

Manuscript Number:

Title: Understanding the bacterial communities of hard cheese with blowing defect

Article Type: Research Paper

Keywords: Grana Padano cheese; blowing defect; Clostridium; Lactobacillus; Illumina-based sequencing; lysozyme.

Corresponding Author: Prof. Pier Sandro Cocconcelli,

Corresponding Author's Institution: Istituto di Microbiologia

First Author: Daniela Bassi, Researcher

Order of Authors: Daniela Bassi, Researcher; Edoardo Puglisi, Researcher; Pier Sandro Cocconcelli

Abstract: The environment of hard cheese encourages bacterial synergies and competitions along the ripening process, which might lead in defects such as clostridial blowing. In this study, Denaturing Gradient Gel Electrophoresis (DGGE), a quantitative Clostridium tyrobutyricum PCR and next-generation Illumina-based sequencing of 16S rRNA gene were applied to study 83 Grana Padano spoiled samples. The aim was to investigate the community of clostridia involved in spoilage, the ecological relationships with the other members of the cheese microbiota, and the effect of lysozyme. Three main genera were dominant in the analysed cheeses, Lactobacillus, Streptococcus and Clostridium, and allowed an assignment at the species level of 94.3% of 4,477,326 high quality sequences. *C. tyrobutyricum* and *C. butyricum* were the most prevalent clostridia. Hierarchical clustering based on the abundance of bacterial genera, revealed three main clusters: one characterized by the highest proportion of Clostridium, a second where Lactobacillus was predominant and the last, dominated by *S. thermophilus*. Ecological relationships among species were found: cheeses characterized by an high abundance of *S. thermophilus* and Lactobacillus rhamnosus were spoiled by *C. tyrobutyricum* while, when *L. delbrueckii* was the most abundant Lactobacillus, *C. butyricum* was the dominant spoiling species. Lysozyme also shaped the bacterial community, reducing *C. tyrobutyricum* in favour of *C. butyricum*. Moreover, this preservative increased the proportion of *L. delbrueckii* and obligate heterofermentative lactobacilli and lowered *L. helveticus* and non-starter species, such as *L. rhamnosus* and *L. casei*.

PIACENZA-CREMONA

Istituto di Microbiologia



UNIVERSITÀ
CATTOLICA
del Sacro Cuore

16 March 2015

Dear prof. M.L. Tortorello,

please find attached the manuscript "Understanding the bacterial communities of hard cheese with blowing defect" by Daniela Bassi, Edoardo Puglisi and Pier Sandro Cocconcelli submitted for possible publication in Food Microbiology.

In this work, an Illumina-based technology for the analysis of 16S rRNA has been used for the first time to determine in 83 cheese samples with blowing defect the community of clostridia involved in spoilage and the ecological relationships with the other members of the cheese microbiota. This new methodology based on 16S rDNA amplification together with high-throughput sequencing technology (HTS) made possible to assign at the species level the 94.3% of the achieved sequences. Moreover, this technique, together with a genus-specific *Clostridium* Cluster I PCR-DGGE and a qPCR on *C. tyrobutyricum* allowed investigating the main agents of spoiling defect, the effect of lysozyme as preservative and the ecological relationships among species. We believe that this work, due to high number of samples and the three different methodological approaches (qPCR, DGGE and NGS) improves actual knowledge on *Clostridium* species ecology in spoiled hard cheese and might help a better understanding of interactions among bacterial species during anomalous fermentation processes of cheeses.

Your sincerely

Prof. Pier Sandro Cocconcelli
Istituto di Microbiologia
Facoltà di Agraria - Università Cattolica del Sacro Cuore

Highlights

- A study of 83 Grana Padano cheese samples with blowing defect
- Community of clostridia involved in spoilage have been studied using DGGE and qPCR
- A NGS approach to analyse the ecological relationships among members of the cheese microbiota
- *C. tyrobutyricum* and *C. butyricum* were the most prevalent clostridia
- Lysozyme shaped the bacterial community

1 **Understanding the bacterial communities of hard cheese with blowing defect**

2

3

4 Daniela Bassi, Edoardo Puglisi , Pier Sandro Cocconcelli*

5 *Istituto di Microbiologia, Università Cattolica del Sacro Cuore, via Emilia Parmense 84, 29100 Piacenza/via Milano*

6 *24, 26100 Cremona, Italy*

7

8

9

10

11

12

13

14

15

16

17 corresponding author: pier.cocconcelli@unicatt.it

18 **Abstract**

19 The environment of hard cheese encourages bacterial synergies and competitions along the ripening
20 process, which might lead in defects such as clostridial blowing. In this study, Denaturing Gradient
21 Gel Electrophoresis (DGGE), a quantitative *Clostridium tyrobutyricum* PCR and next-generation
22 Illumina-based sequencing of 16S rRNA gene were applied to study 83 Grana Padano spoiled
23 samples. The aim was to investigate the community of clostridia involved in spoilage, the
24 ecological relationships with the other members of the cheese microbiota, and the effect of
25 lysozyme. Three main genera were dominant in the analysed cheeses, *Lactobacillus*, *Streptococcus*
26 and *Clostridium*, and allowed an assignment at the species level of 94.3% of 4,477,326 high quality
27 sequences. *C. tyrobutyricum* and *C. butyricum* were the most prevalent clostridia. Hierarchical
28 clustering based on the abundance of bacterial genera, revealed three main clusters: one
29 characterized by the highest proportion of *Clostridium*, a second where *Lactobacillus* was
30 predominant and the last, dominated by *S. thermophilus*. Ecological relationships among species
31 were found: cheeses characterized by an high abundance of *S. thermophilus* and *Lactobacillus*
32 *rhamnosus* were spoiled by *C. tyrobutyricum* while, when *L. delbrueckii* was the most abundant
33 *Lactobacillus*, *C. butyricum* was the dominant spoiling species. Lysozyme also shaped the bacterial
34 community, reducing *C. tyrobutyricum* in favour of *C. butyricum*. Moreover, this preservative
35 increased the proportion of *L. delbrueckii* and obligate heterofermentative lactobacilli and lowered
36 *L. helveticus* and non-starter species, such as *L. rhamnosus* and *L. casei*.

37 **Key words:** Grana Padano cheese, blowing defect, *Clostridium*, *Lactobacillus*, Illumina-based
38 sequencing, lysozyme.

39

40

41

42

43 **1. Introduction**

44 Long ripening times characteristic of hard cheeses production create the favourable conditions for
45 microbial communities able to survive along the dairy process. This could result in beneficial
46 effects or, alternatively, in spoilage defects that damage the final product. Late blowing is one of the
47 most frequent problems still affecting hard cheese production in dairy factories. Spore-forming
48 clostridia are considered the main agents of such damages (Coulon et al., 1991; Guericke, 1993;
49 Ingham et al., 1998; Klijn et al., 1995; Vissers, 2007) thanks to their spores surviving attitude to
50 heat treatment and their added capacity to germinate during cheese ripening. Different clostridial
51 species have been associated to spoiling of hard cheeses, firstly *Clostridium tyrobutyricum* as the
52 main agent (Bergère and Sivelä, 1990; Klijn et al., 1995; Le Bourhis et al., 2007b; Nishihara et al.,
53 2014) followed by *Clostridium sporogenes*, *Clostridium butyricum*, *Clostridium beijerinckii*, and
54 less frequently *Clostridium cochlearium*, *Clostridium perfringens*, *Clostridium septicum* (Le
55 Bourhis et al., 2007b; Lycken and Borch, 2006; Reindl et al., 2014). All these microorganisms,
56 alone or in association, have been related during time to the blowing problem, but few data are
57 available about their dynamic changes in the cheese shape and their relationships all along the
58 ripening period.

59 To reduce losses connected to clostridia spoilage, preservatives such as nitrate and lysozyme,
60 which consistently modify the cheese environment in which microorganisms survive and multiply,
61 are added to milk during hard cheese manufacture (Lodi and Stadhouders, 1990; Stadhouders,
62 1990). The use of nitrate in milk to prevent late blowing defect was also common in hard cheese-
63 making, particularly in Emmental production (Devoyod, 1975; Korenekova et al., 2000), but its
64 employ was banned after the European Food Safety Authority (EFSA) proposed to reduce levels of
65 nitrosamines in food products (EFSA, 2010). In Italy, Grana Padano (GP) hard cheese is produced
66 from raw cow's milk added with natural whey starter cultures and protected from clostridia spoilage
67 by lysozyme addition. Therefore, the use of this preservative may be responsible of different

68 bacterial dynamics in cheese, depending also on dairy conditions and on microbial milk
69 contamination.

70 The main bacterial populations associated to the cheese alteration process have been
71 previously studied with the use of cultivation-dependent techniques and PCR-based approaches, but
72 on a small number of samples (Cocolin et al., 2004; Garde et al., 2011; Le Bourhis et al., 2007b).
73 The new methodologies based on 16S rDNA amplification together with high-throughput
74 sequencing technology (HTS) have the advantage to determine, with a high resolution power, the
75 bacterial communities present in food environments: published examples include cheese (De
76 Filippis et al., 2014; Ercolini et al., 2012; Masoud et al., 2011; Masoud et al., 2012), seafood
77 (Kiyohara et al., 2012; Koyanagi et al., 2011; Roh et al., 2010) and dry fermented sausages (Polka
78 et al., 2015). The recent development in Illumina technology, coupled with multiplexing
79 approaches, allows analysing up two 16S rRNA hypervariable regions in thousands sequences per
80 sample, an approach that was shown to gain an almost complete coverage of the bacterial
81 communities of fermented foods, with correct taxonomical assignment at the species level for more
82 than 95% of the analysed reads (Polka et al., 2015). However, these approaches have not been yet
83 applied for the study of microbial communities involved in late blowing spoilage of hard cheese.

84 The aim of the present work was to assess the microbial communities of spoiled hard cheese
85 using new high-throughput sequencing technologies (HTS) associated to quantitative and
86 qualitative cultivation-independent techniques. Respectively, a PCR-Denaturing Gradient Gel
87 Electrophoresis (PCR-DGGE) approach to obtain a qualitative characterization of clostridia
88 heterogeneity, a TaqMan qPCR on single *C. tyrobutyricum* species and a HTS approach based on
89 Illumina MiSeq sequencing of the V3-V4 16S rRNA of total bacteria were applied to a total of 91
90 Grana Padano cheese samples from nine production facilities, with or without the addition of
91 lysozyme and with blowing defects appearance at different ripening times. Information gathered
92 from this study could be useful to assess the effect of lysozyme as a preservative and to measure the

93 effects of ripening time on the evolution of clostridia and their relationships with other bacterial
94 species present in the cheese paste.

95

96 **2. Materials and methods**

97 ***2.1. Cheese sampling***

98 Eighty-three samples of hard cheese, with anomalous pastry defects and cavities were collected
99 from the Grana Padano cheese production area in Northern Italy. The samples had different
100 ripening times corresponding to the appearance of defect and ranging from 1 to 23 months. In
101 addition, eight hard cheeses without defect were added as negative controls. A total of 56 cheese
102 samples were produced with the addition of lysozyme, while 35 samples were negative to the
103 additive. Collected cheese portions were 50 cm large and 20-30 cm high and were cut directly from
104 the entire cheese shapes; for each cheese, a representative sample was grounded and suddenly
105 frozen at -40°C in our laboratory until processing.

106 For all analyses reported in the paper, samples were labelled according to this legend: the first
107 two letters followed by a three numbers code indicate the production area, the following number
108 indicates the ripening time in months and “+” or “-“ refer to the use or not of lysozyme.

109

110 ***2.2. Bacterial strains and genomic DNA isolation***

111 Reference type and isolated clostridial and non-clostridial strains used in this study are listed in
112 table 1. For DNA extraction, 1 ml of culture was collected and centrifuged at 8,000g for 5 min.
113 Genomic DNA was isolated from the pelleted cells using the Nucleospin Tissue DNA Isolation Kit
114 (Macherey Nagel, Germany) and DNA was verified on a 1% agarose gel containing SYBR®YSafe
115 (Invitrogen Corporation Life technologies).

116

117 ***2.3. DNA extraction from cheese samples***

118 Total bacterial DNA was extracted from 50 g of grounded cheese for each sample; samples were
119 homogenized in 125- μ m filter stomacher bags (Biochek, Foster City, Calif.) with 50 ml of distilled
120 water. The filtered homogenates were collected and centrifuged at 5,000 rpm for 10 minutes at 4°C
121 to obtain pellets successively processed using the bead-technology based FastDNA® SPIN kit and
122 the FastPrep® Instrument (Qbiogene, Inc., CA) according to the protocol previously described
123 (Bassi et al., 2013a). DNA was then finally suspended in 100 μ L of DES solution for further
124 applications. All extracted nucleic acids were examined at a 0.8% electrophoresis agarose gel, and
125 quantified using the picogreen method of the Quant-iT™ HS ds-DNA assay kit (Invitrogen,
126 Paisley, UK) in combination with the QuBit™ fluorometer.

127 128 **2.4.16S-based Clostridium cluster-I specific PCR and DGGE analysis**

129 The PCR-DGGE strategy was based on a first amplification of a 235 bp fragment specific for the
130 order of *Clostridiales* in the V3-V4 region of the 16S rRNA gene, using primers DGGECl2 f (5'-
131 GCGGCGTGCCTAATACATGC-3') and P4 r (5'- ATCTACGCATTTACCGCTAC-3'). Primers
132 were designed after aligning more than 150 16S rDNA sequences among representatives of the
133 genera *Clostridium*, *Ruminococcus*, *Lactobacillus*, *Bacillus*, *Streptococcus*, *Leuconostoc*, and
134 *Enterococcus* that could frequently be present in dairy products in order to avoid cross-reactions.
135 PCR amplifications were performed in a final volume of 25 μ l, which included 12.5 μ l of 2X
136 MasterMix PCR (Promega), 0.4 μ M primers, and 3 μ l of genomic DNA. Template DNA was
137 generally diluted 10-fold and 100-fold to minimize PCR inhibitors. Reactions were heated to 95°C
138 for 5 min and cycled at 94°C for 30 sec, 60°C for 30 sec and 72°C for 30 sec. Cycles were repeated
139 35 times for all samples. Finally, 5 μ l of each PCR product was used for visualization on a 2%
140 agarose gel containing SYBR® Safe (Invitrogen Corporation Life technologies). Negative (without
141 DNA) and positive (with DNA from reference strains) controls were included in each amplification
142 run. 24 *Clostridium* strains, and 30 non-clostridial species (*Bacillus*, *Lactobacillus*, *Leuconostoc*,
143 *Listeria*, *Staphylococcus*, *Enterococcus*, *Streptococcus*, *Ruminococcus*) were selected and used for

144 the sensitivity test of the PCR protocol (Table 1). A nested PCR on the first DNA template was then
145 performed using primers DGGE2Cl f clamped at 5' with a 33bp GC sequence (5'-
146 CGCCCGCCGCGCGCGGGCGGGGCGGGGCGGGGGCC) and P4 r. The reaction was performed in
147 a 25 µl volume, with 12.5 µl of 2X MasterMix PCR (Promega), 0.1 µM of each primer and 1 µl of
148 DNA obtained in the previous PCR reaction. The amplification cycle was the same of the first PCR
149 reaction except for the annealing temperature of 63°C and the extension performed at 72 °C for 1
150 min. Denaturing gradient gel electrophoresis (DGGE) was performed using an INGENY phorU-2
151 (Ingeny International BV, Netherlands) DGGE system. A portion of each PCR template was loaded
152 on a 8% (w/v) acrylamide gel containing a 45-65% (w/v) denaturant gradient of urea and
153 formamide and electrophoresed at 90 V, 60°C for 16 h in 1X TAE buffer. Together with samples,
154 amplicons of the V3-V4 region of different *Clostridium* reference strains were loaded on the gel as
155 a ladder. Denaturing gradient gels were stained with 1X SYBR Green I (Roche, Milan, Italy) for 15
156 min and analysed under UV illumination. Selected DGGE bands were punched from the gel and
157 transferred in 50 µl of sterile water to let them diffuse overnight at 4°C. Only products migrating as
158 a single band, were PCR amplified with the original primer without GC-clamp, purified and sent to
159 a commercial sequencing facility (BMR Genomics, Padova, Italy) for sequencing. The sequences
160 were blasted in RDP database (<http://rdp.cme.msu.edu>) and in the GenBank using the NCBI
161 BLAST program (Altschul et al., 1997).

162 Banding patterns of DGGE profiles were analysed with Fingerprinting II software (Bio-Rad
163 Laboratories, Richmond, CA, USA) using the unweight-pair group method with averages
164 (UPGMA) for the generation of dendrograms.

165

166 **2.5.Real-time PCR analysis for *C. tyrobutyricum* enumeration**

167 The optimised TaqMan qPCR protocol for *C. tyrobutyricum* enumeration previously described
168 (Bassi et al., 2013b) was applied to all cheese samples. Reactions were performed with a
169 LightCycler® 480 Real-Time PCR System (Roche Diagnostics Corporation) using a LightCycler®

170 480 Probes Master kit. The analysis of variance using Tukey method and 95% confidence was
171 applied with R software v 3.1.1 in order to demonstrate significant differences respectively between
172 the effect of lysozyme and the ripening period with *C. tyrobutyricum* counts obtained in real-time
173 PCR. Data were considered significant with a *P*-value < 0.05.

174

175 **2.6. Illumina 16S rRNA next generation sequencing**

176 A high-throughput sequencing approach was applied on 40 samples selected among the total 91 in
177 order to have 20 lysozyme negative samples and 20 lysozyme positive with different ripening time
178 and clostridia composition. Two control samples without late blowing defect were also included.
179 The bacterial V3-V4 16S rRNA region was amplified with the primer pairs 343F (5'-
180 TACGGRAGGCAGCAG-3') and 802R (5'-TACNVGGGTWTCTAATCC-3') using the Phusion
181 Flash High-Fidelity Master Mix (Thermo Fisher Scientific, Inc. Waltham, MA, USA). PCR
182 reactions were performed in 25 µL using 12.5 µL of Phusion Master Mix, 0.5 µL of each primer,
183 0.1 ng of template DNA and nucleases-free PCR grade water. In order to allow a simultaneous
184 analysis of several samples in the same sequencing run, each sample was tagged by adding to the
185 forward primer a nine nucleic acids extension at its 5' end, where the first seven bases served as
186 sample index for multiplexing and the next two as linker bases designed in order to not match any
187 bacterial sequence entry at these position in the Ribosomal Database Project (RDP). This approach
188 was adopted in order to minimize potential biases possibly introduced by the indexed primer
189 extension. These biases were further addressed by adopting a two-step PCR program as
190 implemented in Berry et al. (Berry et al., 2011), with a first step in which untagged primers amplify
191 the template DNA for 23 cycles, and a second step where 1 µL of product of the first PCR is used
192 as a template for 10 final cycles using the barcoded primers. For both steps the PCR conditions used
193 were: initial 4 min at 94°C, cycles made up of 30'' of denaturation at 94 °C, 30'' of primers
194 annealing at 50 °C and 30'' of primers elongation at 72 °C, followed by a final elongation step of
195 10 min at 72 °C. The PCR products of the 2nd step for all samples were multiplexed in a single

196 pool in equimolar amounts on the basis of the QuBit quantification data. The PCR products pool
197 was then purified using the solid phase reverse immobilization (SPRI) method of the Agencourt®
198 AMPure® XP kit (Beckman Coulter, Italy, Milano) and sequenced at Fasteris SA (Geneva,
199 Switzerland). The TruSeq™rDNA sample preparation kit (Illumina Inc., San Diego, CA) was
200 applied for the amplicon library preparation, while the sequencing reaction was performed with a
201 MiSeq Illumina instrument (Illumina Inc, San Diego, CA) with V3 chemistry, generating 300 bp
202 paired-end reads.

203

204 *2.7. Sequences data preparation and analyses*

205 The amplicons pool generated as described in the previous section was sequenced by Fasteris SA
206 (Geneva, Switzerland) using the MiSeq Control Soft. V2.3.0.3, the RTA v1.18.42.0 and the
207 CASAVA v1.8.2 for base calling and Illumina barcode demultiplexing. Since most of the V3-V4
208 amplicons regions are shorter than 500 bp (Vasileiadis et al., 2012) and the sequencing was
209 performed with 300bp x2 paired-reads, the “PANDAseq” script with the restrictions of at least 30
210 bp of overlap between read pairs and two mismatches allowed was applied in order to reconstruct
211 the full amplicons for the majority of sequences. Assembled sequences were then demultiplexed
212 according to the sample indexes and primers using the fastx-toolkit
213 (http://hannonlab.cshl.edu/fastx_toolkit/).

214 Sequences were then analysed with Mothur v.1.33.0 (Schloss et al., 2009) in order to
215 remove: (i) sequences with large homopolymers (>10), (ii) sequences that aligned outside the
216 targeted V3-V4 region, (iii) chimeric amplicons and (iv) sequences that the Bayesian classifier
217 identified as not belonging to Bacteria by using the ribosomal database project (RDP) training set
218 with a 80% bootstrap cutoff.

219 Downstream sequence analyses were then performed using both the operational taxonomic
220 unit (OTU) and the taxonomy based approach. The OTU and the taxonomy matrices were obtained
221 using Mothur v.1.33.0 (Schloss et al., 2009), while all the other statistical analyses were performed

222 in R v 3.1.1. In the OTU approach, sequences were first aligned against the SILVA database
223 bacterial reference alignment (Pruesse et al., 2007) using the NAST algorithm and a kmer search
224 approach (Schloss, 2010). The resulting aligned sequences were then hierarchically clustered into 3
225 % distance defined OTUs using the average linkage algorithm. If not stated differently, OTUs
226 having a sum of their overall abundance across all samples of less than 0.1% were grouped together
227 in a one “rare” OTU group. For the taxonomy based analyses, sequences were hierarchically
228 classified in different taxa using the Greengenes database. A manual amendment of the database
229 was carried out by retrieving on RDP and adding to the database all species level sequences
230 available of type strains (1700 taxonomically annotated sequences in total) belonging to genera that
231 were frequently found in the samples after a first taxonomical analysis of sequences.

232 The composition of PCR products in OTUs and taxonomical groups were used for estimating
233 the associated α - and β - diversity of the analysed samples for the two previously mentioned
234 approaches. Calculated α -diversity indices included the inverse Simpson’s index (D) (Heip et al.,
235 2001) and the observed richness (S), while the Good’s coverage estimate was calculated for
236 assessing the percentage of captured diversity by the devoted sequencing effort (Good, 1953).

237

238 ***2.8. Accession number and data availability.***

239 Raw sequences in read-pairs format were submitted to the National Centre for Biotechnology
240 Information (NCBI) Sequence Read Archive (SRA - Bioproject PRJNA277133).

241

242

243

244 **3. Results**

245 **3.1. Development of a genus specific PCR for DGGE analysis of *Clostridium* cluster I species**

246 A genus specific PCR for the amplification of variable region of 16S rRNA gene was developed in
247 order to analyse the community of clostridia in spoiled cheeses and to identify the species involved
248 in late blowing defect. The choice of the *Clostridium* specific primers was based on the analysis of
249 the 16S rRNA gene sequence of the order *Clostridiales* and other genera included in the Ribosomal
250 Database Project (RDP-II) (Cole et al., 2005). This analysis revealed that members of the
251 *Clostridium* cluster I, differently from related taxonomical units of *Firmicutes*, presented a deletion
252 of about 17bp (positions 583-600) + 15 bp (positions 604-618) in the V3 region of *16S rRNA* gene.
253 This allowed to design DGGE12 (position 480-500 bp) and P4 (position 694-715 bp)
254 oligonucleotides targeted to the sequence upstream and downstream the deletion. *In silico* analysis
255 indicated that the DGGE12-P4 primer couple, is able to amplify almost uniquely 16S rRNA gene
256 from *Clostridium* species. To confirm the specificity of the method, 24 strains belonging to the
257 *Clostridium* cluster I and other 30 strains belonging to *Firmicutes* were tested. The PCR protocol
258 allowed amplification of DNA from clostridial strains but not from strains of other genera. Upon
259 DNA amplification with DGGE12 and P4 primers, each *Clostridium* reference strains used
260 produced a single DGGE band and these bands constituted a normalization ladder.

261

262 **3.2. Detection of clostridial populations in cheese samples by PCR-DGGE analysis**

263 Putative *Clostridium* cluster I-specific amplicons have been obtained for 78 out of 91 (86%) cheese
264 samples. Single band migration patterns were observed for the clostridia type strains used as a
265 marker, thus allowing the identification in cheese samples of *C. butyricum*, *C. sporogenes*, *C.*
266 *tyrobutyricum*, *C. perfringens* and *C. septicum* as the most commonly isolated species. Differently,
267 only one of the eight cheese samples without defects and positive to lysozyme used as negative
268 control, showed a band related to *C. butyricum* in the DGGE analysis. The most indicative bands
269 for each sample were excised from the acrylamide gel and sequenced. These showed a high

270 homology (>99%) with 16S rRNA genes from *Clostridium* species present in databases. Clostridia
271 distribution in cheese samples with blowing defect was mainly affected by the presence or absence
272 of lysozyme in cheeses (Table 2). Rare correlations have been found with the ripening times that
273 varied from 1 to 28 months from the appearance of defect. In general, *C. tyrobutyricum* was the
274 species most frequently found in the total samples, followed by *C. butyricum*, *C. sporogenes*, *C.*
275 *perfringens* and *C. septicum*. In the 56 samples (50 spoiled cheeses and 6 negative controls) where
276 lysozyme was added to milk, we found most frequently *C. butyricum* (50% of the samples)
277 followed by *C. tyrobutyricum* (30%), *C. perfringens* (23%), *C. sporogenes* (21%) and *C. septicum*
278 (5%). On the contrary, in the 35 cheeses (33 spoiled cheeses and 2 negative controls) made without
279 lysozyme the higher clostridial contamination was represented by *C. tyrobutyricum* (77%) that was
280 found to be the dominant species in the majority of samples and, in most cases, it was the only
281 detected species. In addition, *C. butyricum* contamination was observed in 8 samples (22%), *C.*
282 *sporogenes* in 6 samples (17%), *C. perfringens* in 3 samples (8%) and *C. septicum* in only 2
283 samples (5%). In all analysed samples, *C. sporogenes* was never found alone but it appeared in
284 constant association with *C. butyricum* or *C. tyrobutyricum*. Interestingly, mutual exclusive
285 dominance of *C. butyricum* or *C. tyrobutyricum* was observed in all but four of the samples,
286 independently from the ripening time.

287

288 **3.3. Quantitative analysis of *C. tyrobutyricum***

289 Since from PCR-DGGE analysis, *C. tyrobutyricum* appeared to be the dominant species in most of
290 the samples, real-time quantitative PCR was applied for its enumeration in cheeses with late
291 blowing defect (Bassi et al., 2013b). This method was applied on the 91 cheese samples and, 60%
292 of them showed positivity to *C. tyrobutyricum*; in cheese samples manufactured with lysozyme,
293 45% were positive to *C. tyrobutyricum* with counts ranging from 2.0 to 9.2 log CFU/50 g. A higher
294 prevalence of this species was observed in cheese without lysozyme (83% of samples) where only 6
295 samples were negative and counts varying from 2.0 to 8.9 log CFU/50 g (table 2). These

296 observations were confirmed by the statistical analysis of variance that revealed a significant effect
297 of lysozyme ($P<0.01$) on the *C. tyrobutyricum* prevalence. No significant effect of ripening time on
298 *C. tyrobutyricum* counts was observed.

300 **3.4.NGS assessment of bacterial communities in Grana Padano cheese**

301 To obtain a deeper view of the microbiota of spoiled cheese, 20 samples with lysozyme and 20
302 without lysozyme (19 spoiled and one control cheeses in both cases), were randomly selected, the
303 DNA extracted and the V3-V4 region of 16S rRNA genes amplified using universal primers for
304 *Bacteria*. The sequencing of amplicons pooled in equimolar amounts from the analysed samples
305 resulted in 7,539,211 paired-end sequence reads, reduced to 5,045,205 after assembly and
306 demultiplexing on the basis of sample indexes, with an average of 126,130 sequences per sample.
307 The 4.9% of these sequences was discarded because of quality issues (length, presence of
308 homopolymers or ambiguous bases), 0.6% did not align on the V3-V4 targeted region and 1.8%
309 were identified as chimeras. After these screening steps, 4,477,326 sequences were retained and
310 used for further analyses. The taxonomical identification was carried out by comparing the obtained
311 sequences with those included in the GreenGenes database, containing 202,421 aligned reference
312 sequences. The database was manually curated and amended, in order to increase the representation
313 of reference sequences belonging to the genera mostly represented in the cheese samples after a first
314 taxonomical assignment. A total of 1700 type strain 16S full-length sequences were thus
315 downloaded from RDP and added to the database. Alignment to this improved database showed that
316 among 4,477,326 high quality filtered sequences, 97.2% were correctly classified at the genus level
317 and 94.3% at the species level. The number of sequences per samples was downscaled to 32,566
318 sequences per sample, in order to avoid biases related to α -diversity and β -diversity estimations
319 when analysing and comparing samples having unequal size (Gihring et al., 2012; Lundin et al.,
320 2012). A number of richness and diversity indexes were calculated for all samples from OUT data,
321 and results are reported in Supplementary table S1. No significant differences according to ripening

322 time, lysozyme or defect were found for all analysed indexes (data not shown). The Good's
323 coverage showed an average value per sample of 98.2% (standard deviation 0.2%), thus indicating
324 that the even with the reduction of sequences to downscale all samples at 32,566 sequences, most of
325 the bacterial diversity in the cheese samples was still covered. The observed richness (i.e., the total
326 number of OTUs) ranged between 720 and 907 OTUs, with 25 OTUs covering the 99.9% of the
327 total bacterial communities. Chao richness index had an average value of 2901, with 291 of
328 standard deviation, while the evenness, as estimated by Shannon index was 1.78 ± 0.3 .

329 The eight most abundant bacterial genera found in the analysed cheese samples with and
330 without lysozyme, which account for 95% of the total bacterial populations, are reported in Figure
331 1, together with a hierarchical clustering of samples according to the average linkage algorithm.
332 Eight bacterial genera in total, namely *Lactobacillus* (65.3%), *Streptococcus* (14.4%), *Clostridium*
333 (9.54%), *Brevibacterium* (1.5%), *Enterococcus* (0.97%), *Staphylococcus* (0.96%), *Acinetobacter*
334 (0.77%) and *Chryseobacterium* (0.5%), were present at 5% or higher percentages at least in one of
335 the analysed samples. Consistently with cultivation based studies in cheese (Monfredini et al., 2012;
336 Pogačić et al., 2013), most of the sequence of phylum *Firmicutes* belonged to the *Lactobacillus*,
337 which covered in 19 out of 40 samples more than 90% of the total diversity. Sequences assigned to
338 *Clostridium* were detected in all samples, although in different amounts. Specifically this genus
339 covered almost 40% of the detected bacterial diversity in the samples VI616_3-, TN303_6-,
340 BS634_1-, PC544_3+, PC544_7+. The two control samples without late blowing defect (TN303_7-
341 and MN425_6+) contains DNA from clostridia, although in low amount. In general, samples
342 clustering, as depicted in Figure 1, was not affected by either cheese factory, lysozyme use, ripening
343 time or blowing defect. Only the sample CN911_11- presented a complete different bacterial
344 community, as it was dominated by two species *Brevibacterium aureum* and *Staphylococcus*
345 *equorum*. Hierarchical clustering based on the abundance of bacterial genera, revealed three main
346 clusters (fig. 1). Cluster A, composed by 13 samples, was characterized by the highest proportion of
347 sequences assigned to the *Clostridium* genus and a relevant presence of *Streptococcus* sequences,

348 ascribed to *S. thermophilus*. The genus *Lactobacillus* was predominant in the 20 samples of cluster
349 B, a cluster that included also the two negative controls TN303_7- and MN425_6+. This cluster
350 was also characterized by a low number of clostridial sequences. In Cluster C, composed by three
351 samples, more than 50% of the sequences were assigned to *S. thermophilus*. Interestingly,
352 comparing the data achieved by quantitative detection of *C. tyrobutyricum* with the results of this
353 cluster analysis, a correlation between the abundance of *S. thermophilus* and the *C. tyrobutyricum*
354 counts was observed. In the samples characterized by a relevant presence of *S. thermophilus*
355 (cluster A and C), the average value of *C. tyrobutyricum* achieved by real time PCR in cluster C
356 was 6.5 log cells/g, while in cluster A the mean count was 3.4 log cells/g.

357 A detailed analysis on relative abundances of species was then carried out on the two most
358 relevant genera, *Lactobacillus* (Figure 2) and *Clostridium* (Figure 3), being the first the responsible
359 for primary fermentation and ripening process, and the second involved in the insurgences of
360 defects in hard cheese. The diversity of the *Lactobacillus* population in the analysed samples was
361 captured by data achieved with NGS: DNA from *L. helveticus* and *L. delbrueckii*, the dominant
362 species in the primary fermentation of Grana Padano cheese (Neviani et al., 2009; Zago et al.,
363 2011), which reached the highest levels in the first days of ripening and then rapidly declined, were
364 detected in all samples. Different species of non-starter lactic acid bacteria (NSLAB), mainly
365 obligate and facultative heterofermentative species, such as *L. rhamnosus*, *L. casei*, *L. fermentum*,
366 and the obligate heterofermentative *L. buchneri*, were also detected in all analysed samples.
367 Hierarchical clustering of samples according to the average linkage algorithm identified three main
368 clusters on the basis of the *Lactobacillus* species relative abundance (Fig. 2): in cluster Lb-A, *L.*
369 *delbrueckii* was the predominant species, cluster Lb-B was characterized by a relevance of *L.*
370 *rhamnosus*, while *L. casei* was highly represented in cluster Lb-C.

371 Metastats models were employed to assess if the use of lysozyme had significant effects on
372 the relative proportions of *Lactobacillus* (Figure 4a) and *Clostridium* species (Figure 4b). Cheese
373 samples with lysozyme presented significantly higher proportion of *L. delbrueckii*, *L. parabuchneri*

374 and *L. parafarraginis* and significantly lower levels of *L. buchneri*, and *L. paracasei*, when
375 compared with samples without the additive. Moreover, lower numbers of *L. helveticus* and *L.*
376 *rhamnosus* were found in cheese with the lytic enzyme.

377 Seven species of clostridia, namely *C. sporogenes*, *C. butyricum*, *C. disporicum*, *C.*
378 *perfringens*, *C. difficile*, *C. sordelii* and *C. tyrobutyricum* were found in the defected cheeses and,
379 among them, the most abundant were *C. butyricum* and *C. tyrobutyricum*. When hierarchical cluster
380 analysis (Fig. 3) was applied to the clostridial community, three clusters were defined mainly of the
381 ratio between the two species *C. butyricum* and *C. tyrobutyricum*. Thus, *C. butyricum* was highly
382 predominant in cluster CI-A, cluster CI-B contained samples characterized by similar amounts of
383 sequences assigned to *C. butyricum* and *C. tyrobutyricum*, while cluster in CI-C, more than 80% of
384 the analysed sequences were identified as *C. tyrobutyricum*. Other clostridial species identified in
385 the analysed cheese samples were detected in lower percentages. Consistently with the data
386 achieved using the DGGE analysis, the presence of lysozyme in the cheese paste significantly
387 influenced the relative presence of *C. butyricum* and *C. tyrobutyricum*, as shown by the Metastats
388 analysis reported in fig 4b. In the absence of this lytic enzyme, *C. tyrobutyricum* dominated the
389 clostridial community, accounting for more than 50% of the DNA sequences assigned to this genus.
390 When lysozyme was added as preservative, a shift in the dominance was observed and *C. butyricum*
391 was the most present species.

392

393 **4. Discussion**

394 A crucial step for the reduction of cheese blowing defects caused by clostridia is to investigate the
395 community of these anaerobic sporeformers cause of spoilage and their ecological relationships
396 with the other members of the cheese microbiota (Doyle et al., 2015). Many factors may affect
397 together or independently the ripening process of Grana Padano cheese, such as the dairy
398 technology, the use of autochthonous starters and additives, the environmental parameters (Bittante
399 et al., 2011; Neviani and Gatti, 2013); moreover, in the case of Grana Padano cheese, its size

400 (approximately 40 kg) and long ripening times (up to 24 months) make difficult the understanding
401 of bacterial dynamics, including those of clostridia, in a so complex food matrix. In general the
402 quality of cheese is mostly dependent on the LABs fermentations, that influence its sensorial and
403 aromatic features (Steele et al., 2013); therefore, the understanding of interactions, still lacking,
404 between these last and clostridia might be of help to better clarify the blowing defect.

405 In order to provide an updated insight of the major clostridial species and the other microbial
406 populations affecting the cheese environment, 83 GP cheese samples removed from the ripening
407 process, due to the presence of the typical signs, eyes, fractures and blowing, were studied using
408 DNA-based approaches. Moreover, eight cheeses without defects were considered in the study as
409 negative controls. The majority of cheeses were produced using egg-white lysozyme as food
410 additive; this muramidase is intentionally added to milk to inhibit clostridia in the outgrowth step
411 after germination, with the aim of preventing the late blowing defects of hard cheese (Wasserfall
412 and Teuber, 1979). The culture-independent analysis was carried out using three different
413 approaches, a *Clostridium* cluster-I specific DGGE, a quantitative qPCR enumeration of *C.*
414 *tyrobutyricum* and a next-generation Illumina-based sequencing on 16S rRNA gene, which
415 provided firstly a qualitative and then a quantitative picture of the entire bacterial community of
416 spoiled cheese.

417 The number of OTUs determined in the cheese samples was high, with an average of 804
418 (Table S1). Other studies relying on NGS analyses for cheese revealed maximum OTU values of
419 192 for Mozzarella (Ercolini et al., 2012), of 73 OTUs for Parmigiano Reggiano and 64 for Grana
420 Padano (De Filippis et al., 2014). This difference in outcome may be explained by the fact that our
421 study was based on Illumina sequencing technology, while the ones here cited on 454
422 pyrosequencing technology, which has a lower throughput. The high coverage values that we
423 obtained, together with a an assignment at the species level of 94.3% of all analysed sequences,
424 prove that Illumina NGS of 16S amplicons is a reliable approach for a thorough analyses of cheese
425 bacterial communities, as already demonstrated for fermented meats (Połka et al., 2015) and

426 beverages (Bokulich et al., 2012). It must also be stressed out that, among all the OTUs retrieved,
427 just a fraction of maximum 26 OTUs covers the 99.9% of the total bacterial community. This result
428 points to a major microbiome constituted by a few selected groups (Figure 1), and a complex rare
429 microbiome whose possible role in ripening and final cheese characteristics still needs to be
430 addressed. The diversity and richness indexes calculated for our samples were instead in
431 agreement with published reports (De Filippis et al., 2014; Ercolini et al., 2012).

432 The advantage of the PCR-DGGE approach selectively amplifying *Clostridiales* members
433 was the possibility to analyse almost exclusively species belonging to the cluster I of the
434 *Clostridium* genus, reducing the interference of other bacteria, such as lactic acid bacteria, which
435 dominate the microbial communities of dairy products. *Clostridium* cluster-I is a wide taxonomic
436 group which contains all the clostridial species identified as agents of cheese spoilage, such as *C.*
437 *beijerinckii*, *C. butyricum*, *C. sporogenes*, *C. tyrobutyricum* (Cocolin et al., 2004; Ingham et al.,
438 1998; Le Bourhis et al., 2007b). Moreover, this approach reduced the need of additional steps, such
439 as pasteurization and selective medium enrichment, previously applied to the study of late blowing
440 in hard cheese (Cocolin et al., 2004), which may introduce biases in the analysis of the clostridial
441 communities. Results obtained with *Clostridium* cluster I-specific PCR-DGGE analysis provided a
442 qualitative picture of the dominant clostridia biodiversity and suggested that clostridia, in case of
443 blowing defect, are nearly ubiquitous members of the GP cheese ecosystem and are strictly related
444 to the spoilage event. In general, outputs of HTS data regarding clostridial communities were in
445 accordance with those obtained by PCR-DGGE analysis even targeting different 16S hyper-variable
446 regions and with a major resolution and quantitative power for the NGS approach. The major
447 performance of NGS as compared to PCR-DGGE was recently proven for Italian dry-fermented
448 sausages (Połka et al., 2015) and for Oscypek, a traditional Polish smoked cheese (Alegría et al.,
449 2012). Analyses of NGS data defined three main clusters, the first (CI-A) characterized by the
450 dominant presence of *C. butyricum*, the second (CI-B) by a co-presence of *C. butyricum* and *C.*
451 *tyrobutyricum* and the third (CI-C) by a higher prevalence of *C. tyrobutyricum* (Fig. 3). These data

452 indicated that late blowing defect in this hard cheese was primarily caused by this two butyric spore
453 formers. Moreover, although the two analytical methods were not totally comparable, the
454 quantitative data obtained from qPCR on *C. tyrobutyricum* were in accordance with the results of
455 NGS analysis. Thus, samples belonging to the dominant *C. tyrobutyricum* cluster (fig. 3, Cl-C)
456 presented an average value of 7.45 Log target genome/50g, higher than the other two clusters,
457 corresponding to 4.40 Log target genome/50g and 2.86 Log target genome/50g for Cl-A and Cl-B
458 respectively.

459 *Clostridium* cluster I-specific PCR-DGGE data demonstrated that *C. tyrobutyricum*,
460 previously described as the main responsible of hard cheese spoilage (Klijn et al., 1995; Le Bourhis
461 et al., 2005), was the most frequent species, hosted in the 50% of total samples. According to the
462 literature, *C. tyrobutyricum* is commonly associated with late blowing defects appearing in cheese
463 after 6-8 months of ripening (Klijn et al., 1995; Le Bourhis et al., 2005), while *C. butyricum* is most
464 frequently involved in early blowing (Bottazzi, 2001). However, in our samples we found no
465 significant correlations between the ripening time and the prevalence of a particular *Clostridium*
466 species. It has been demonstrated (Le Bourhis et al., 2007a) that the presence of other clostridia,
467 such as *C. sporogenes* and *C. beijerinckii*, synergistically stimulate the germination and growth of
468 *C. tyrobutyricum* in the early stages of Emmental type cheese. Differently, our analysis of Grana
469 Padano cheese samples showed that that the simultaneous presence of *C. sporogenes* and *C.*
470 *tyrobutyricum* did not induce an early insurgence of the defect.

471 Lysozyme, rather, seemed to influence bacterial distribution both in terms of *Clostridium* and
472 *Lactobacillus*, the two most abundant genera, together with *Streptococcus*, found in the analysed
473 cheeses. Among clostridia, *C. tyrobutyricum* was negatively affected by lysozyme, as demonstrated
474 by DGGE, NGS data and by quantitative results obtained with species-specific TaqMan real-time
475 method (Bassi et al., 2013b). In particular, the quantitative data achieved by real-time PCR, which
476 indicated a significant higher number of *C. tyrobutyricum* in absence than in presence of lysozyme
477 (4.6 Log target genome/50g, vs 2.5 Log target genome/50g) (Table 2), were confirmed by the

478 Metastats modelling of NGS data at genus level (Figure 4b). Recent experiments in milk and RCM
479 medium proved that lysozyme was particularly effective in limiting *C. tyrobutyricum* cells and
480 spores growth respect to other clostridia (Ávila et al., 2014). Otherwise, *C. butyricum* resulted the
481 most prevalent species when lysozyme was added to milk, as shown by DGGE data (Table 2) and
482 confirmed by the relative abundance analysis of 16S rRNA gene sequence (Figure 3). This
483 observation confirmed the *in vitro* data (Ávila et al., 2014) of high resistance of *C. butyricum* to this
484 additive. *C. butyricum* has been rarely associated to late blowing defects in hard cheese using
485 cultivation based approaches and non-quantitative PCR methods: Kljin et al. (1995) did not detect
486 this species in samples of Gouda cheese by PCR, while Cocolin et al. (2004) found this species in
487 only 2 out of 17 samples by DGGE. Our data, based on both genus specific DGGE and NGS
488 approaches revealed that *C. butyricum* is a common component of the clostridial population
489 involved in late blowing of hard cheese. *C. sporogenes*, which was detected in approximately 20%
490 of the analysed samples, always associated to *C. butyricum* or *C. tyrobutyricum* and present in low
491 amounts, seemed to be independent on lysozyme addition in both DGGE and NGS analyses. In the
492 same way, Garde et al. (2011) demonstrated that in spoiled Manchego cheese samples with
493 lysozyme *C. sporogenes* was the dominant species. This observation reflects a diversity in the
494 clostridia community of different hard cheeses. Among the other clostridia species detected in
495 lower amounts, *C. perfringens* too was more prevalent in cheese samples with lysozyme. The
496 presence of this species does not rise safety concern; indeed, food-borne disease caused by this
497 pathogenic bacterium is related to the ingestion of 10^8 vegetative cells of the enterotoxin producing
498 strains and only a limited fraction of strains produce enterotoxins. Moreover, no cases of outbreaks
499 have been described from cheese and this organism is widely present in a variety of foodstuffs
500 (EFSA, 2005). Shifts in the *Lactobacillus* community were also observed in the presence of
501 lysozyme for both bacterial species of the primary lactose fermentation and NSLABs. An increase
502 in the proportion of *L. delbrueckii* and obligate heterofermentative NSLABs was detected in the

503 presence of added muramidase, while in its absence *L. helveticus* and other NSLABs, such as *L.*
504 *rhamnosus*, *L. casei* and *L. buchneri*, were more abundant (Figure 4a).

505 When the quantitative NGS data of lactic acid bacteria, *Lactobacillus* and *Streptococcus*, and
506 of clostridia were analysed by hierarchical clustering, several ecological relationships were found.
507 Most of the samples from cluster C1-C, characterized by a high abundance of *C. tyrobutyricum*,
508 showed also the dominance of *S. thermophilus* and *L. rhamnosus*. These data were confirmed by the
509 qPCR analysis of *C. tyrobutyricum* that demonstrated higher amounts of this species in all samples
510 of this cluster. A possible relationship between *S. thermophilus* and *C. tyrobutyricum* could be
511 explained by the effect on germination produced by L(+) lactate, the stereoisomer of lactate
512 produced in cheese. This end product of lactose metabolism by *S. thermophilus* alone or in
513 association with amino acids, such as alanine, is the most effective germinant for *C. tyrobutyricum*
514 (Bassi et al., 2009) and other clostridia such as *C. botulinum* (Peck et al., 2011). *L. rhamnosus*, was
515 found as dominant species in samples of cluster C1-C, in accordance to other study that showed in
516 GP an increase in number over ripening up to the sixth month (Pogačić et al., 2013). Thus, this L(+) lactate
517 producer might further stimulate the germination of *C. tyrobutyricum*. Differently from other
518 clostridia detected in cheese, *C. tyrobutyricum* is able to use lactate as energy source. This species
519 metabolises both D(-) and L(+) racemic forms of lactic acid, although D(-) lactate acid is more
520 rapidly dissimilated to butyric acid when cells grow in RCM medium (Huchet et al., 1997). These
521 authors suggested that the use of *Lactobacillus* producing L(+) in cheese fermentation may reduce
522 the growth rate of *C. tyrobutyricum* and the spoilage of Emmentaler cheese. Differently our data
523 indicate that *C. tyrobutyricum* is most abundant when there is a dominance of *S. thermophilus* and
524 *L. rhamnosus*, two L(+) lactate producers. The key enzymes of this metabolism are NAD
525 independent lactate dehydrogenases, which convert lactate to pyruvate. The analysis of the *C.*
526 *tyrobutyricum* genomes (Bassi et al., 2013c; Jiang et al., 2013) revealed the presence of genes
527 coding for both D and L lactate dehydrogenases and of a L-lactate permease. It is still not clear if
528 the ratio between the two racemic forms affect the vegetative cycle of *C. tyrobutyricum* in cheese.

529 All the samples dominated by *C. butyricum* (cluster Cl-A) were also clustered in Lb-A, where *L.*
530 *delbrueckii* was the most abundant *Lactobacillus*. Based on the energetic metabolism, it cannot be
531 explained why the presence of *L. delbrueckii* relates to a higher prevalence of *C. butyricum*.

532 It is now well recognised that bacteria interact in the food communities, influencing the final
533 quality of fermented products. Our data indicated that in hard cheese presenting the typical signs of
534 late blowing, the spoiling clostridial species coexisted with the LAB populations and often they did
535 not dominate in quantitative terms the cheese microbiota. Some preliminary hints deriving from the
536 NGS analysis of the bacterial community indicated a correlation between the inhabitant LAB, both
537 natural starters and NSLAB, and the prevalence of different species of clostridia. Moreover, our
538 study showed that the use of lysozyme, added to affect spore germination and their vegetative cell
539 outgrowth, shaped the species composition of the cheese bacterial communities of both LAB and
540 butyric clostridia. Further studies aimed to analyse the interactions LAB-clostridia might provide in
541 the future a deeper knowledge of the ecological relationships in cheese microbiota and new
542 approaches for the reduction of hard cheese spoilage.

543

544

545 **Acknowledgements**

546 The research was supported by grants from the Ministero delle politiche agricole, alimentari e
547 forestali (MIPAAF) national project “Filigrana” D.M. 25741/7303/11, Consorzio per la Tutela del
548 Formaggio Grana Padano, Italy and Regione Lombardia ERSAF national project “GPL-free”
549 D.11/08/2011 n° 7656. The authors are grateful to Sotirios Vasileiadis for the designing of the
550 primer-linker-index constructs and for the setup of several data analysis methods used for the
551 elaboration of the Illumina sequences.

552

553

554

555 **Figure Legends**

556 **Table 1.** Specificity of the *Clostridiales* PCR amplification using a range of clostridial and non-
557 clostridial species.

558 **Table 2.** Data obtained from *Clostridium* cluster I-specific PCR-DGGE analysis and qPCR for *C.*
559 *tyrobutyricum* enumeration in cheese samples with and without added lysozyme.

560 **Figure 1.** Hierarchical clustering of classified sequences using the average linkage algorithm at
561 genus classification level for taxa participating with ≥ 5 % in at least one sample. Taxa with lower
562 participations were added to the “other” sequence group. Three main clusters have been defined
563 (A, B and C).

564 **Figure 2.** Hierarchical clustering of classified sequences using the average linkage algorithm at the
565 species classification level limited to sequences belonging to *Lactobacillus* genus. Taxa with
566 participations lower than 1% were added to the “other” sequence group. Three main clusters have
567 been defined with Lb-A, Lb-B and Lb-C.

568 **Figure 3.** Hierarchical clustering of classified sequences using the average linkage algorithm at the
569 species classification level limited to sequences belonging to *Clostridium* genus. Taxa with
570 participations lower than 1% were added to the “other” sequence group. Three main clusters have
571 been defined with Cl-A, Cl-B and Cl-C.

572 **Figure 4.** Metastats model assessing the effects of lysozyme on the relative abundances of
573 *Lactobacillus* (a) and *Clostridium* sequences (b). Genera showing significant differences are
574 highlighted with an asterisk.

575 **Table S1.** Number of sequences per sample, coverage and diversity indexes calculated on the
576 1,204,942 high quality sequences analyzed in the casing samples.

577

578

579

580

581

582

583

584

585

REFERENCES

- 586 Alegría, Á., Szczesny, P., Mayo, B., Bardowski, J., Kowalczyk, M., 2012. Biodiversity in Oscypek,
587 a traditional Polish cheese, determined by culture-dependent and-independent approaches.
588 Appl Environ Microb 78, 1890-1898.
- 589 Ávila, M., Gómez-Torres, N., Hernández, M., Garde, S., 2014. Inhibitory activity of reuterin,
590 nisin, lysozyme and nitrite against vegetative cells and spores of dairy-related *Clostridium*
591 species. Int J Food Microbiol 172, 70-75.
- 592 Bassi, D., Cappa, F., Cocconcelli, P.S., 2009. A combination of a SEM technique and X-ray
593 microanalysis for studying the spore germination process of *Clostridium tyrobutyricum*. Res
594 Microbiol 160, 322-329.
- 595 Bassi, D., Cappa, F., Cocconcelli, P.S., 2013a. Array-based transcriptional analysis of
596 *Clostridium sporogenes* UC9000 during germination, cell outgrowth and vegetative life. Food
597 Microbiol 33, 11-23.
- 598 Bassi, D., Fontana, C., Zucchelli, S., Gazzola, S., Cocconcelli, P.S., 2013b. TaqMan real time-
599 quantitative PCR targeting the phosphotransacetylase gene for *Clostridium tyrobutyricum*
600 quantification in animal feed, faeces, milk and cheese. Int Dairy J 33, 75-82.
- 601 Bassi D, Fontana C, Gazzola S, Pietta E, Puglisi E, Cappa F, Cocconcelli PS. Zhu L, 2013c. Draft
602 genome sequence of *Clostridium tyrobutyricum* UC7086, isolated from Grana Padano cheese
603 with late-blowing defect. Genome A 1:e00614-00614.
- 604 Bergère, J.L., Sivelä, S., 1990. Detection and enumeration of clostridial spores related to cheese
605 quality-classical and new methods. Bull Int Dairy Fed, 18-23.
- 606 Berry, D., Mahfoudh, K.B., Wagner, M., Loy, A., 2011. Barcoded primers used in multiplex
607 amplicon pyrosequencing bias amplification. Appl Environ Microb 77, 7846-7849.
- 608 Bittante, G., Cologna, N., Cecchinato, A., De Marchi, M., Penasa, M., Tiezzi, F., Endrizzi, I.,
609 Gasperi, F., 2011. Monitoring of sensory attributes used in the quality payment system of
610 Trentingrana cheese. J Dairy Sci 94, 5699-5709.
- 611 Bokulich, N.A., Bamforth, C.W., Mills, D.A., 2012. Brewhouse-resident microbiota are
612 responsible for multi-stage fermentation of American coolship ale. PLoS ONE 7, e35507.
- 613 Bottazzi, V., 2001. Gonfiore precoce del formaggio Grana da fermentazione butirrica. Scienza e
614 Tecnica Lattiero-Casearia 52, 339-350.
- 615 Cocolin, L., Innocente, N., Biasutti, M., Comi, G., 2004. The late blowing in cheese: a new
616 molecular approach based on PCR and DGGE to study the microbial ecology of the alteration
617 process. Int J Food Microbiol 90, 83-91.
- 618 Coulon, J.B., Varignier, M., Darne, D., 1991. Contamination butyrique du lait de vache: étude
619 dans les exploitations de Haute-Loire (1). Inra Productions Animales 4, 369-372.
- 620 De Filippis, F., La Storia, A., Stellato, G., Gatti, M., Ercolini, D., 2014. A selected core microbiome
621 drives the early stages of three popular Italian cheese manufactures. PLoS ONE 9, e89680.
- 622 Devoyod, J.J., 1975. [Use of nitrates in cheese-making], Ann Nutr Aliment, pp. 789-792.
- 623 Doyle, C.J., Gleeson, D., Jordan, K., Beresford, T.P., Ross, R.P., Fitzgerald, G.F., Cotter, P.D., 2015.
624 Anaerobic sporeformers and their significance with respect to milk and dairy products. Int J
625 Food Microbiol 197, 77-87.
- 626 EFSA, 2005. Opinion of the Scientific Panel on Biological Hazards on the request from the
627 Commission related to *Clostridium* spp in foodstuffs. The EFSA Journal 199, 1-65.

628 EFSA, 2010. Opinion of the scientific panel on food additives and nutrient sources added to
629 food on the request from the European Commission related to scientific evidence for a
630 revision of the maximum limits on nitrites in food adopted in Directive 2006/52/EC. Question
631 number: EFSA-Q-2010-00097. Adopted on 11 March 2010. . The EFSA Journal 8.

632 Ercolini, D., De Filippis, F., La Stora, A., Iacono, M., 2012. "Remake" by high-throughput
633 sequencing of the microbiota involved in the production of water buffalo mozzarella cheese.
634 Appl Environ Microb 78, 8142-8145.

635 Garde, S., Arias, R., Gaya, P., Nuñez, M., 2011. Occurrence of *Clostridium* spp. in ovine milk and
636 Manchego cheese with late blowing defect: identification and characterization of isolates. Int
637 Dairy J 21, 272-278.

638 Gihring, T.M., Green, S.J., Schadt, C.W., 2012. Massively parallel rRNA gene sequencing
639 exacerbates the potential for biased community diversity comparisons due to variable library
640 sizes. Environ Microbiol 14, 285-290.

641 Good, I.J., 1953. The population frequencies of species and the estimation of population
642 parameters. Biometrika 40, 237-264.

643 Guericke, S., 1993. Laktatvergarende Clostridien bei der Kaseherstellung. Dtsch Milchwirtsch
644 48, 735-739.

645 Heip, C.H.R., Herman, P.M.J., Soetaert, K., 2001. Indices of Diversity and Evenness. Océanis 24,
646 61-87.

647 Huchet, V., Thuault, D., Bourgeois, C.M., 1997. The stereoselectivity of the use of lactic acid by
648 *Clostridium tyrobutyricum*. Food Microbiol 14, 227-230.

649 Ingham, S.C., Hassler, J.R., Tsai, Y.-W., Ingham, B.H., 1998. Differentiation of lactate-fermenting,
650 gas-producing *Clostridium* spp. isolated from milk. Int J Food Microbiol 43, 173-183.

651 Jiang, L., Zhu, L., Xu, X., Li, Y., Li, S., Huang, H., 2013. Genome sequence of *Clostridium*
652 *tyrobutyricum* ATCC 25755, a butyric acid-overproducing strain. GenomeA 1, e00308-00313.

653 Kiyohara, M., Koyanagi, T., Matsui, H., Yamamoto, K., Take, H., Katsuyama, Y., Tsuji, A.,
654 Miyamae, H., Kondo, T., Nakamura, S., 2012. Changes in microbiota population during
655 fermentation of narezushi as revealed by pyrosequencing analysis. Biosci Biotech Bioch 76,
656 48-52.

657 Klijn, N., Nieuwenhof, F.F., Hoolwerf, J.D., Van Der Waals, C.B., Weerkamp, A.H., 1995.
658 Identification of *Clostridium tyrobutyricum* as the causative agent of late blowing in cheese by
659 species-specific PCR amplification. Appl Environmental Microb 61, 2919-2924.

660 Korenekova, B., Kottferova, J., Korenek, M., 2000. The fate of added nitrate used in the
661 manufacture of Emmental cheese. Food Addit Contam 17, 373-377.

662 Koyanagi, T., Kiyohara, M., Matsui, H., Yamamoto, K., Kondo, T., Katayama, T., Kumagai, H.,
663 2011. Pyrosequencing survey of the microbial diversity of 'narezushi', an archetype of modern
664 Japanese sushi. Lett Appl Microbiol 53, 635-640.

665 Le Bourhis, A.-G., Doré, J., Carlier, J.-P., Chamba, J.-F., Popoff, M.-R., Tholozan, J.-L., 2007a.
666 Contribution of *C. beijerinckii* and *C. sporogenes* in association with *C. tyrobutyricum* to the
667 butyric fermentation in Emmental type cheese. Int J Food Microbiol 113, 154-163.

668 Le Bourhis, A.-G., Doré, J., Carlier, J.-P., Chamba, J.-F., Popoff, M.-R., Tholozan, J.-L., 2007b.
669 Contribution of *C. beijerinckii* and *C. sporogenes* in association with *C. tyrobutyricum* to the
670 butyric fermentation in Emmental type cheese. Int J Food Microbiol 113, 154-163.

671 Le Bourhis, A.-G., Saunier, K., Doré, J., Carlier, J.-P., Chamba, J.-F., Popoff, M.-R., Tholozan, J.-L.,
672 2005. Development and validation of PCR primers to assess the diversity of *Clostridium* spp. in
673 cheese by temporal temperature gradient gel electrophoresis. Appl Environ Microb 71, 29-38.

674 Lodi, R., Stadhouders, J., 1990. The use of lysozyme to control butyric acid fermentation. Bull
675 Int Dairy Fed, 51-54.

676 Lundin, D., Severin, I., Logue, J.B., Östman, Ö., Andersson, A.F., Lindström, E.S., 2012. Which
677 sequencing depth is sufficient to describe patterns in bacterial α - and β -diversity? *Environ*
678 *Microbiol Rep* 4, 367-372.

679 Lycken, L., Borch, E., 2006. Characterization of *Clostridium* spp. isolated from spoiled
680 processed cheese products. *J Food Protect* 69, 1887-1891.

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

685 Masoud, W., Vogensen, F.K., Lillevang, S., Abu Al-Soud, W., Sørensen, S.J., Jakobsen, M., 2012.
686 The fate of indigenous microbiota, starter cultures, *Escherichia coli*, *Listeria innocua* and
687 *Staphylococcus aureus* in Danish raw milk and cheeses determined by pyrosequencing and
688 quantitative real time (qRT)-PCR. *Int J Food Microbiol* 153, 192-202.

689 Monfredini, L., Settanni, L., Poznanski, E., Cavazza, A., Franciosi, E., 2012. The spatial
690 distribution of bacteria in Grana-cheese during ripening. *Syst Appl Microbiol* 35, 54-63.

691 Neviani, E., De Dea Lindner, J., Bernini, V., Gatti, M., 2009. Lysozyme resistance of lactic acid
692 bacteria. *Latte* 26, 90-91.

693 Neviani, E., Gatti, M., 2013. Microbial evolution in raw-milk, long-ripened cheeses: Grana
694 Padano and Parmigiano Reggiano case, in: Randazzo, C.L., Caggia, C., Neviani, E. (Eds.), *Cheese*
695 *Ripening: Quality, Safety and Health Aspects*, New York, pp. 133-148.

696 Nishihara, M., Takahashi, H., Sudo, T., Kyoi, D., Kawahara, T., Ikeuchi, Y., Fujita, T., Kuda, T.,
697 Kimura, B., Yanahira, S., 2014. Multilocus variable-number of tandem repeat analysis (MLVA)
698 for *Clostridium tyrobutyricum* strains isolated from cheese production environment. *Int J Food*
699 *Microbiol* 190, 61-65.

700 Peck, M.W., Stringer, S.C., Carter, A.T., 2011. *Clostridium botulinum* in the post-genomic era.
701 *Food Microbiol* 28, 183-191.

702 Pogačić, T., Mancini, A., Santarelli, M., Bottari, B., Lazzi, C., Neviani, E., Gatti, M., 2013. Diversity
703 and dynamic of lactic acid bacteria strains during aging of a long ripened hard cheese
704 produced from raw milk and undefined natural starter. *Food Microbiol* 36, 207-215.

705 Połka, J., Rebecchi, A., Pisacane, V., Morelli, L., Puglisi, E., 2015. Bacterial diversity in typical
706 Italian salami at different ripening stages as revealed by high-throughput sequencing of 16S
707 rRNA amplicons. *Food Microbiol* 46, 342-356.

708 Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glöckner, F.O., 2007.
709 SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA
710 sequence data compatible with ARB. *Nucleic Acid Res* 35, 7188-7196.

711 Reindl, A., Dzieciol, M., Hein, I., Wagner, M., Zangerl, P., 2014. Enumeration of clostridia in goat
712 milk using an optimized membrane filtration technique. *J Dairy Sci* 97, 6036-6045.

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

716 Schloss, P.D., 2010. The effects of alignment quality, distance calculation method, sequence
717 filtering, and region on the analysis of 16S rRNA gene-based studies. *PLOS Comput Biol* 6,
718 e1000844.

719 Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A.,
720 Oakley, B.B., Parks, D.H., Robinson, C.J., 2009. Introducing mothur: open-source, platform-
721 independent, community-supported software for describing and comparing microbial
722 communities. *Appl Environ Microb* 75, 7537-7541.

723 Stadhouders, J., 1990. Prevention of butyric acid fermentation by the use of nitrate. *Bull Int*
724 *Dairy Fed*, 40-46.

725 Steele, J., Broadbent, J., Kok, J., 2013. Perspectives on the contribution of lactic acid bacteria to
726 cheese flavor development. *Curr Opin Biotech* 24, 135-141.

727 Vasileiadis, S., Puglisi, E., Arena, M., Cappa, F., Cocconcelli, P.S., Trevisan, M., 2012. Soil
728 Bacterial Diversity Screening Using Single 16S rRNA Gene V Regions Coupled with Multi-
729 Million Read Generating Sequencing Technologies. *PLoS ONE* 7.

730 Vissers, M., 2007. Modeling to control spores in raw milk. Wageningen Universiteit.

731 Wasserfall, F., Teuber, M., 1979. Action of egg white lysozyme on *Clostridium tyrobutyricum*.
732 *Appl Environ Microb* 38, 197-199.

733 Zago, M., Fornasari, M.E., Carminati, D., Burns, P., Suárez, V., Vinderola, G., Reinheimer, J.,
734 Giraffa, G., 2011. Characterization and probiotic potential of *Lactobacillus plantarum* strains
735 isolated from cheeses. *Food Microbiol* 28, 1033-1040.

736

737

738

Table 1

Bacterial strains	Isolate number^a	Source	<i>Clostridiales</i>-specific PCR amplification
<i>Bacillus cereus</i>	UC9027	Cheese	negative
<i>Bacillus coagulans</i>	UC9021	Cheese	negative
<i>Bacillus coagulans</i>	UC9021	GP Cheese	negative
<i>Bacillus sp.</i>	UC9024	Cheese	negative
<i>Bacillus subtilis</i>	UC9028	Cheese	negative
<i>Bacillus thuringiensis/cereus</i>	UC8553	GP cheese	negative
<i>Clostridium cochlearium</i>	DSM1285T	Type strain	positive
<i>Clostridium acetobutylicum</i>	LMG5710T	Type strain	positive
<i>Clostridium botulinum</i>	UC9031	Animal feed	positive
<i>Clostridium butyricum</i>	DSM10702 ^T	Type strain	positive
<i>Clostridium butyricum</i>	UC9041	Cheese	positive
<i>Clostridium butyricum</i>	UC9045	GP cheese	positive
<i>Clostridium perfringens</i>	UC9042	Silage	positive
<i>Clostridium perfringens</i>	UC9043	Silage	positive
<i>Clostridium sporogenes</i>	ATCC7955 ^T	Type strain	positive
<i>Clostridium sporogenes</i>	UC9015	Milk	positive
<i>Clostridium tyrobutyricum</i>	DSM2637 ^T	Type strain	positive
<i>Clostridium tyrobutyricum</i>	UC7086	GP Cheese	positive
<i>Clostridium tyrobutyricum</i>	ATCC25755 ^T	Type strain	positive
<i>Clostridium tyrobutyricum</i>	UC9030	GP Cheese	positive
<i>Clostridium tyrobutyricum</i>	UC9036	GP Cheese	positive
<i>Clostridium tyrobutyricum</i>	UC9037	GP Cheese	positive
<i>Clostridium tyrobutyricum</i>	UC9014	GP Cheese	positive
<i>Clostridium tyrobutyricum</i>	UC9044	Milk	positive
<i>Clostridium tyrobutyricum</i>	UC8981	Milk	positive
<i>Clostridium tyrobutyricum</i>	UC8984	Ricotta cheese	positive
<i>Clostridium tyrobutyricum</i>	UC8988	Ricotta cheese	positive
<i>Clostridium tyrobutyricum</i>	UC9008	GP cheese surface	positive
<i>Clostridium tyrobutyricum</i>	UC9038	GP cheese	positive

<i>Clostridium tyrobutyricum</i>	UC9040	GP cheese	positive
<i>Enterococcus faecalis</i>	UC8823	Sausage	negative
<i>Enterococcus faecium</i>	UC8821	Sausage	negative
<i>Enterococcus faecium</i>	UC9029	Cheese	negative
<i>Lactobacillus acidophilus</i>	UC10110	Cheese	negative
<i>Lactobacillus curvatus</i>	UC8266	Taleggio cheese	negative
<i>Lactobacillus helveticus</i>	UC8900	Cheese	negative
<i>Lactobacillus helveticus</i>	UC10206	GP cheese	negative
<i>Lactobacillus plantarum</i>	UC10119	Cheese	negative
<i>Lactobacillus rhamnosus</i>	UC8509	GP cheese	negative
<i>Lactobacillus rhamnosus</i>	UC8514	GP cheese	negative
<i>Lactobacillus sakei</i>	UC8705	Sausage	negative
<i>Lactobacillus sakei</i>	UC8168	Casera cheese	negative
<i>Leuconostoc mesenteroides</i>	UC8133	Scimudin Cheese	negative
<i>Leuconostoc mesenteroides</i>	UC8182	Casera cheese	negative
<i>Listeria monocytogenes</i>	UC8159	Gorgonzola cheese	negative
<i>Ruminococcus flavefaciens</i>	ATCC19208 ^T	Type strain	negative
<i>Ruminococcus gnavus</i>	ATCC29149 ^T	Type strain	negative
<i>Staphylococcus equorum</i>	UC8190	Scimudin cheese	negative
<i>Staphylococcus xylosus</i>	UC8449	Salami	negative
<i>Staphylococcus saprophyticus</i>	UC8555	GP cheese	negative
<i>Streptococcus macedonicus</i>	UC9022	Cheese	negative
<i>Streptococcus thermophilus</i>	UC10185	GP cheese	negative
<i>Streptococcus thermophilus</i>	UC8547	milk	negative
<i>Streptococcus thermophilus</i>	UC10166	GP cheese	negative

^a UC, Università Cattolica del Sacro Cuore culture collection; ATCC, American Type Culture Collection; DSM, German Collection of Microorganisms.

Table 2

LYSOZYME POSITIVE SAMPLES						
Sample	<i>C. butyricum</i>	<i>C. sporogenes</i>	<i>C. perfringens</i>	<i>C. septicum</i>	<i>C. tyrobutyricum</i>	TaqMan Log target genome/50g <i>C. tyr</i>
BS603_28+						0.00
BS607_7+	+	+				0.00
BS607_8+	+	+	+			0.00
BS619_10+a	+					0.00
BS619_10+b			+			0.00
BS619_10+c			+			0.00
BS623_8+a	+					0.00
BS623_8+b	+					0.00
BS628_2+a	+	+	+			0.00
BS628_2+b	+	+				0.00
MN427_7+	+			+	+	0.00
MN437_6+	+	+			+	0.00
MN449_4+						0.00
MN449_7+	+	+	+			0.00
MN460_5+					+	0.00
MN460_8+	+	+	+			0.00
MN460_9+		+				0.00
MN481_11+a			+			0.00
MN481_11+b						0.00
MN481_14+					+	0.00
PC508_7+	+					0.00
PC540_5+						0.00
PC540_6+						0.00
PC544_7+	+					0.00
PC544_8+	+					0.00
BS627_5+						0.00
CN903_6+						0.00
MN425_6+						0.00
MN443_7+						0.00
PC521_6+	+					0.00
PC544_5+						0.00
MN449_5+			+		+	2.00
MN449_8+			+		+	2.04
BS628_6+			+		+	3.70
MN460_8+		+			+	4.49
PC508_8+	+					4.56
BG506_4+	+					4.60
PC508_11+	+		+			4.87
PC544_7+	+		+			4.94
PC544_6+	+			+		4.96
MN437_5+	+	+			+	4.98
BS634_1+	+	+			+	5.00
PC544_3+	+					5.11

VI618_8+						5.15
BG506_12+	+					5.38
BS618_10+a	+					5.43
BS618_10+b				+		5.48
CR101_10+b			+		+	5.66
PC508_7+	+	+				5.88
CR101_10+a	+					6.54
PC508_9+	+					6.54
PC508_4+a					+	7.98
CN907_9+					+	8.00
CN907_8+					+	8.15
CN907_7+					+	8.56
PC508_4+b					+	8.92

LYSOZYME NEGATIVE SAMPLES

Sample	<i>C. butyricum</i>	<i>C. sporogenes</i>	<i>C. perfringens</i>	<i>C. septicum</i>	<i>C. tyrobutyricum</i>	TaqMan Log target genome /50g <i>C. tyr</i>
TN302_4-					+	0.00
TN302_6-					+	0.00
TN305_12-	+				+	0.00
TN307_16-						0.00
TN303_7-						0.00
TN305_6-						0.00
MN460_8-			+		+	2.04
BS628_9-	+	+				2.39
BS654_8-					+	2.52
BS634_1-	+					2.97
BS611_9-	+	+			+	3.11
BS654_7-			+		+	3.18
BS654_6-					+	3.51
TN307_13-					+	4.28
TN305_12-						4.46
TN307_15-	+	+			+	4.49
TN307_16-	+	+				4.80
TN307_13-	+				+	4.81
TN302_5-					+	4.91
TN305_11-a		+			+	4.93
TN305_11-b					+	5.38
PC508_12-	+				+	5.46
TN304_9-					+	6.52
VI616_5-					+	6.71
PD704_6-					+	6.72
TN303_5-					+	6.72
VI616_8-					+	7.11
TN303_10-					+	7.23
TN303_6-					+	7.40
VI616_4-					+	7.40
TN304_22-				+	+	7.51
VI625_10-					+	7.62

VI616_3-	+		+	8.23
CN911_11-		+	+	8.28
VI616_2-			+	9.20

Figure 1

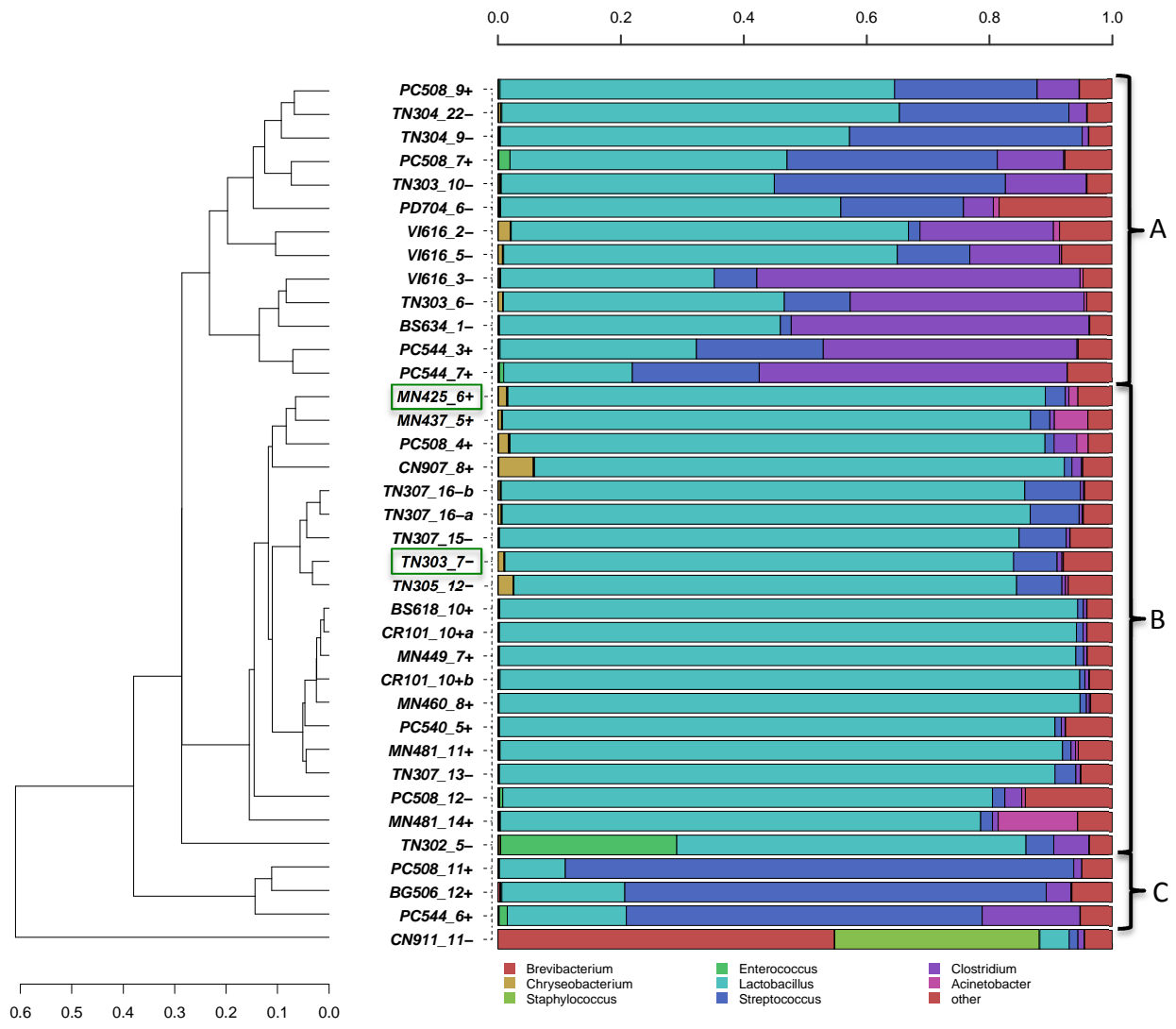


Figure 3

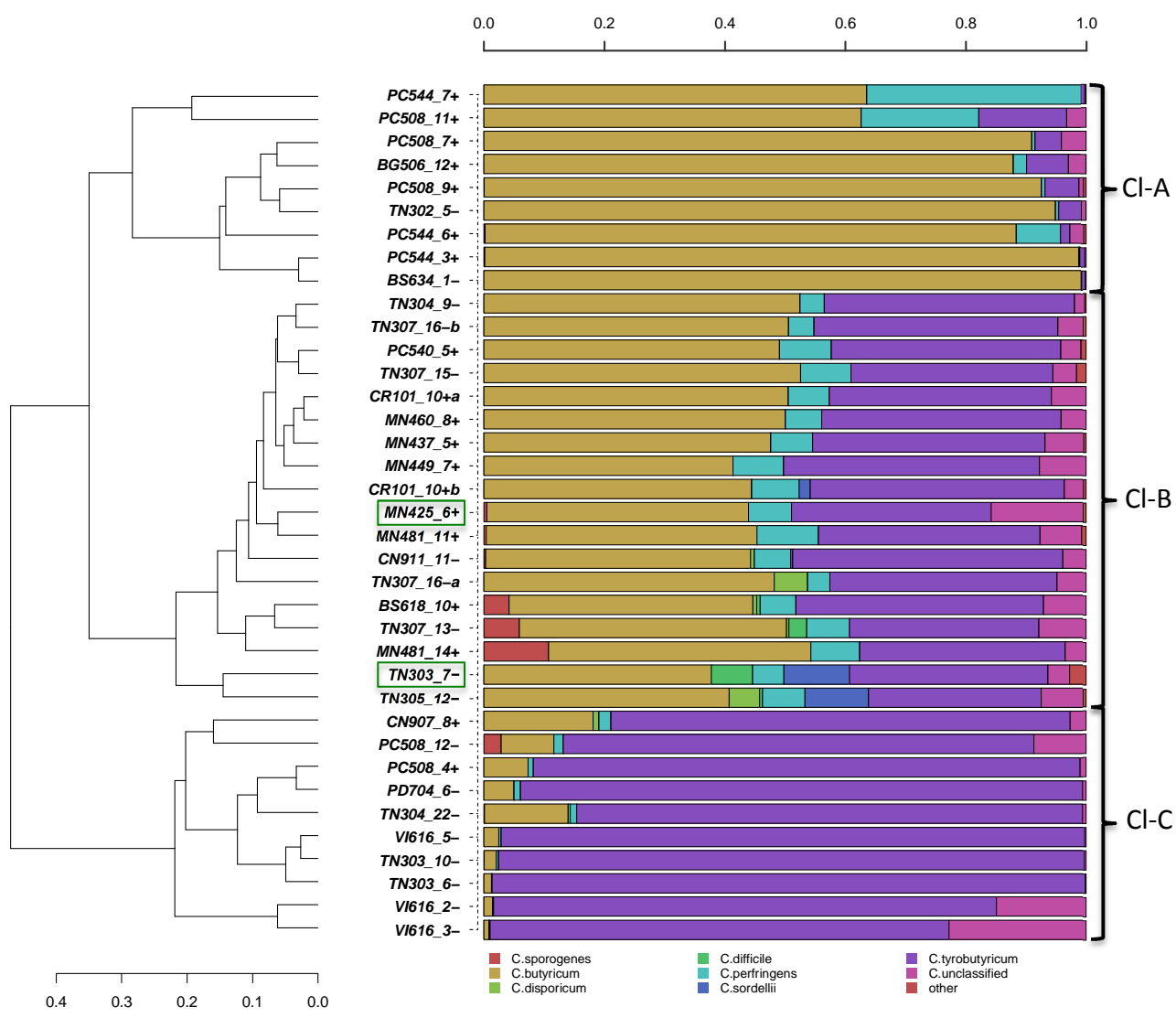
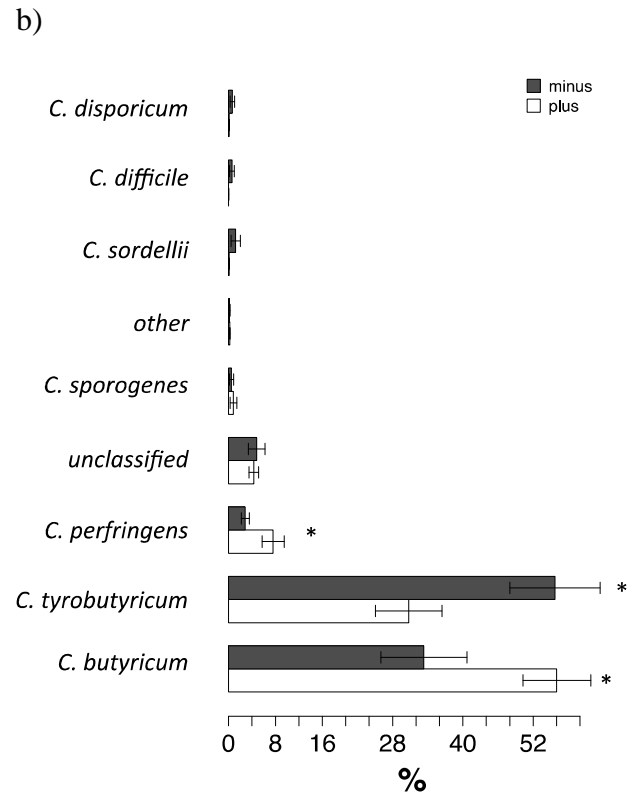
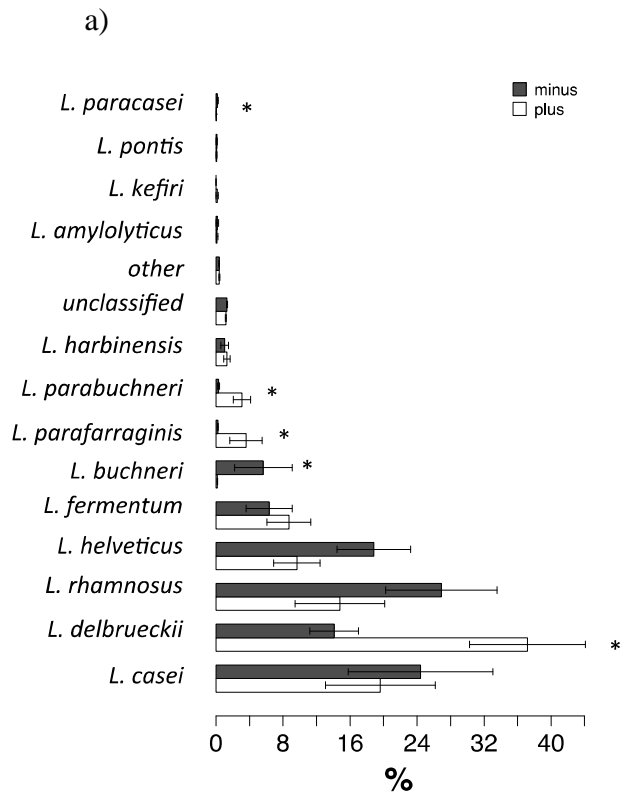


Figure 4



Supplementary Table S1

Sample	Nseqs	coverage	Sobs	Simpson eveness	inverted Simpson	Chao	Shannon eveness	np Shannon
PC_minus_12	32566	0,981	801	0,00405	3,24	3261,46	0,301	2,121
TN_minus_5	32566	0,983	731	0,00610	4,46	2556,60	0,295	2,040
TN_minus_6	32566	0,980	850	0,00399	3,39	3415,43	0,261	1,876
TN_minus_10	32566	0,981	822	0,00487	4,01	2837,37	0,272	1,940
TN_minus_12	32566	0,982	813	0,00218	1,78	2446,64	0,206	1,496
TN_minus_16	32566	0,982	766	0,00242	1,85	2503,11	0,202	1,454
TN_minus_16bis	32566	0,982	778	0,00261	2,03	2892,12	0,211	1,523
TN_minus_15	32566	0,982	781	0,00388	3,03	2920,29	0,258	1,826
TN_minus_13	32566	0,981	847	0,00305	2,59	3006,50	0,233	1,685
PD_minus_6	32566	0,981	837	0,00642	5,37	3049,76	0,363	2,540
VI_minus_5	32566	0,980	859	0,00333	2,86	3053,62	0,272	1,954
VI_minus_3	32566	0,980	873	0,00512	4,47	3103,10	0,312	2,228
VI_minus_2	32566	0,980	837	0,00496	4,15	3346,28	0,329	2,319
CN_minus_11	32566	0,979	907	0,00269	2,44	3007,01	0,218	1,612
TN_minus_22	32566	0,982	785	0,00454	3,56	2659,68	0,254	1,800
TN_minus_9	32566	0,983	739	0,00506	3,74	2770,15	0,272	1,896
BS_minus_1	32566	0,982	770	0,00365	2,81	2653,57	0,228	1,624
TN_minus_7	32566	0,981	827	0,00220	1,82	3279,57	0,221	1,600
MN_plus_7	32566	0,980	820	0,00289	2,37	3056,53	0,228	1,651
MN_plus_8	32566	0,982	751	0,00201	1,51	3214,89	0,158	1,162
CR_plus_10	32566	0,982	782	0,00273	2,13	3162,28	0,200	1,448
CR_plus_10bis	32566	0,982	766	0,00347	2,66	2764,93	0,224	1,597
BS_plus_10	32566	0,979	905	0,00229	2,08	3015,07	0,204	1,523
MN_plus_11	32566	0,980	882	0,00216	1,91	3118,82	0,210	1,552
MN_plus_14	32566	0,981	807	0,00513	4,14	2872,61	0,295	2,085
BG_plus_12	32566	0,980	868	0,00365	3,16	3291,67	0,276	1,987
PC_plus_11	32566	0,980	847	0,00210	1,78	3222,31	0,194	1,435
PC_plus_7	32566	0,981	791	0,00560	4,43	3106,78	0,302	2,124
CN_plus_8	32566	0,982	751	0,00239	1,79	2931,65	0,203	1,457
PC_plus_4	32566	0,983	720	0,00197	1,42	2361,51	0,158	1,152
PC_plus_9	32566	0,983	734	0,00406	2,98	2393,16	0,260	1,822
PC_plus_5	32566	0,983	755	0,00309	2,33	2396,38	0,220	1,565
PC_plus_7	32566	0,982	772	0,00728	5,62	2653,03	0,332	2,308
PC_plus_6	32566	0,981	798	0,00464	3,70	2909,51	0,295	2,078
PC_plus_3	32566	0,981	834	0,00478	3,99	2577,45	0,293	2,082
MN_plus_5	32566	0,982	785	0,00468	3,68	2691,56	0,279	1,968
MN_plus_6	32566	0,982	790	0,00252	1,99	2851,50	0,232	1,657