaceae* and *Enterobacteriaceae* were the most abundant contaminant
298 microbiota in the milk, while *Lactobacillaceae*, *Streptococcaceae* and *Leuconostocaceae* were the
299 predominant microbial families in the curds and cheese. There are some number of notable observations
300 among the subdominant populations. Some records from sheep feces, sheep wool, soil and grass
301 contaminants like those that *Brochothrix* sp. and *Rothia* sp. are also detected at milk samples nevertheless
302 there is no trace of them in curd and cheese. Furthermore, *Streptococcus* sp. especially *Streptococcus*
303 *thermophiles* were detected exclusively in curd samples while its population decreased remarkably in
304 cheese.
305 The microbial co-occurrence/exclusion patterns also proved these results, as there was a strong co-
306 exclusion effect between *Lactobacillaceae and Enterobacteriaceae,* as well as between *Lactobacillaceae*
307 and *Staphylococcaceae*. *Lactobacillus* species co-occurred each other. These results confirmed that the
308 LAB microbial populations tended to dominate the cheese microbiota and to limit the development of
309 spoilage bacteria, as has recently been demonstrated for other types of cheeses (Alessandria et al., 2016).
310 This diversity and dynamics of prokaryotes in this study are in agreement with the results that had
311 previously been reported for Liqvan and other semi hard raw milk based cheeses (Bozoudi et al., 2016;
312 Ryssel et al., 2015; Masoud et al., 2011; Rantsiou et al., 2008) although some differences were noted
313 which may related to difference in samples origin and producers.
314 The use of predicted metagenomes has proved to be useful for the observation of the putative gene
315 repertoires in the analyzed samples. The milk samples were characterized by gene profiles that were
316 related to the aminoacid metabolism, and high proteolitic activity due to the spoilage bacteria, while
317 putative genes characterized the curd and cheese related to carbohydrate depletion, which was presumably
318 associated with the nature of the highly competitive LAB that dominated during the ripening period.
319 In conclusion, in the present work, the study of microbial population with culture independent techniques
320 has provided a better understanding of the bacterial structure of Liqvan cheese during its processing and
321 ripening.
322

Acknowledgments

The authors would like to thank the Iranian Biological Resource Center (IBRC), ACECR for their partial financial support. We also dedicate our special thanks to the Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d'Aosta, S.C. Controllo Alimenti e Igiene delle Produzioni (Turin, Italy) for performing the 16S rRNA sequencing.

References

Aldrete-Tapia, A., Escobar-Ramírez, M.C., Tamplin, M.L., Hernández-Iturriaga, M., 2014. High-throughput sequencing of microbial communities in Poro cheese, an artisanal Mexican cheese. Food Microbiol. 44, 136–141. doi:10.1016/j.fm.2014.05.022.

Alessandria, V., Ferrocino, I., De Filippis, F., Fontana, M., Rantsiou, K., Ercolini, D., Cocolin, L., 2016. Microbiota of an Italian Grana-Like Cheese during Manufacture and Ripening, Unraveled by 16S rRNA-Based Approaches. Appl. Environ. Microbiol. 82(13), 3988-95.

Alessandria, V., Dolci, p., Rantsiou, K., Pattono, D., Dalmasso, A., Civera, T., Cocolin, L., 2010. Microbiota of the Planalto de Bolona: an artisanal cheese produced in uncommon environmental conditions in the Cape Verde Islands. World. J. Microbiol. Biotechnol. 26(12), 2211-2221.

Alegría, A., Szczesny, P., Mayo, B., Bardowski, J., Kowalczyk, M., 2012. Biodiversity in Oscypek, a Traditional Polish Cheese, Determined by Culture-Dependent and -Independent Approaches. Appl. Environ. Microbiol. 78(6), 1890-1898.

Barouei, J., Karbassi, A., B. Ghoddusi, H., Mortazavi, A., 2008. Lactic microflora present in Liqvan ewes' milk cheese. Int. J. Food. Prop. 11(2), 407-414.

Bassi, D., Puglisi, E., Cocconcelli,P.S., 2015. Understanding the bacterial communities of hard cheese with blowing defect. Food. Microbiol. 52, 106-118.

Bautista-Gallego, J., Alessandria, A., Fontana, M., Bisotti, S., Taricco, S., Dolci, P., Cocolin, L., Rantsiou, K., 2014. Diversity and functional characterization of *Lactobacillus* spp. isolated throughout the ripening of a hard cheese. Int. J. Food. Microbiol. 181, 60-66.

Bokulich, N., Amiranashvili, L., Chitchyan, K., Ghazanchyan, N., Darbinyan, K., Gagelidze, N., Sadunishvili, T., Goginyan, V., Kvesitadze, G., Torok, T., A. Mills, D., 2015. Microbial biogeography of the transnational fermented milk matsoni. Food. Microbiol. 50, 12-19.

361 Bozoudi, D., Torriani, S., Zdragas, A., Litopoulou-Tzanetaki, E., 2016. Assessment of microbial diversity
362 of the dominant microbiota in fresh and mature PDO Feta cheese made at three mountainous areas of
363 Greece. LWT-food. Sci. Technol. 72, 525-533.

364

365 Carraro, L., Maifreni, M., Bartolomeoli, I., Martino, M. E., Novelli, E., Frigo, F., Marino, M., Cardazzo,
366 B., 2011. Comparison of culture-dependent and –independent methods for bacterial community
367 monitoring during Montasio cheese manufacturing. Res. Microbiol. 162, 231-239.

368

369 Cocolin, L., Alessandria, L., Dolci, P., Gorra, R., Rantsiou, K., 2013. Culture independent methods to
370 assess the diversity and dynamics of microbiota during food fermentation. Int. J. Food. Microbiol. 167,
371 29-43.

372 Dalmasso, A., Dolores Soto del Rio, M.D.L., Civera, T., Pattono, D., Cardazzo, B., Bottero, M.T., 2016.
373 Characterization of microbiota in Plaisentif cheese by high-throughput sequencing. LWT – Food. Sci.
374 Technol. 69, 490-496.

375 De Pasquale, I., Di Cagno, R., Buchin, S., De Angelis, M., Gobbetti, M., 2016. Spatial distribution of the
376 metabolically active microbiota within Italian PDO ewes' milk cheeses. PLoS One 11, 1–23.
377 doi:10.1371/journal.pone.0153213.
378

379 Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14, 927–930.

380

381 Dugat-Bony, E., Straub, C., Teissandier, A., Onésime, D., Loux, V., Monnet, C., Irlinger, F., Landaud, S.,
382 Leclercq-Perlat, M.-N., Bento, P., Fraud, S., Gibrat, J.-F., Aubert, J., Fer, F., Guédon, E., Pons, N.,
383 Kennedy, S., Beckerich, J.-M., Swennen, D., Bonnarme, P., 2015. Overview of a surface-ripened cheese
384 community functioning by meta-omics analyses. PLoS One 10, e0124360.
385 doi:10.1371/journal.pone.0124360.
386

387 Edalatian, M.R., Habibi Najafi, M.B., Mortazavi, A., Mayo, B., 2012. The biodiversity and evolution of
388 lactic flora during ripening of the Iranian semisoft Lighvan cheese. Int. J. Dairy. Technol. 65(1), 81-89.
389

390 Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and
391 speed of chimera detection. Bioinformatics. 27(16), 2194-2200. doi: 10.1093/bioinformatics/btr381
392 [PMID 21700674].
393

394 Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics. 26,
395 2460–2461.
396

397 Escobar-Zepeda, A., Sanchez-Flores, A., Baruch, M.Q., 2016. Metagenomic analysis of a Mexican
398 ripened cheese reveals a unique complex microbiota. Food. Microbiol. 57, 116-127.
399

400 Ferrocino, I., Greppi, A., Storia, A.L., Rantsiou, K., Ercolini, D., Cocolin, L., 2016. Impact of nisin-
401 activated packaging on microbiota of beef burgers during storage. Appl. Environ. Microbiol. 82(2), 549-
402 559.
403

404 Garofalo, C., Osimani, A., Milanović, V., Aquilanti, L., De Filippis, F., Stellato, G., Di Mauro, S.,
405 Turchetti, B., Buzzini, P., Ercolini, D., Clementi, F., 2015. Bacteria and yeast microbiota in milk kefir
406 grains from different Italian regions. Food. Microbiol. 49, 123-33.

407

408 Higuchi, R., Fockler, C., Dollinger, G., Watson, R., 1993. Kinetic PCR analysis: real-time monitoring of
409 DNA amplification reactions. Nat. Biotechnol. 11, 1026-1030.

410

411 Kafili, T., Razavi, S.H., Emam Djomeh, Z., Naghavi, M.R., Álvarez-Martín, P., Mayo, B., 2009.
412 Microbial characterization of Iranian traditional Lighvan cheese over manufacturing and ripening via
413 culturing and PCR-DGGE analysis: identification and typing of dominant lactobacilli. Eur. Food. Res.
414 Technol. 229, 83–92.

415

416 Klindworth, A., Pruesse, E., Schweer, T., Peplles, J., Quast, C., Horn, M., Oliver Glöckner, F., 2013.
417 Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation
418 sequencing-based diversity studies. Nucleic. Acids. Res. 41(1).

419

420 Langille, M. G.I., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J., Clemente, J.C.,
421 Burkepile, D.E., Vega Thurber, R.L., Knight, R., Beiko, R.G., Huttenhower, C., 2013. Predictive
422 functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat. Biotechnol.
423 8, 1-10.

424

425 Magoc, T., Salzberg, S., 2011. FLASH: Fast length adjustment of short reads to improve genome
426 assemblies. Bioinformatics. 27(21), 2957-2963.

427

428 Masoud, W., Takamiya, M., K. Vogensen, F., Lillevang, S., Al-Soud, W.A., Sørensen, S.J., Jakobsen,
429 M., 2011. Characterization of bacterial populations in Danish raw milk cheeses made with different
430 starter cultures by denaturating gradient gel electrophoresis and pyrosequencing. Int. Dairy. J. 21, 142-
431 148.

432

433 Minervini, F., Lattanzi, A., Angelis, M.D., Celano, G., Gobbetti, M., 2015. House microbiotas as sources
434 of lactic acid bacteria and yeasts in traditional Italian sourdoughs. Food. Microbiol. 52, 66-76.

435

436 Muyzer, G., De Waal, EC., Uitterlinden, AG., 1993. Profiling of complex microbial populations by
437 denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for
438 16S rRNA. Appl. Environ. Microbiol. 59:695–700.

439

440 Navidghasemizad, S., Hesari, J., Saris, P., Nahaei, M.R., 2009. Isolation of lactic acid bacteria from
441 Lighvan cheese, a semihard cheese made from raw sheep milk in Iran. Int. J. Dairy. Technol. 62(2), 260-
442 264.

443

444 Ndoye, B., Rasolofo, E.A., LaPointe, G., Roy, D., 2011. A review of the molecular approaches to
445 investigate the diversity and activity of cheese microbiota. Dairy. Sci. Technol. 91, 495–524.

446

447 O'Sullivan, D.J., Cotter, P.D., O'Sullivan, O., Giblin, L., McSweeney, P.L.H., Sheehan, J.J., 2015.
448 Temporal and Spatial Differences in Microbial Composition during the Manufacture of a Continental-
449 Type Cheese. Appl. Environ. Microbiol. 81, 2525–2533.

Rantsiou, K., Urso, R., Dolci, P., Comi, G., Cocolin, L., 2008. Microflora of Feta cheese from four Greek manufacturers. Int. J. Food. Microbiol. 126 (1-2), 36-42.

Ryssel, M., Johansena, P., Al-Soud, W.A., Sørensen, S., Arneborg, N., Jespersen, L., 2015. Microbial diversity and dynamics throughout manufacturing and ripening of surface ripened semi-hard Danish Danbo cheeses investigated by culture-independent techniques. Int. J. Food. Microbiol. 215, 124-130.

Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microb 73(16): 5261-5267.

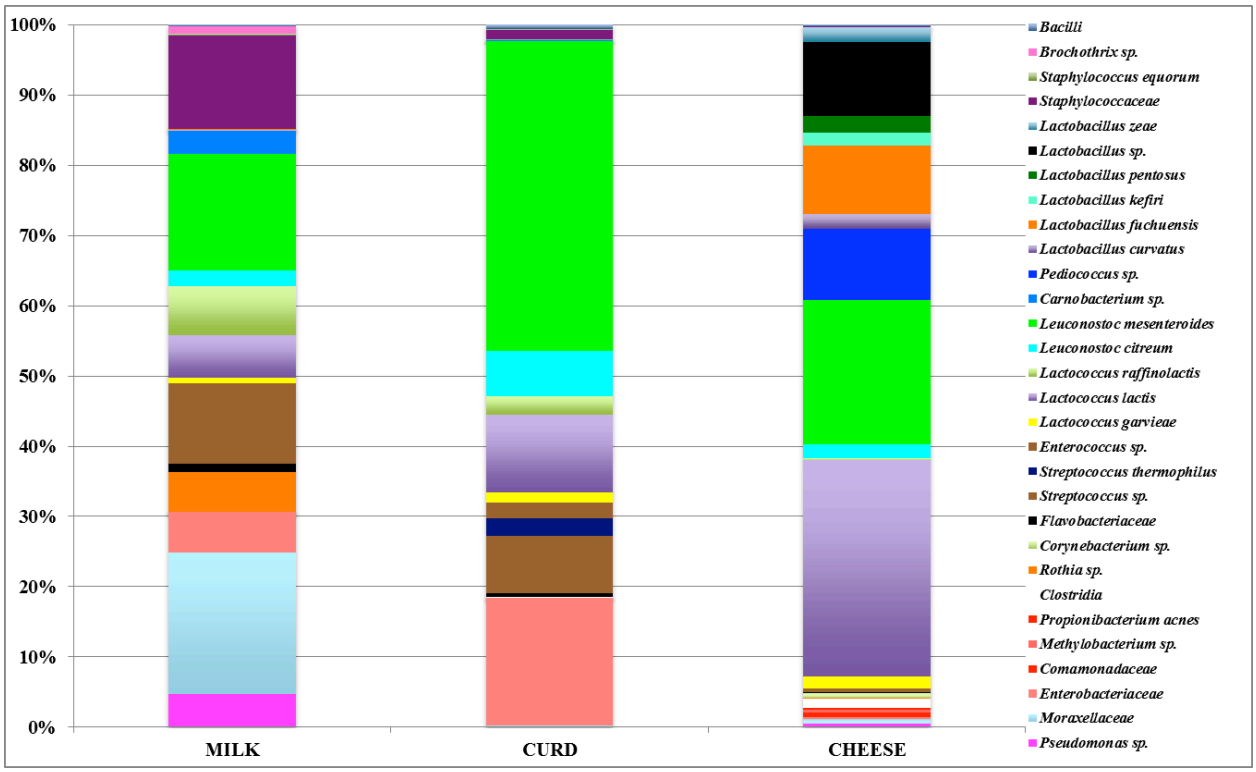
462 **Table 1:** Identification of the selected bands from bacterial DGGE gels and their occurence in the samples
463 considered in this study

Bacterial species	**Stage of manufacturing**		
	Milk	Curd	Cheese
Lactococcus lactis ssp. *lactis*	+	+	+
Lactococcus lactis ssp. *cremoris*	-	+	+
Leuconostoc mesenteroides	+	+	+
Lactococcus raffinolactis	+	+	-
Lactobacillus pentosus	-	-	+
Lactococcus garvieae	-	+	+
Enterobacter aerogenes	+	+	_
Pseudomonas fluorescens	+	-	-
Lactobacillus fuchuensis	-	-	+
Staphylococcus sp.	+	-	-
Streptococcus thermophilus	-	+	-

464
465
466
467
468
469
470
471
472
473
474
475
476
477

478 **FIG 1:** Relative abundance of the major taxonomic groups detected by means of 16S sequencing. Only OTUs with
479 an incidence above 0.5% in at least two samples are shown. The abundance of OTUs from the milks, curds and
480 cheeses was averaged.

481



482

483

484

485

486

487

488

489

490

491

492

493

494

FIG 2: Significant co-occurrence and co-exclusion relationships between bacterial OTUs. Spearman's rank correlation matrix of OTUs with > 0.5% abundance in at least 2 samples. Strong correlations are indicated with large squares, whereas weak correlations are indicated with small squares. The color of the scale bar denotes the nature of the correlation, with 1 indicating a perfectly positive correlation (dark blue) and -1 indicating a perfectly negative correlation (dark red). Only significant correlations (FDR < 0.05) are shown.

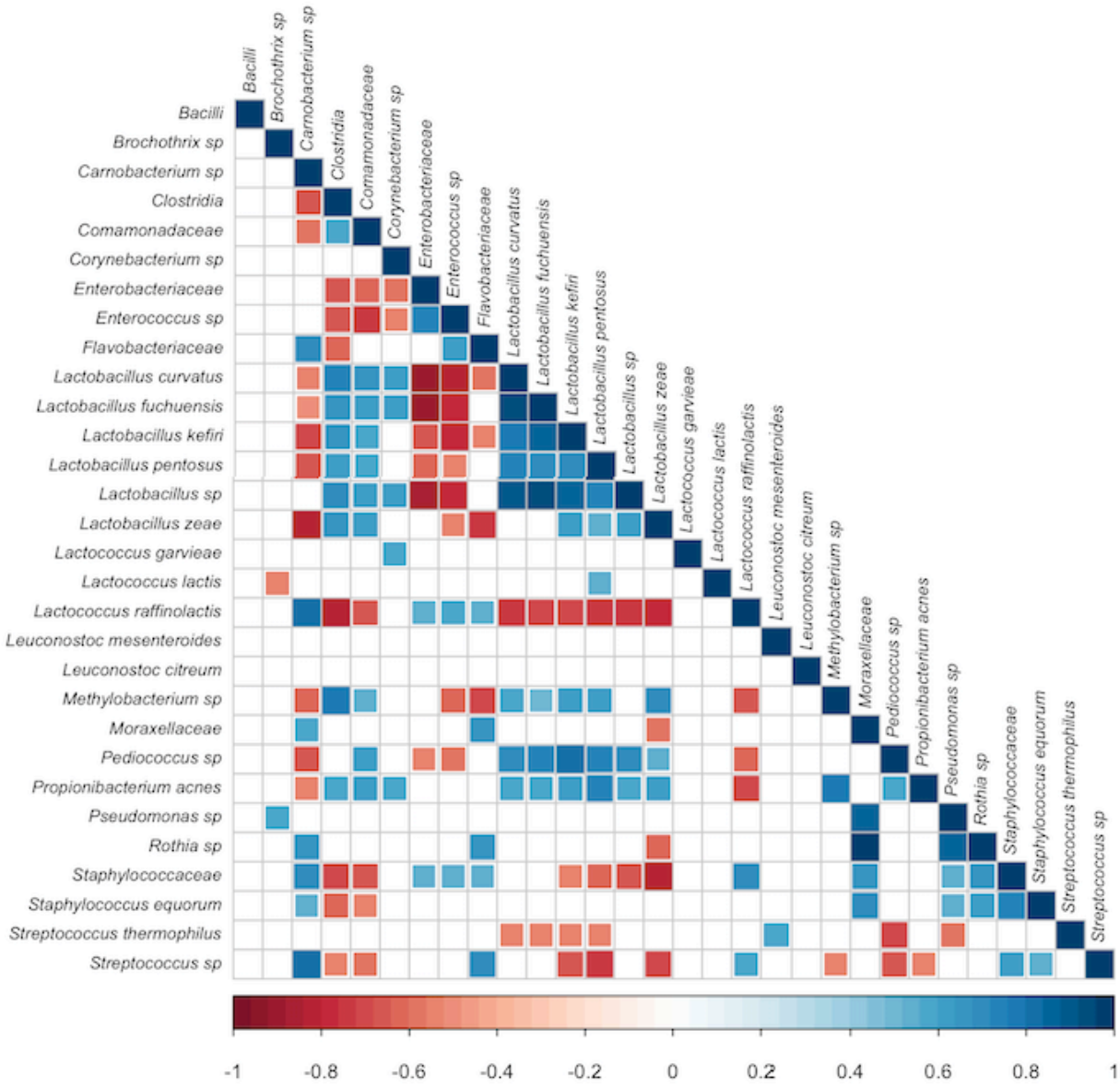


FIG 3: Heat plot of the abundances of genes presumptively belonging to carbohydrate (blue squares), amino acid (red squares) and lipid (green squares) metabolism pathways in the milk (upper blue bar), curd (upper green bar) and cheese (upper red bar). Rows and columns are clustered by means of Ward linkage hierarchical clustering. The intensity of the colors represents the degree of correlation between the samples and KO as measured by Spearman's correlations.