



UNIVERSITÀ
DEGLI STUDI
DI UDINE

Università degli studi di Udine

Metagenomic profiles of different types of Italian high-moisture Mozzarella cheese

Original

Availability:

This version is available <http://hdl.handle.net/11390/1146739> since 2020-03-05T13:37:10Z

Publisher:

Published

DOI:10.1016/j.fm.2018.12.007

Terms of use:

The institutional repository of the University of Udine (<http://air.uniud.it>) is provided by ARIC services. The aim is to enable open access to all the world.

Publisher copyright

(Article begins on next page)

Manuscript Details

Manuscript number	FM_2018_722_R1
Title	Metagenomic profiles of different types of Italian high-moisture Mozzarella cheese
Article type	Research Paper

Abstract

The microbiota of different types of Italian high-moisture Mozzarella cheese produced using cow or buffalo milk, acidified with natural or selected cultures, and sampled at the dairy or at the mass market, was evaluated using a Next Generation Sequencing approach, in order to identify possible drivers of the bacterial diversity. Cow Mozzarella and buffalo Mozzarella acidified with commercial cultures were dominated by *Streptococcus thermophilus*, while buffalo samples acidified with natural whey cultures showed similar prevalence of *L. delbrueckii* subsp. *bulgaricus*, *L. helveticus* and *S. thermophilus*. Moreover, several species of non-starter lactic acid bacteria were frequently detected. The diversity in cow Mozzarella microbiota was much higher than that of water buffalo samples. Cluster analysis clearly separated cow's cheeses from buffalo's ones, the former having a higher prevalence of psychrophilic taxa, and the latter of *Lactobacillus* and *Streptococcus*. A higher prevalence of psychrophilic species and potential spoilers was observed in samples collected at the mass retail, suggesting that longer exposures to cooling temperatures and longer production-to-consumption times could significantly affect microbiota diversity. Our results could help in detecting some kind of thermal abuse during the production or storage of mozzarella cheese.

Keywords	Mozzarella cheese; Microbiota; Next Generation Sequencing; Psychrotrophs; Metagenomics
Corresponding Author	Marilena Marino
Corresponding Author's Institution	University of Udine
Order of Authors	Marilena Marino, Giorgia Dubsy de Wittenau, Elena Saccà, Federica Cattonaro, Alessandro Spadotto, nadia innocente, Slobodanka Radovic, Edi Piasentier, Fabio Marroni

Submission Files Included in this PDF

File Name [File Type]

Cover letter Rev1.docx [Cover Letter]
Answer_to_Reviewers.docx [Response to Reviewers]
Highlights.docx [Highlights]
Paper_Mozzarella_NGS_Rev1.docx [Manuscript File]
Figure 1.docx [Figure]
Figure 2.docx [Figure]
Figure 3.docx [Figure]
Figure 4.docx [Figure]
Figure S1.docx [Figure]
Table 1.docx [Table]
Table 2.docx [Table]
Table 3.docx [Table]

Submission Files Not Included in this PDF

File Name [File Type]

Table S1.xlsx [Table]

Table S2.xlsx [Table]

To view all the submission files, including those not included in the PDF, click on the manuscript title on your EVISE Homepage, then click 'Download zip file'.

Editor-in-Chief
Food Microbiology

23rd October 2018

Dear Editor,

We hereby submit a revised form of the manuscript entitled “*Metagenomic profiles of different types of Italian high-moisture Mozzarella cheese*” by Marino et al. to be considered for publication as an original research paper in *Food Microbiology*.

The paper has been corrected according the referees’ suggestions. We highlighted by colour each change made in the text as raised in the reviewer comments, and provided a separate suitable rebuttal to each reviewer comment.

We hope you find our manuscript suitable for publication and look forward to hearing from you.

Sincerely,

Marilena Marino

Dipartimento di Scienze Agroalimentari, Ambientali e Animali, University of Udine, Italy

6 **Reviewer 1**

7
8 The paper FM_2018_772 “Metagenomic profiles of different types of Italian high-moisture Mozzarella
9 cheese“ – evaluated the microbiota, by Illumina MiSeq approach, of different types of Italian high-moisture
10 Mozzarella cheese produced using cow or buffalo milk, acidified with natural or selected cultures, and
11 sampled at the dairy or at the mass market.

12 The study is interesting and showed in one investigation the microbiota of several types of mozzarella cheese
13 produced with different procedures.

14 I have some advices in order to improve the manuscript:

15
16 So, I think that a very important aspect that the authors should avoid is to say that by such analysis is
17 possible to differentiate the PDO mozzarella cheeses by other kind of mozzarellas. In order to validate such
18 affirmation a largest panel of samples is necessary, which includes mozzarellas of other regions and
19 produced also in different seasons and so and so....

20 The important aspect evidenced by this investigation is the possibility to detect some irregularity during the
21 production of mozzarella di bufala. I think that this message should be stressed. So, not the origin, but the
22 safety and quality of the product are the main aim of the study. I think that the safety of the consumers is
23 more important than such a marketing issue referred to the labels of PDO and brothers....

24 **Thank you for evidencing this important point. We now shifted the emphasis towards safety and quality by**
25 **changing the last paragraph of the abstract and part of the introduction.**

26
27 **As a result of our changes, in the text we do not mention anymore our intention to discriminate PDO from**
28 **non-PDO (a task for which as the reviewer correctly pointed out a more detailed study would be needed),**
29 **and we only focus on the importance of characterizing the microbiota of mozzarella cheese, especially in**
30 **virtue of safeguarding consumer's health. The discrimination PDO/non-PDO is still present in the paper just**
31 **as one of the many variables between mozzarella cheese samples.**

32
33
34 **Minor issues:**

35 - Please I think that it is necessary just one sentence, perhaps in the Materials and Methods section, to
36 explain clearly the various acronyms (BNCG, BDN, CC etc) of the mozzarellas used in this study. Yes, they
37 are explained, but around the text.

38 **We added a numbered list in paragraph 2.1, explaining the acronyms.**

39
40 - Caption of the Figure 3, please write down “natural whey culture“ instead of NWC.

41 **Done.**

42
43 - Please put the legend in the Figures 1, 2 3 and Table 2 where you explain the meaning of the acronyms
44 (BNCG, BDN, CC etc). In this way the reader do not need to go back to manuscript text to find the meaning
45 of these codes.

46 **Done.**
47
48
49
50
51
52
53
54
55
56
57
58
59

60
61
62 **Reviewer 2**
63

64 Manuscript FM_2018_722 Metagenomic profiles of different types of Italian high-moisture
65 Mozzarella cheese.
66

67 Marilena Marino et colleagues analyzed the composition of the microbiota of different types of high-
68 moisture Mozzarella cheese by using an NGS approach.
69

70 The study's objective is well justified and presents significant interest not just for microbiologists but
71 for the entire community of investigators interested in food sciences.

72 They used, in order to identify possible drivers of the bacterial diversity, samples with different
73 characteristics in term of type of milk (water-buffalo and cow), acidification (by natural whey culture
74 or selected starter), certification status (PDO or non-PDO), and sampling point (local or mass retailer).
75

76 Just a few minor points that need to be addressed by the authors.

77 L87. Please, used biota instead of flora.
78

79 **Done**
80
81

82 L101. Were the samples collected from the same day/period of cheesemaking?
83

84 If so it might have been statistically more valid to sample over 3 different days and not in the same
85 day to see if contamination from the environment play an important role.

86 **Sampling was performed at the point of sale, with no strict control on the time passed from the**
87 **beginning of cheesemaking; this has the disadvantage that cheese sampled at local stores might tend**
88 **to have shorter times than cheese samples at mass retailers, but it has the advantage to produce a**
89 **realistic picture of what the consumers have the opportunity to buy.**
90

91 **To clarify, we specify at the beginning of paragraph 2.1 that we sampled the cheese at local shops or**
92 **in supermarkets, i.e. when the cheese is on sale.**
93
94

95 Were the samples collected and analyzed in triplicate?
96

97 **While we understand that having triplicates would be desirable, we conducted this pilot study on**
98 **individual samples. For this reason, differences were only tested between groups and never between**
99 **individual samples (for which no valid analysis can be conducted, due to the lack of replicates).**
100

101 Staphylococci and micrococci were different between the cheeses. Were the data analyzed?
102

103 **We think that the reviewer refers to what was shown in Figure 1. Actually, both Staphylococcaceae**
104 **and Micrococcaceae are present at very low prevalence in all samples, but the colors were similar to**
105 **those of Streptococcaceae and Moraxellaceae, respectively, which are present (although with**
106 **different abundances) in all samples. We modified the graph so that we hope the reader can better**
107 **understand differences between samples.**
108
109

110 L233. It's statistically significant those difference in the microbial diversity?
111

112 **Thank you for the point. We didn't test the difference. We now performed a very naïve t-test**
113 **comparing the Chao1 diversity in Cow and in Buffalo samples and we found that the difference is**
114 **statistically significant. We added the results of the test to the discussion and to the newly added**
115 **Table S2 (see below).**
116
117
118

119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177

L326. It is not clear “Metagenomics approach can confidently discriminate cow Mozzarella from buffalo Mozzarella”. Please, add more details.

We meant that the use of metagenomics can leverage differential abundance of bacterial species in mozzarella cheese to discriminate cow and buffalo mozzarella. We changed the sentence to “Metagenomics approach can leverage differential abundance of bacterial species to confidently discriminate cow Mozzarella from buffalo Mozzarella”.

Can you add more details about statistical analysis and related results to measure the sequencing diversity, included Chao1 richness, Shannon diversity, and Good's coverage results, as well as monitoring results for sequencing abundance (rarefaction)?

We added Table S2 reporting Chao1 richness (again), Shannon’s diversity, Good’s coverage, and Chao1 on data rarefied at 20000 reads, together with the results of the t-test to assess the significance of the diversity between Cow and Buffalo mozzarella cheese. Results of the test are discussed in the main Manuscript in the 2nd paragraph of the discussion.

178
179
180 **Reviewer 3**
181

182 The manuscript entitled “Metagenomic profiles of different types of Italian high-moisture
183 Mozzarella cheese” provides very interesting findings related to the microbiota found in types of
184 Mozzarella cheese differently manufactured. It includes a comprehensible discussion about the
185 undesirable effects that could be produced in cow Mozzarella with longer refrigeration times, and
186 longer production-to-consumption times caused to the presence of psychrotrophic bacteria.
187 Moreover, results show the importance of keeping the traditional procedures used by the PDO
188 cheesemakers. I recommend strongly the publication of this manuscript. However, I have the
189 following comments to improve it.
190
191

- 192
193 1. Abstract, page 2, line 25, *Corynebacterium* belongs to the psychrotrophic genera. Results
194 show that buffalo Mozzarella was enriched with *Lactococcus*, *Streptococcus*, and *Weissella*,
195 instead. Thank you, we corrected the abstract accordingly. We did not mention *Weissella* in
196 the abstract because it is a rare genus and we decided not to report extensive results for rare
197 genera.
198
- 199 2. The second highlight exceeds the maximum number of characters. Try dividing it into two
200 sentences. Thank you, we followed your suggestion.
201
- 202 3. Page 4, line 87, page 9, line 232 and page 11, lines 280 and 287, use microbiota instead of
203 microflora or flora. Thank you, we followed your suggestion.
204
- 205 4. Page 4, lines 90-93, a reference is missing for that paragraph. We added the missing
206 reference.
207
- 208 5. Page 5, line 122, please give a more explicit description of the second amplification step.
209 Which flow-cell binding domains and unique indices? We added the missing information.
210
- 211 6. Page 5, line 124, what does SRA stands for? SRA stands for Sequence Reads Archive
212 (<https://www.ncbi.nlm.nih.gov/sra>). We added the full name in the methods section.
213
- 214 7. Table 1, consider changing the following column descriptions: Reads/sample, Identified
215 OTUs/sample and Estimated OTUs/sample[§] Then in the table notes [§] (Chao, 1984). We
216 followed reviewer’s suggestion.
217
- 218 8. I suggest that Figure S1 should not be supplementary, but part of the main manuscript. We
219 followed reviewer’s suggestion. Figure S1 is now Figure 1. Numbers of all other figures
220 have been shifted accordingly.
221
- 222 9. Page 7, line 156, use approximation instead of proxy. We followed reviewer’s suggestion.
223
- 224 10. Page 7, line 171, in table S1 there are 74 OTUs, instead of 75. We incorrectly counted the
225 header as an OTU. We now report the correct number of OTUs (74).
226
- 227 11. In the caption of Figure 3, for intelligibility, use natural whey culture instead of NWC.
228 Done.
229
- 230 12. Table 2, please state that BM is buffalo Mozzarella and CM is cow Mozzarella. In the table
231 notes, explain what FDR means. We followed reviewer’s suggestion.
232
- 233 13. Page 10, line 265, the paper by Martino et al. (2013) refers to a bacteriocin produced
234 by *Pediococcus pentosaceus*. In the analyzed Mozzarella cheeses, *Pediococcus* was not
235 identified. Instead, there are plenty of references related to bacteriocins produced by the
236 NSLAB found in this study. Done
14. The paper from Delorme et al. 2015 is not mentioned in the text. The reference was deleted
from the references’ list

Highlights

- Metagenomics clearly allows to distinguish cow Mozzarella from buffalo Mozzarella
- Cow Mozzarella show a higher bacterial diversity
- Cow mozzarella show a large presence of psychrophilic species
- Sampling point (local or mass retail) is a possible driver of bacteria diversity

1
2
3 **1 Metagenomic profiles of different types of Italian high-moisture**
4
5 **2 Mozzarella cheese**
6
7

8 4 Marilena Marino^{a,*}, Giorgia Dubsky de Wittenau^b, Elena Saccà^a, Federica Cattonaro^b,
9 5 Alessandro Spadotto^b, Nadia Innocente^a, Slobodanka Radovic^b, Edi Piasentier^a, Fabio Marroni^{b,*}
10 6

11
12
13 ^a Dipartimento di Scienze Agroalimentari, Ambientali e Animali, Università di Udine, via
14 8 Sondrio 2/A, 33100 Udine, Italy

15
16 ^b IGA Technology Services s.r.l., Via J. Linussio 51, 33100, Udine, Italy
17

18 10

19 11 Corresponding author:

20
21 12 Marilena Marino: Tel. +39 0432 558150; E-mail address: marilena.marino@uniud.it
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56

57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112

14 **Abstract**

15 The microbiota of different types of Italian high-moisture Mozzarella cheese produced using cow
16 or buffalo milk, acidified with natural or selected cultures, and sampled at the dairy or at the
17 mass market, was evaluated using a Next Generation Sequencing approach, in order to identify
18 possible drivers of the bacterial diversity. Cow Mozzarella and buffalo Mozzarella acidified with
19 commercial cultures were dominated by *Streptococcus thermophilus*, while buffalo samples
20 acidified with natural whey cultures showed similar prevalence of *L. delbrueckii* subsp.
21 *bulgaricus*, *L. helveticus* and *S. thermophilus*. Moreover, several species of non-starter lactic acid
22 bacteria were frequently detected. The diversity in cow Mozzarella microbiota was much higher
23 than that of water buffalo samples. Cluster analysis clearly separated cow's cheeses from
24 buffalo's ones, the former having a higher prevalence of psychrophilic taxa, and the latter of
25 *Lactobacillus* and *Streptococcus*. A higher prevalence of psychrophilic species and potential
26 spoilers was observed in samples collected at the mass retail, suggesting that longer exposures to
27 cooling temperatures and longer production-to-consumption times could significantly affect
28 microbiota diversity. Our results could help in detecting some kind of thermal abuse during the
29 production or storage of mozzarella cheese.

31 **Keywords**

- 32 High-moisture Mozzarella cheese
- 33 Microbiota
- 34 Next Generation Sequencing
- 35 Psychrotrophs
- 36 Metagenomics

37 1 Introduction

38 High-moisture Mozzarella is one of the most popular unripened cheeses on the market. It
39 belongs to the cheese category “Pasta Filata”, which refers to a unique processing step of curd
40 plasticization and stretching, during which the acidified curd is soaked in hot water or salt brine
41 until a plastic consistency is achieved. The hot plastic curd is then kneaded and stretched to
42 produce a homogeneous cheese with a fiber-like structure. Right after production Mozzarella
43 cheese is packaged in liquid and stored under refrigerated conditions for up to 5 days (Gorrasi et
44 al., 2016). Many varieties of high-moisture Mozzarella cheese exist on the market, usually
45 produced using cow’s or buffalo’s milk. Regarding buffalo Mozzarella cheese, the Protected
46 Designation of Origin (PDO) has been assigned to Mozzarella di Bufala Campana by the
47 European Commission in 1996. The PDO territory, in which raw buffalo milk has to be produced
48 and processed, currently includes some areas in the Italian regions of Campania and Lazio. The
49 highly valued PDO Mozzarella di Bufala Campana cheese is traditionally made from Italian
50 Mediterranean buffalo (*Bubalus bubalis*, river type) milk acidified by adding a natural whey
51 culture (NWC) starter obtained from the batch of the previous day with the technique called
52 backslopping. The specific and highly appreciated features of the final product originate mainly
53 from the quality of raw materials used during processing, the agri-ecosystem of the production
54 area, and the traditional processing technology (Ercolini et al., 2012). Non-PDO buffalo
55 Mozzarella cheeses can also be produced, e.g. using or transforming milk coming from regions
56 outside of the borders of the PDO geographical area, or acidifying curd with selected commercial
57 starter cultures (CS). The cheaper and more widespread cow’s milk Mozzarella cheese is instead
58 produced using raw or pasteurized cow’s milk that is acidified using a variety of methods,
59 including citric acid addition and/or biological acidification carried out mainly by selected
60 commercial starters. Both NWC and CS have the main function to ensure a rapid acidification of
61 the curd, by synthesizing enough lactic acid to demineralize and transform the curd into the state
62 that undergoes stretching in hot water at the target pH (de Candia et al., 2007).

63 During the last decades, several methodologies have been applied to characterize Mozzarella
64 cheese **with the aim to ensure high quality and safety standards**. Polymerase chain reaction
65 (PCR) has been employed to detect species-specific DNA sequences in milk and cheese
66 (Lopparelli et al., 2007), and isoelectric focusing, reversed-phase liquid chromatography, mass
67 spectrometry and enzymatic assays to check the presence of specific buffalo and cow proteins in
68 milk and cheese (Addeo et al., 2009; Hurley et al., 2006). Recently, a metabolomic approach

169
170 69 based on gas-chromatography mass-spectrometry coupled with the analysis of the composition
171 70 of predominant cultivable microbiota has been used to discriminate different types of Mozzarella
172 71 cheese and to protect the authenticity of PDO Mozzarella di Bufala Campana cheese (Pisano et
173 72 al., 2016). Due to the high water content and relatively high pH, microbial spoilage of
174 73 Mozzarella cheese might occur, caused by proteolytic and/or lipolytic microorganisms that can
175 74 cause unwanted modifications of the texture, off-odors or discolorations (Andreani et al., 2014;
176 75 Segat et al., 2014). In the last decade, the food microbiology has been deeply revolutionized by
177 76 the use of Next Generation Sequencing (NGS) technologies, which can provide a thorough
178 77 analysis of microbial diversity present in a food sample, producing much deeper output than
179 78 more commonly used culture-independent approaches (Chen et al., 2017; Marino et al., 2017).
180 79 Currently, only two studies have been carried out to study the microbial diversity of Mozzarella
181 80 cheese using an NGS approach (Ercolini et al., 2012; Guidone et al., 2016). However, the
182 81 **microbiota** of the buffalo and the cow Mozzarella cheese has been studied in separate papers,
183 82 which makes it difficult to understand the potential of NGS-based metagenomics in
184 83 distinguishing products obtained with milk of different animal origins and different technologies.
185 84 Moreover, the only study carried out on cow Mozzarella cheese analyzed the cheese microbiota
186 85 after a 5-d refrigerated storage, which could have favored the growth of psychrotrophic
187 86 microorganisms and hence modified to some extent the composition of the native microbiota of
188 87 Mozzarella cheese (Guidone et al., 2016).
189 88 The objective of this study was to analyze the composition of the microbiota of different types of
190 89 high-moisture Mozzarella cheese by using an NGS approach. In order to identify possible drivers
191 90 of the bacterial diversity, samples with different characteristics in term of type of milk (water-
192 91 buffalo and cow), acidification (by natural whey culture or selected starter), certification status
193 92 (PDO or non-PDO), and sampling point (local or mass retailer) were included in the study.
194 93

194 94 **2 Materials and Methods**

195 95 *2.1 Samples collection*

196 96 Thirty-nine samples of high-moisture Mozzarella cheese were collected **in local or mass retailers**
197 97 to maximize the variability of factors potentially affecting the cheese microbiota composition,
198 98 namely type of milk, acidification system, certification status, and sampling point (Table 1).
199 99 Three main groups of buffalo Mozzarella and cow Mozzarella samples were collected as
200 100 follows: (i) 15 PDO Mozzarella cheese produced with buffalo milk and acidified with NWC, and

225
226 101 purchased at local dairies in the main districts of the production area, (ii) 11 PDO Mozzarella
227
228 102 cheese produced in the PDO area with buffalo milk and acidified with NWC, but collected in
229
230 103 supermarkets, and (iii) 13 non-PDO Mozzarella cheese collected in supermarkets, including
231
232 104 buffalo Mozzarella acidified with CS, buffalo Mozzarella acidified with NWC, and cow's milk
233 105 Mozzarella acidified with CS.

234 106 For the aim of the present work, samples were classified as follows:

- 236 107 1) BDN: Buffalo mozzarella with PDO certification and acidified with Natural Whey
237 Culture (15 samples)
- 239 109 2) BDNG: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture
240 and collected at mass retailers (11 samples)
- 242 111 3) BNNG: Buffalo mozzarella without certification, acidified with Natural Whey Culture
243 and collected at mass retailers (3 samples)
- 245 113 4) BNCG: Buffalo mozzarella without certification, acidified with commercial starters and
246 collected at mass retailers (2 samples)
- 248 115 5) CC: Cow mozzarella acidified with commercial starters and collected at mass retailers (8
250 samples)

253 118 2.2 DNA extraction and sequencing

255 119 Immediately after collection, all samples were frozen (- 20 °C). First, 50 mg were split off to be
256
257 120 incubated for 90 min at 65 °C with 600 µL of CTAB Buffer, 30 µL of Proteinase K and 2 µL of
258
259 121 RNase Solution (Promega, WI) and then were centrifuged to collect 300 µL of the lysate to be
260
262 122 used as input for the total DNA extraction. The Maxwell® 16 Instrument (Promega, WI) with
261 Maxwell® 16 FFS Kit (Promega, WI) were used for all samples.

263 124 The bacterial diversity was obtained by the library preparation and sequencing of the 16S rRNA
264
265 125 gene. The following two amplification steps were performed: an initial PCR amplification using
266
267 126 16S locus specific PCR primers (16S-341F 5'-CCTACGGGNGGCWGCAG-3' and 16S-805R
268
269 127 5'-GACTACHVGGGTATCTAATCC-3') and a subsequent amplification integrating relevant
270
271 128 flow-cell binding domains (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' for
272
273 130 the For primer and 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' for the
274
275 131 reverse overhang) and unique indices selected among those available Nextera XT Index Kits
276
277
278
279
280 combined according to manufacturer's instructions (Illumina, CA). Libraries were sequenced in

281
282 132 a MiSeq (Illumina, CA) in paired end with 300-bp read length. Raw reads are available on
283
284 133 [Sequence Reads Archive under the accession](#) SRP156292.

285 134
286
287 135 *2.3 Data analysis*

288 136 Reads were de-multiplexed based on Illumina indexing system. Sequences were analyzed using
289 137 QIIME 1.5.0 (Caporaso et al., 2010). After filtering based on read quality and length (minimum
290 138 quality = 25 and minimum length = 200), Operational Taxonomic Units (OTUs) defined by a
291 139 97% of similarity were picked using the Uclust v1.2.22q method (Edgar, 2013) and the
292 140 representative sequences were submitted to the RDP classifier (Wang et al., 2007) to obtain the
293 141 taxonomy assignment and the relative abundance of each OTU using the Greengenes 16S rRNA
294 142 gene database (McDonald et al., 2012). [Alpha-diversity analysis was performed using QIIME](#)
295 143 [1.5.0](#) (Caporaso et al., 2010) [and R](#) (R Core Team, 2018); [the following alpha-diversity indexes](#)
296 144 [were computed: Chao1](#) (Chao, 1984), [Good's coverage](#) (Good, 1953), [and Shannon's diversity](#)
297 145 [index](#) (Shannon, 1948). Selection of the OTUs for downstream analysis was performed by
298 146 requiring that the OTUs represented at least 0.1% of at least one study sample. Clustering was
299 147 performed using the R function heatmap.2 on the read counts normalized using DESeq on the 50
300 148 most represented OTUs (Anders and Huber, 2010). Differential abundance of OTUs across
301 149 categories of samples was tested using the differential_abundance.py routine implemented in
302 150 QIIME (Caporaso et al., 2010). The routine returns results of Fisher's exact test and the fit of a
303 151 zero inflated Gaussian model (fitZIG) (Paulson et al., 2013). An OTU was considered to be
304 152 differentially present in two samples if the adjusted p-value (FDR) was lower than 0.05.
305 153 Partial Least Squares – Discriminant Analysis (PLS-DA) was applied by Unscramble X 10.4
306 154 (CAMO software AS, Oslo, Norway) to check the efficacy of the relative abundance of the most
307 155 represented OTUs in discriminating the Mozzarella samples according to the method of
308 156 acidification and the market target (Chevallier et al., 2006). With this aim, the PLS-DA model
309 157 was built between the OTUs matrix and the cheese matrix, which was created by defining three
310 158 dummy variables, one for each Mozzarella type considered: acidified by commercial starters
311 159 (BNCG and CC groups) and acidified by natural whey culture, locally (BDN group) or large
312 160 scale distributed (BDNG and BNNG groups).

313 161
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336

3 Results

Summary statistics of the sequencing results for all samples are reported in Table 1. Briefly, a total of 4,511,861 paired end reads were sequenced, with an average of 115,689 reads per sample (range 45,170-216,852). The number of identified OTUs per sample ranged 463 to 1,569 with an estimated number of OTUs (Chao, 1984) ranging 795 to 2,623. The estimated number of OTUs is an approximation of within sample diversity (Chao, 1984) and was significantly higher in Mozzarella samples produced with cow's milk than in all other samples ($p < 0.05$, pairwise Wilcoxon-Mann-Whitney test, Figure 1).

Identification at the species or genus level was obtained for 47% and 48% of OTUs, respectively, and only 4% of OTUs were identified only at the family level. Twenty-six families were present at abundance $> 0.1\%$ in at least one sample, with *Lactobacillaceae* and *Streptococcaceae* being the most prevalent in all samples. Figure 2 shows the distribution of the most abundant ($> 0.1\%$) families in all the samples. Cow Mozzarella samples (CC samples) and buffalo Mozzarella acidified with CS (BNCG samples) were dominated by *Streptococcaceae*, which ranged 47-85% and 86-90% in CC and BNCG samples, respectively. *Lactobacillaceae* were instead detected at lower prevalence, ranging 0-11%. Conversely, samples acidified with NWC (i.e. BDN, BDNG and BNNG) showed usually a higher prevalence of *Lactobacillaceae* (18-80% of identified OTUs), and *Streptococcaceae* were also abundant (13-71%). Some non-lactic families, namely *Enterobacteriaceae*, *Flavobacteriaceae*, *Moraxellaceae*, and *Pseudomonadaceae*, were present in all samples.

74 OTUs were present with an abundance of at least 0.1% in at least one sample (Table S1). The most represented OTUs belonged to the species of *Streptococcus thermophilus*, *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Lactobacillus helveticus*. CC and BNCG samples were dominated by *Streptococcus thermophilus*, and the two thermophilic lactobacilli were relatively rare. The second most abundant OTUs in CC samples belonged to the genus *Acinetobacter*, followed by *Pseudomonas*. In BDN, BNNG and BNCG samples, i.e. water buffalo Mozzarella acidified with NWC, the prevalence of *S. thermophilus*, *L. delbrueckii* subsp. *bulgaricus* and *L. helveticus* was quite similar and taken together these three species represented the vast majority of identified OTUs (65.90-98.49%). Many other lactic acid bacteria (LAB) belonging to the genera *Lactococcus*, *Lactobacillus*, *Leuconostoc*, *Streptococcus* and *Weissella* were generally detected at high frequencies (0.78-14.40%). Also, in these groups of samples, the presence of *Acinetobacter* was detected at relatively high levels (about 2.40% of identified OTUs). In

393
394 194 addition to *Acinetobacter*, a variety of other psychrotrophic genera (including *Corynebacterium*,
395 195 *Flavobacterium*, *Chryseobacterium*, *Pseudomonas*, *Shewanella*, *Escherichia*, and *Enterobacter*)
396 196 were found in all samples, although with different abundances within the groups.
397
398 197 Figure 3 shows the heatmap of the Mozzarella samples clustered by Euclidean distance
399 198 computed based on the 50 more abundant OTUs. The cluster clearly separated cow Mozzarella
400 199 from buffalo Mozzarella. In addition, the samples of buffalo Mozzarella obtained using
401 200 commercial starters (BNCG samples) were in an intermediate position between cow and buffalo
402 201 milk Mozzarella samples.
403
404 202 Considering the strong separation between cow Mozzarella and buffalo Mozzarella observed in
405 203 the cluster, an enrichment test contrasting the samples belonging to the two categories was
406 204 carried out to identify the species responsible for the differentiation. Table 2 shows the list of
407 205 differentially abundant species between buffalo and cow Mozzarella samples. Buffalo
408 206 Mozzarella samples showed a higher prevalence of *Lactobacillus* species, in addition to the
409 207 species *Streptococcus equinus*, while no significant difference in abundance of *Streptococcus*
410 208 *thermophilus* was observed. Conversely, in cow Mozzarella samples a higher prevalence of
411 209 several psychrophilic taxa, including *Brochothrix*, *Erwinia*, *Flavobacterium*, *Pseudomonas*, and
412 210 *Shewanella*, as well as thermophilic and spore-forming genera, as *Anoxybacillus flavithermus* and
413 211 *Thermus thermophilus*, was observed.
414
415 212 To further characterize differences in the microbiota of Mozzarella acidified with NWC (BDN,
416 213 BDNG and BNNG samples), a cluster analysis was carried out after removing samples obtained
417 214 with CS. The analysis produced two main clusters (Figure 4). Cluster 1 included most samples
418 215 purchased from local market (12 out of 15 samples), and Cluster 2 comprises most samples (10
419 216 out of 14 samples) collected in supermarkets. A similar conclusion can be drawn from the score
420 217 plot of PLS-DA (Figure S1), where along Factor 1, CC and BNCG are clearly discriminated
421 218 from buffalo Mozzarella obtained by acidification with NWC. Moreover, along Factor 2, NWC
422 219 samples show the tendency to split into two populations according to the market target.
423 220 To further investigate the differences between samples belonging to Cluster 1 and Cluster 2, we
424 221 performed a test for differential enrichment in OTUs between the clusters. Results are listed in
425 222 Table 3. Several psychrotrophic genera were overrepresented in Cluster 2, including
426 223 *Acinetobacter*, *Chryseobacterium*, *Citrobacter*, *Corynebacterium*, and *Pseudomonas*.
427 224 Conversely, in Cluster 1 *Lactococcus* spp., *Streptococcus vestibularis* and *Weissella viridescens*
428 225 were present with significantly higher prevalence than in Cluster2.

226 4 Discussion

227 We collected thirty-nine samples of cow and buffalo Mozzarella cheese from local and mass
228 market and submitted to culture-independent NGS, in order to get an in-depth quantitative
229 picture of the structure of the bacterial populations and to identify possible drivers of the
230 bacterial diversity. We identified a much higher number of OTUs compared to previous studies
231 on buffalo or cow Mozzarella cheese, probably because of the higher number of reads obtained.
232 The estimated number of OTUs in buffalo Mozzarella was similar to previous estimates (Ercolini
233 et al., 2012), whereas for cow Mozzarella was higher than previously reported (Guidone et al.,
234 2016).

235 The Chao1 diversity index in cow Mozzarella microbiota was higher than that of buffalo
236 Mozzarella (Table S2) (two tailed t-test, $p < 10^{-4}$), the same was true for the Good's coverage
237 ($p = 0.0255$) and for the Shannon index, although the latter difference was not statistically
238 significant. These observations are in contrast with the data presented in the only recent report on
239 the microbiological profile of Mozzarella cheese produced with buffalo and cow milk, in which
240 the authors identified a larger number of species in buffalo mozzarella (Pisano et al., 2016).

241 However, the authors isolated a very small number of strains and explored the Mozzarella
242 diversity using only culture-based techniques, which are known to have low sensitivity and may
243 lead to an underestimation of microbial diversity present in food environments. Cow Mozzarella
244 samples were all acidified with commercial starters, which is known to reduce the diversity of
245 microbiota in cheese (Coppola et al., 2001). However, in this study, a different observation was
246 made. In fact, with the exception of the lactic starter microbiota, the samples of cow's milk
247 Mozzarella were characterized by a higher microbial diversity than buffalo Mozzarella samples.
248 This might be attributed to a different microbial composition of bovine milk compared to that of
249 the buffalo. In fact, although at present there are no data comparing the composition of the milk
250 microbiota of the two species obtained using NGS techniques, an overview derived from a
251 number of separate studies allowed evidencing a greater number of bacterial genera present in
252 cow's milk (Quigley et al., 2013).

253 The starter composition has a major effect on the microbiota of the final Mozzarella. In water
254 buffalo samples acidified with NWC the species *L. delbrueckii* subsp. *bulgaricus*, *L. helveticus*
255 and *S. thermophilus* were present with a similar prevalence. The presence of these species
256 reflects the microbial composition of NWC used for acidification processes, where these LAB
257 assure lactose fermentation, curd ripening and formation of a typical aroma profile. Mozzarella

505
506 258 samples produced with cow milk were instead dominated by *S. thermophilus*, confirming
507
508 259 previous findings (Guidone et al., 2016; Pisano et al., 2016), whereas the two thermophilic
509
510 260 lactobacilli were less represented within the microbiota. This could be related to the use of
511
512 261 commercial starters, which is quite common in cow Mozzarella, and usually consist of *S.*
513
514 262 *thermophilus* alone or associated with *L. delbrueckii* in smaller concentrations, in order to avoid
515
516 263 the risk of excessive secondary proteolysis that might take place in a high moisture environment
517
518 264 (Pisano et al., 2016). Anyway, *L. helveticus* and *L. delbrueckii* subsp. *bulgaricus* were present as
519
520 265 sub-dominant LAB. Galactose-fermenting *L. helveticus* could help to reduce the accumulation of
521
522 266 galactose, which is not fermented by *S. thermophilus* and *L. delbrueckii* subsp. *bulgaricus*, thus
523
524 267 reducing the risk of non-enzymatic browning on cooking (Ma et al., 2013). Similar
525
526 268 considerations can be made for the BNCG samples (buffalo Mozzarella acidified with
527
528 269 commercial starters). These samples, however, clustered in an intermediate position between
529
530 270 buffalo and cow samples when the 50 more abundant OTUs are considered, suggesting that both
531
532 271 the type of starter (natural or selected) and the type of milk are possible drivers of bacterial
533
534 272 diversity in Mozzarella cheese.

530 273 Several species of non-starter lactic acid bacteria (NSLAB) belonging to the genera
531
532 274 *Lactobacillus*, *Lactococcus*, *Streptococcus*, *Leuconostoc*, and *Weissella* were frequently detected
533
534 275 in all groups of Mozzarella samples. NSLAB do not contribute to acidification during
535
536 276 cheesemaking, but they can play a significant role during ripening by using residual lactose and
537
538 277 other carbohydrates, citrate, peptide and aminoacids, giving rise to volatile aroma compounds.
539
540 278 Moreover, they can exert protective effects by producing bacteriocins and other antimicrobial
541
542 279 compounds (Ristagno et al., 2012). Recently, different amounts of some metabolites (namely
543
544 280 threonine and lactic acid dimer) were linked to different levels of NSLAB in buffalo and cow
545
546 281 Mozzarella cheese (Pisano et al., 2016). Several lactobacilli were more abundant in buffalo
547
548 282 Mozzarella produced with natural cultures, probably coming from NWC (De Filippis et al.,
549
550 283 2014). Another species more frequent in buffalo Mozzarella compared to cow mozzarella is
551
552 284 *Streptococcus equinus*. It is a commensal inhabitant of the gastrointestinal tract of mammals, but
553
554 285 also an opportunistic pathogen of humans and animals (Jans et al., 2014). Except for what is
555
556 286 reported in this study, not much is known about the presence of *S. equinus* in dairy processing
557
558 287 environments.

554 288 A large variety of psychrotrophic species belonging to different bacterial families were detected
555
556 289 in all samples. However, most of the psychrotrophic genera (e.g. *Anoxybacillus*, *Brochothrix*,

561
562 290 *Flavobacterium*, *Pseudomonas*, *Shewanella* and *Thermus*) were more abundant in cow
563 291 Mozzarella than in buffalo Mozzarella samples. These genera have been evidenced in NGS
564 292 separate studies in buffalo Mozzarella (Ercolini et al., 2012) and cow Mozzarella (Guidone et al.,
565 293 2016), nevertheless this is the first report in which the prevalence of some microbial taxa has
566 294 been differently associated to one type of cheese. Psychrotrophic populations are commonly
567 295 present as minor **components** in raw milk from several species, including cows, sheep, and goats,
568 296 but can become the most abundant genera in refrigerated milk. The higher prevalence of
569 297 psychrotrophic bacteria in cow Mozzarella suggests a stronger application of refrigeration during
570 298 the processing and/or storage of cow Mozzarella compared to buffalo's.
571 299 Usually, raw milk is not directly processed after milking and is stored under refrigerated
572 300 conditions until it is delivered to the dairy plant, where an additional storage at low temperature
573 301 for up to 48 h is possible. The excessive proliferation of psychrotolerant microorganisms during
574 302 cold storage increases the risk of milk and cheese spoilage. Indeed, such **species produce**
575 303 thermostable extracellular enzymes, with proteases and lipases being the most important.
576 304 Proteases can degrade milk proteins (mainly casein) producing a grey discoloration, bittering,
577 305 off-flavours, increase in viscosity and gelation, while lipases cause rancidity (Chen et al., 2003).
578 306 Moreover, some psychrotrophic taxa, such as *Pseudomonas* spp. and *Thermus* spp., have been
579 307 associated with cheese discoloration (Andreani et al., 2014; Quigley et al., 2016).
580 308 Thermotolerant taxa, e.g. *Anoxybacillus flavithermus* and *Methylobacterium* spp., were detected in
581 309 Mozzarella samples. *A. flavithermus*, frequently associated to cow Mozzarella samples, is a
582 310 sporeformer that can attach to stainless steel and develop into biofilms, suggesting that an
583 311 environmental contamination could be the source of this taxon in Mozzarella. In fact, spores can
584 312 overcome the pasteurization and, being sticky, attach to the pasteurizer inside in the heat
585 313 recovery portion, where the temperature is lower (Palmer et al., 2010). Thermotolerant species can
586 314 moreover easily survive through the high curd cooking and stretching temperatures, which might
587 315 be the main reason for their presence in Mozzarella cheese.
588 316 The cluster analysis carried out on Mozzarella samples produced using NWC showed two
589 317 distinct groups of samples, named Cluster 1 and Cluster 2. Cluster 1, which contained most
590 318 buffalo Mozzarella purchased from local market, was characterized by a higher prevalence of
591 319 lactic species belonging to the taxa *Lactococcus* spp., *Weissella viridescens* and *Streptococcus*
592 320 *vestibularis*. Psychrotrophic taxa were overrepresented in Cluster 2, which comprises most
593 321 samples collected in supermarkets. Some of these taxa are possibly involved in food and dairy

617
618 322 spoilage (Innocente et al., 2009; Stellato et al., 2015). Moreover, several of the species
619 323 overrepresented in Cluster 2 are potential pathogens. For example, *Plesiomonas* genus includes
620 324 species that might be associated to foodborne disease (Janda et al., 2016). *Enterobacter*
621 325 *hormaechei* is a pathogenic *Enterobacter* that has been previously isolated in cheese (Pangallo et
622 326 al., 2014). In general, it has been observed that Cluster 2 is enriched in bacteria usually
623 327 associated to lower quality compared to Cluster 1. Incidentally, Cluster 2 is the one containing
624 328 the higher proportion of products marketed on mass distribution circuit, while in Cluster 1 the
625 329 majority of products are marketed locally. One possible explanation of our findings is that
626 330 products marketed in the mass distribution circuit might experience longer exposures to
627 331 suboptimal temperatures, as well as longer production-to-consumption times, both potentially
628 332 resulting in a relative increase of psychrotolerant, food spoilage-related organisms. It should be
629 333 noted that, during processing of PDO buffalo Mozzarella, according to the procedural guidelines
630 334 the milk must be delivered to the dairy within the sixteenth hour from the milking, and
631 335 transformed into Mozzarella within the sixtieth hour from the first milking (Gobbetti et al.,
632 336 2018). Thus, it is possible that the PDO Mozzarella cheese, also if locally sold, is produced using
633 337 milk that has undergone more or less prolonged refrigeration. This may be the reason why three
634 338 BDN samples are included in Cluster 2.

635 339 In conclusion, this study confirmed the role of acidification method in the determination of the
636 340 microbiota, with samples using NWC mostly composed by *Lactobacillus* and *Streptococcus*
637 341 species and CS dominated by *Streptococcus* species alone. Metagenomics approach can leverage
638 342 differential abundance of bacterial species to confidently discriminate cow Mozzarella from
639 343 buffalo Mozzarella. Finally, two clusters of samples were identified composed by a majority of
640 344 products sampled at a local retail and in a mass retail, respectively. Differential analysis of the
641 345 microbiota of the two groups revealed that samples collected at mass retail usually have higher
642 346 prevalence of microorganisms related to food spoilage, thus suggesting that the metagenomics
643 347 approach can be a useful method for detecting critical issues in the storage of food products, such
644 348 as Mozzarella cheese.

645 349

673
674 350 **5 References**
675

- 676 351 Addeo, F., Pizzano, R., Nicolai, M.A., Caira, S., Chianese, L., 2009. Fast isoelectric focusing
677 352 and antipeptide antibodies for detecting bovine casein in adulterated water buffalo milk and
678 353 derived Mozzarella cheese. *J. Agric. Food Chem.* 57, 10063–10066. doi:10.1021/jf9020009
680 354 Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. *Genome*
682 355 *Biol.* 11, R106. doi:10.1186/gb-2010-11-10-r106
683 356 Andreani, N.A., Martino, M.E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., Bordin, P.,
685 357 Cardazzo, B., 2014. Tracking the blue: A MLST approach to characterise the *Pseudomonas*
687 358 *fluorescens* group. *Food Microbiol.* 39, 116–126. doi:10.1016/j.fm.2013.11.012
688 359 Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K.,
690 360 Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights,
691 361 D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M.,
692 362 Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunencko, T.,
693 363 Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community
694 364 sequencing data. *Nat. Methods* 7, 335–336. doi:10.1038/nmeth.f.303
695 365 Chao, A., 1984. Non-parametric estimation of the classes in a population. *Scand. J. Stat.* 11,
699 366 265–270. doi:10.2307/4615964
700 367 Chen, G., Chen, C., Lei, Z., 2017. Meta-omics insights in the microbial community profiling and
702 368 functional characterization of fermented foods. *Trends Food Sci. Technol.* 65, 23–31.
703 369 doi:10.1016/j.tifs.2017.05.002
704 370 Chen, L., Daniel, R.M., Coolbear, T., 2003. Detection and impact of protease and lipase
707 371 activities in milk and milk powders. *Int. Dairy J.* 13, 255–275. doi:10.1016/S0958-
708 372 6946(02)00171-1
709 373 Chevallier, S., Bertrand, D., Kohler, A., Courcoux, P., 2006. Application of PLS-DA in
712 374 multivariate image analysis. *J. Chemom.* 20, 221–229. doi:10.1002/cem.994
713 375 Coppola, S., Blaiotta, G., Ercolini, D., Moschetti, G., 2001. Molecular evaluation of microbial
715 376 diversity occurring in different types of Mozzarella cheese. *J. Appl. Microbiol.* 90, 414–
716 377 420. doi:10.1046/j.1365-2672.2001.01262.x
717 378 de Candia, S., De Angelis, M., Dunlea, E., Minervini, F., McSweeney, P.L.H., Faccia, M.,
720 379 Gobbetti, M., 2007. Molecular identification and typing of natural whey starter cultures and
721 380 microbiological and compositional properties of related traditional Mozzarella cheeses. *Int.*
722 381 *J. Food Microbiol.* 119, 182–191. doi:10.1016/j.ijfoodmicro.2007.07.062
723
724
725
726
727
728

- 729
730 382 De Filippis, F., La Storia, A., Stellato, G., Gatti, M., Ercolini, D., 2014. A selected core
731
732 383 microbiome drives the early stages of three popular Italian cheese manufactures. PLoS One
733 384 9, e89680. doi:10.1371/journal.pone.0089680
734
735 385 Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads.
736 386 Nat. Methods 10, 996–998. doi:10.1038/nmeth.2604
737
738 387 Ercolini, D., De Filippis, F., La Storia, A., Iacono, M., 2012. “Remake” by high-throughput
739 388 sequencing of the microbiota involved in the production of water buffalo Mozzarella
740 389 cheese. Appl. Environ. Microbiol. 78, 8142–8145. doi:10.1128/AEM.02218-12
741
742 390 Gobbetti, M., Neviani, E., Fox, P.F., 2018. The most traditional and popular Italian cheeses, in:
743 391 Gobbetti, M., Neviani, E., Fox, P.F. (Eds.), The Cheeses of Italy: Science and Technology.
744 392 Springer International Publishing, Cham, Switzerland, pp. 99–274.
745
746 393 Good, I.J., 1953. The population frequencies of species and the estimation of population
747 394 parameters. Biometrika 40, 237.
748
749 395 Gorrasi, G., Bugatti, V., Tammara, L., Vertuccio, L., Vigliotta, G., Vittoria, V., 2016. Active
750 396 coating for storage of Mozzarella cheese packaged under thermal abuse. Food Control 64,
751 397 10–16. doi:10.1016/j.foodcont.2015.12.002
752
753 398 Guidone, A., Zotta, T., Matera, A., Ricciardi, A., De Filippis, F., Ercolini, D., Parente, E., 2016.
754 399 The microbiota of high-moisture Mozzarella cheese produced with different acidification
755 400 methods. Int. J. Food Microbiol. 216, 9–17. doi:10.1016/j.ijfoodmicro.2015.09.002
756
757 401 Hurley, I.P., Coleman, R.C., Ireland, H.E., Williams, J.H.H., 2006. Use of sandwich IgG ELISA
758 402 for the detection and quantification of adulteration of milk and soft cheese. Int. Dairy J. 16,
759 403 805–812. doi:10.1016/j.idairyj.2005.07.009
760
761 404 Innocente, N., Marino, M., Marchesini, G., Biasutti, M., 2009. Presence of biogenic amines in a
762 405 traditional salted Italian cheese. Int. J. Dairy Technol. 62, 154–160. doi:10.1111/j.1471-
763 406 0307.2009.00479.x
764
765 407 Janda, M.J., Abbott, S.L., McIver, C.J., 2016. *Plesiomonas shigelloides* revisited. Clin.
766 408 Microbiol. Rev. 29, 349–374. doi:10.1128/CMR.00103-15
767
768 409 Jans, C., Meile, L., Lacroix, C., Stevens, M.J.A., 2014. Genomics, evolution, and molecular
769 410 epidemiology of the *Streptococcus bovis*/*Streptococcus equinus* complex (SBSEC). Infect.
770 411 Genet. Evol. 33, 419–436. doi:10.1016/j.meegid.2014.09.017
771
772 412 Lopparelli, R.M., Cardazzo, B., Balzan, S., Giaccone, V., Novelli, E., 2007. Real-time TaqMan
773 413 polymerase chain reaction detection and quantification of cow DNA in pure water buffalo
774
775
776
777
778
779
780
781
782
783
784

- 785
786 414 Mozzarella cheese: Method validation and its application on commercial samples. *J. Agric.*
787
788 415 *Food Chem.* 55, 3429–3434. doi:10.1021/jf0637271
- 789 416 Ma, X., James, B., Balaban, M.O., Zhang, L., Emanuelsson-Patterson, E.A.C., 2013. Quantifying
790
791 417 blistering and browning properties of Mozzarella cheese. Part I: Cheese made with different
792
793 418 starter cultures. *Food Res. Int.* 54, 912–916. doi:10.1016/j.foodres.2013.06.007
- 794 419 Marino, M., Innocente, N., Maifreni, M., Mounier, J., Cobo-Díaz, J.F., Coton, E., Carraro, L.,
795
796 420 Cardazzo, B., 2017. Diversity within Italian cheesemaking brine-associated bacterial
797
798 421 communities evidenced by massive parallel 16S rRNA gene tag sequencing. *Front.*
799 422 *Microbiol.* 8, 2119. doi:10.3389/fmicb.2017.02119
- 800 423 McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., Desantis, T.Z., Probst, A., Andersen,
801
802 424 G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit
803
804 425 ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* 6, 610–618.
805 426 doi:10.1038/ismej.2011.139
- 807 427 Palmer, J.S., Flint, S.H., Schmid, J., Brooks, J.D., 2010. The role of surface charge and
808
809 428 hydrophobicity in the attachment of *Anoxybacillus flavithermus* isolated from milk powder.
810 429 *J. Ind. Microbiol. Biotechnol.* 37, 1111–1119. doi:10.1007/s10295-010-0758-x
- 811 430 Pangallo, D., Šaková, N., Koreňová, J., Puškárová, A., Kraková, L., Valík, L., Kuchta, T., 2014.
812
813 431 Microbial diversity and dynamics during the production of May Bryndza cheese. *Int. J.*
814
815 432 *Food Microbiol.* 170, 38–43. doi:10.1016/j.ijfoodmicro.2013.10.015
- 816 433 Paulson, J.N., Colin Stine, O., Bravo, H.C., Pop, M., 2013. Differential abundance analysis for
817
818 434 microbial marker-gene surveys. *Nat. Methods* 10, 1200–1202. doi:10.1038/nmeth.2658
- 819 435 Pisano, M.B., Scano, P., Murgia, A., Cosentino, S., Caboni, P., 2016. Metabolomics and
820
821 436 microbiological profile of Italian Mozzarella cheese produced with buffalo and cow milk.
822
823 437 *Food Chem.* 192, 618–624. doi:10.1016/j.foodchem.2015.07.061
- 824 438 Quigley, L., O’Sullivan, O., Stanton, C., Beresford, T.P., Ross, R.P., Fitzgerald, G.F., Cotter,
825
826 439 P.D., 2013. The complex microbiota of raw milk. *FEMS Microbiol. Rev.* 37, 664–698.
827 440 doi:10.1111/1574-6976.12030
- 829 441 Quigley, L., Sullivan, D.J.O., Daly, D., Sullivan, O.O., Fitzgerald, G.F., Mcsweeney, P.L.H.,
830
831 442 Giblin, L., Sheehan, J.J., Cotter, D., 2016. *Thermus* and the pink discoloration defect in
832
833 443 cheese. *mSystems* 1, e00023-16. doi:10.1128/mSystems.00023-16
- 834 444 R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for
835
836 445 Statistical Computing, Vienna, Austria.

841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896

446 Ristagno, D., Hannon, J.A., Beresford, T.P., McSweeney, P.L.H., 2012. Effect of a bacteriocin-
447 producing strain of *Lactobacillus paracasei* on the nonstarter microflora of Cheddar cheese.
448 *Int. J. Dairy Technol.* 65, 523–530.

449 Segat, A., Biasutti, M., Iacumin, L., Comi, G., Baruzzi, F., Carboni, C., Innocente, N., 2014. Use
450 of ozone in production chain of high moisture Mozzarella cheese. *LWT - Food Sci.*
451 *Technol.* 55, 513–520. doi:10.1016/j.lwt.2013.10.029

452 Shannon, C.E., 1948. A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423.

453 Stellato, G., De Filippis, F., La Storia, A., Ercolini, D., 2015. Coexistence of lactic acid bacteria
454 and potential spoilage microbiota in a dairy processing environment. *Appl. Environ.*
455 *Microbiol.* 81, 7893–7904. doi:10.1128/AEM.02294-15

456 Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid
457 assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.*
458 73, 5261–5267. doi:10.1128/AEM.00062-07

897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952

460 **Figure captions**

461
462 **Figure 1.** Alpha diversity measured as estimated number of OTUs in five different types of
463 mozzarella cheese. Pairwise difference between groups was assessed using Wilcoxon test. Box-
464 plots labeled with different letters are significantly different from each other.

465 **BDN:** Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG:**
466 Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at
467 mass retailers; **BNNG:** Buffalo mozzarella without certification, acidified with Natural Whey
468 Culture and collected at mass retailers; **BNCG:** Buffalo mozzarella without certification,
469 acidified with commercial starters and collected at mass retailers; **CC:** Cow mozzarella acidified
470 with commercial starters and collected at mass retailers.

471
472 **Figure 2.** Abundance of bacterial families represented by at least 0.1% of reads in at least on
473 sample.

474 **BDN:** Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG:**
475 Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at
476 mass retailers; **BNNG:** Buffalo mozzarella without certification, acidified with Natural Whey
477 Culture and collected at mass retailers; **BNCG:** Buffalo mozzarella without certification,
478 acidified with commercial starters and collected at mass retailers; **CC:** Cow mozzarella acidified
479 with commercial starters and collected at mass retailers.

480
481 **Figure 3.** Clustering of samples based on the 50 most represented species.

482 **BDN:** Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG:**
483 Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at
484 mass retailers; **BNNG:** Buffalo mozzarella without certification, acidified with Natural Whey
485 Culture and collected at mass retailers; **BNCG:** Buffalo mozzarella without certification,
486 acidified with commercial starters and collected at mass retailers; **CC:** Cow mozzarella acidified
487 with commercial starters and collected at mass retailers.

488
489 **Figure 4.** Clustering of samples obtained by the use of Natural Whey Culture.

490 **BDN:** Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG:**
491 Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at

953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008

492 mass retailers; **BNNG**: Buffalo mozzarella without certification, acidified with Natural Whey
493 Culture and collected at mass retailers.

Figure 1

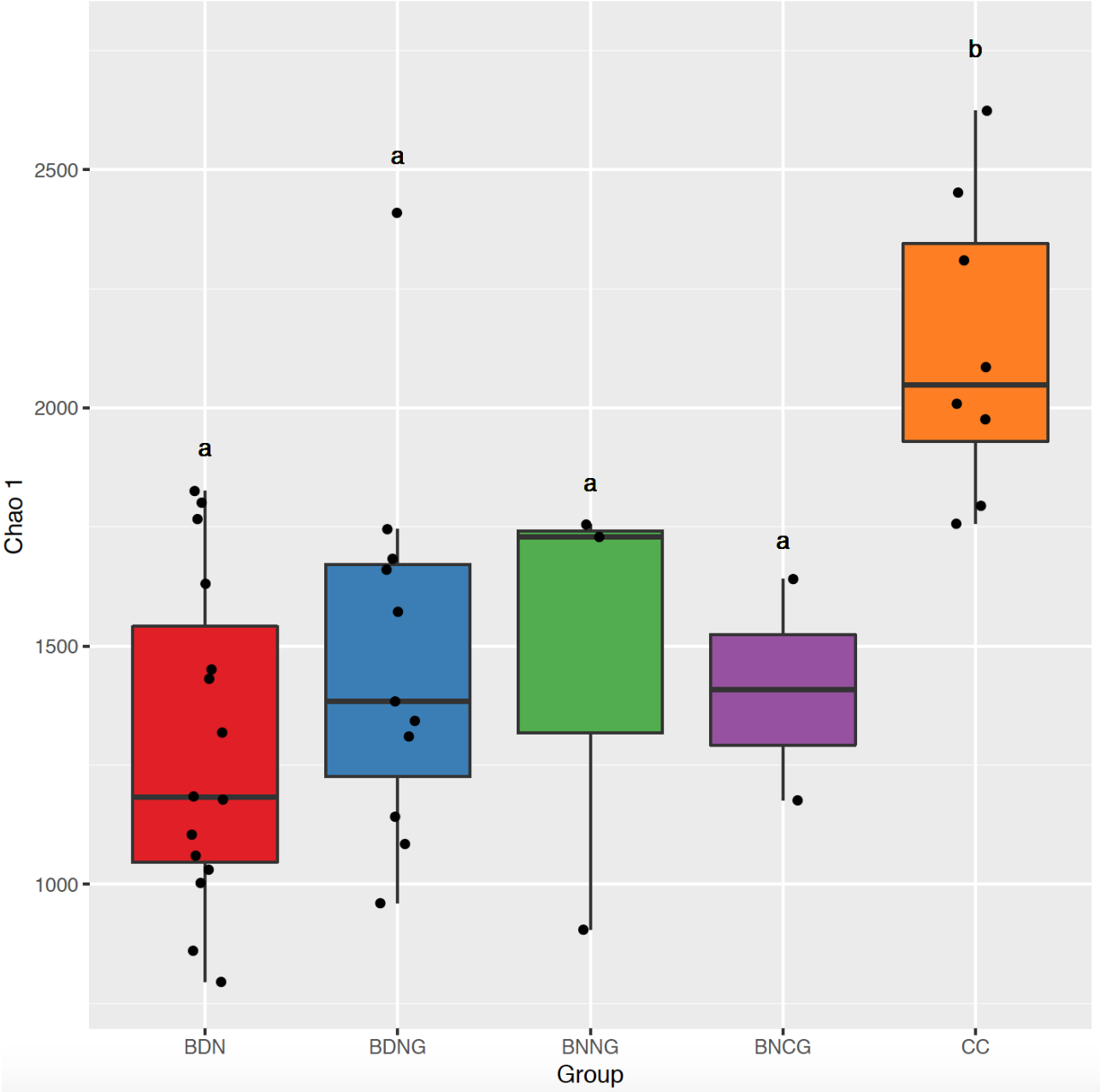


Figure 3

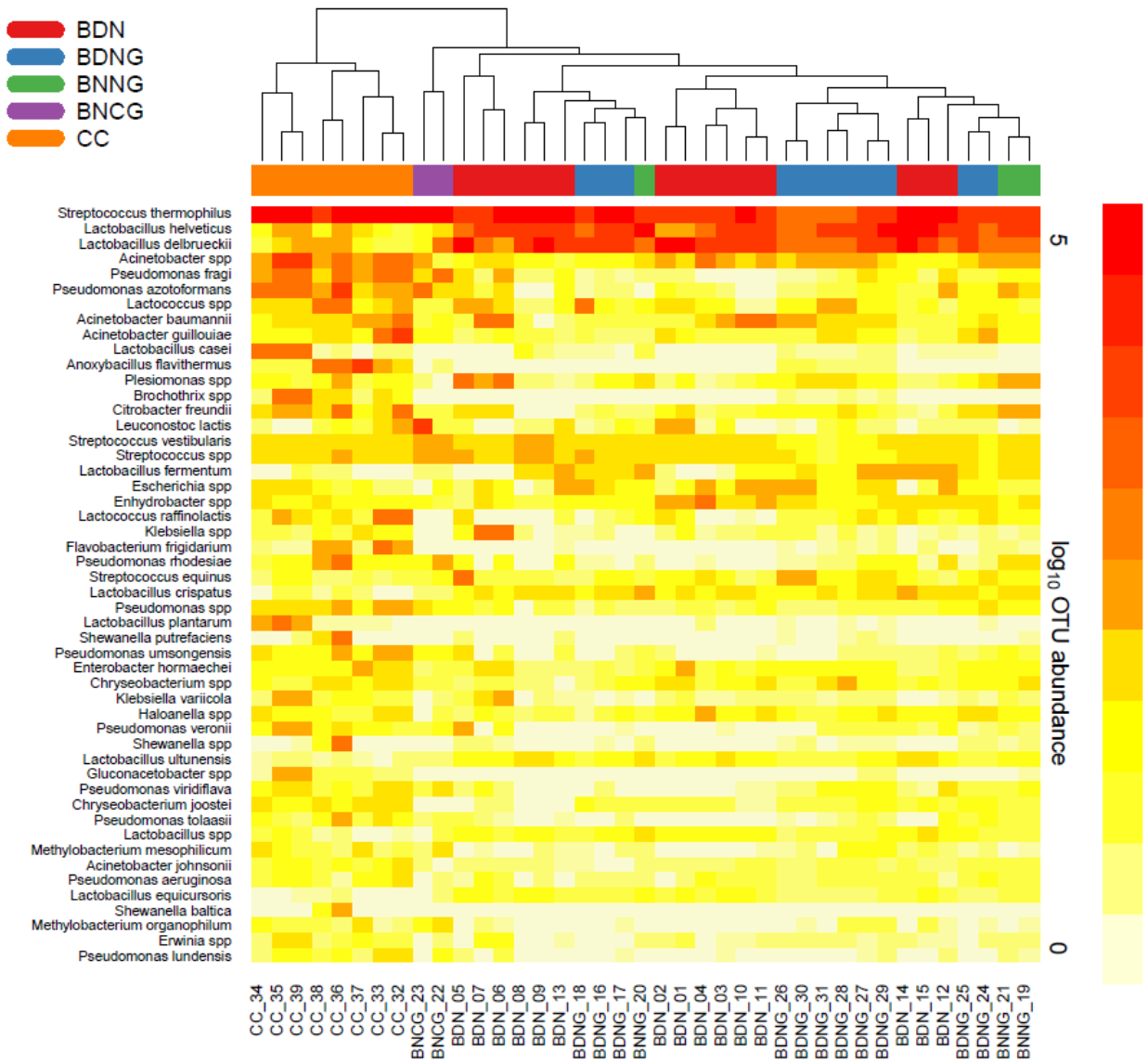
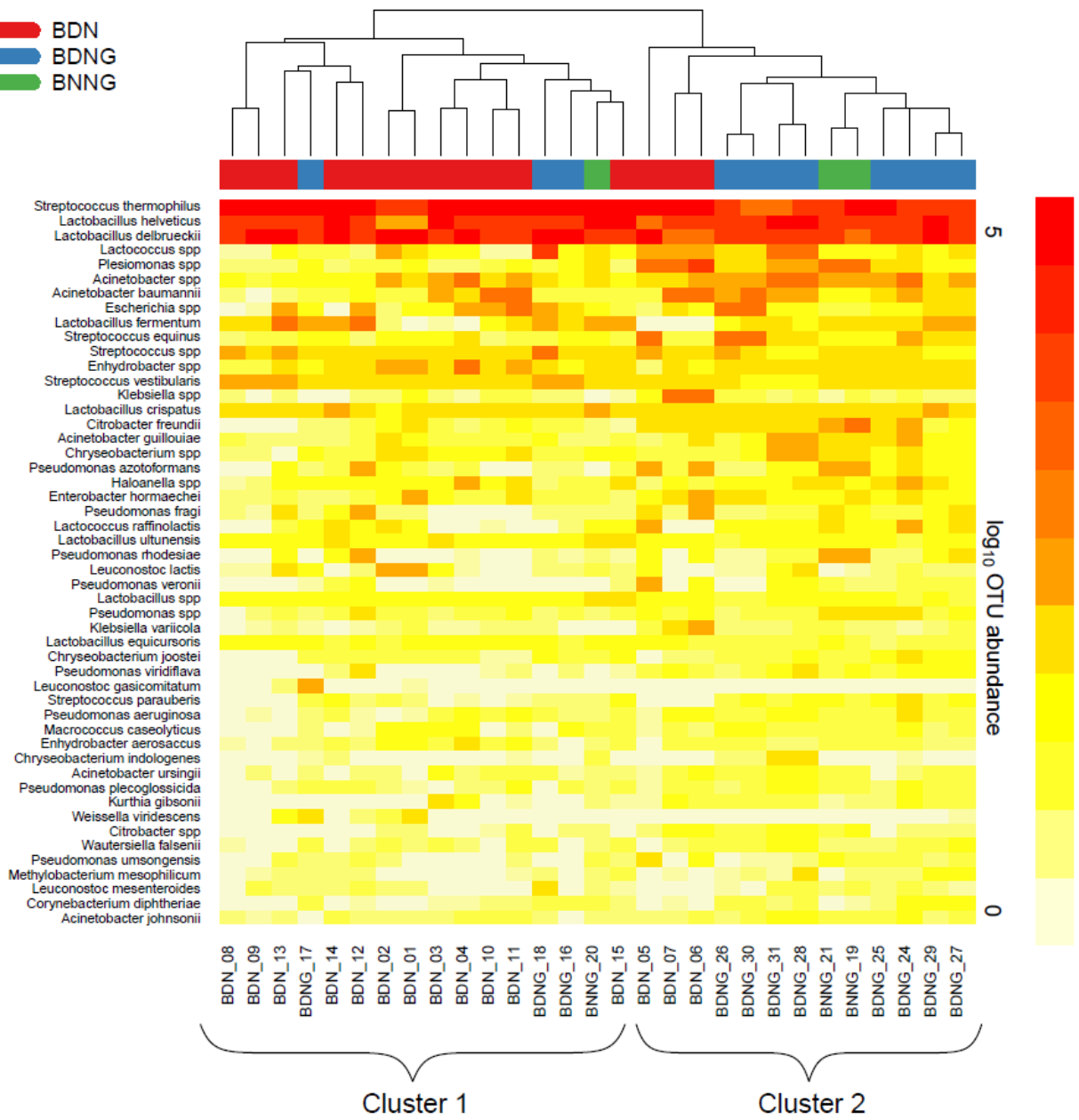


Figure 4

BDN
BDNG
BNNG



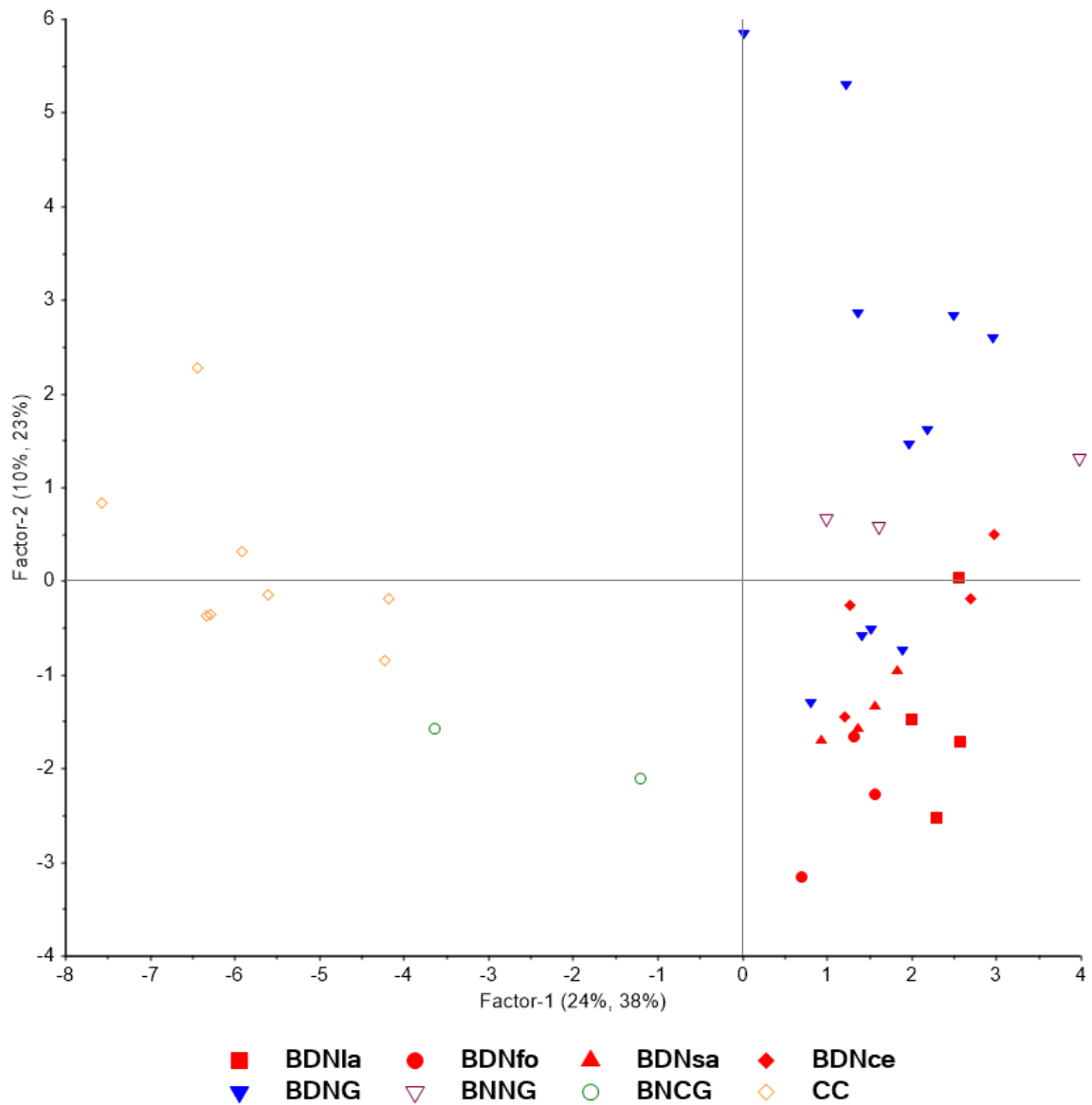


Figure S1: Score-plot of Mozzarella samples from a PLS-DA model of classification based on their OUTF profile. BDNla, BDNfo, BDNsa and BDNce were samples coming respectively from the provinces of Latina, Foggia, Salerno, and Caserta.

1 Table 1. Summary statistics of the study samples

Sample [§]	Type*	Certification [#]	Sampling point [¶]	Acidification [†]	Reads/sample	Identified OTUs/Sample	Estimated OTUs/sample [^]
BDN_01	BM	PDO	L	NWC	76,736	570	1,003
BDN_02	BM	PDO	L	NWC	45,170	463	795
BDN_03	BM	PDO	L	NWC	135,946	747	1,183
BDN_04	BM	PDO	L	NWC	108,787	912	1,430
BDN_05	BM	PDO	L	NWC	100,464	737	1,177
BDN_06	BM	PDO	L	NWC	136,682	1,065	1,825
BDN_07	BM	PDO	L	NWC	115,630	934	1,631
BDN_08	BM	PDO	L	NWC	124,178	688	1,103
BDN_09	BM	PDO	L	NWC	131,778	675	1,031
BDN_10	BM	PDO	L	NWC	126,677	788	1,318
BDN_11	BM	PDO	L	NWC	122,042	918	1,801
BDN_12	BM	PDO	L	NWC	159,580	1,044	1,767
BDN_13	BM	PDO	L	NWC	89,681	609	1,060
BDN_14	BM	PDO	L	NWC	55,518	477	860
BDN_15	BM	PDO	L	NWC	121,617	829	1,451
BDNG_16	BM	PDO	M	NWC	97,084	630	1,084
BDNG_17	BM	PDO	M	NWC	115,616	760	1,342
BDNG_18	BM	PDO	M	NWC	111,893	752	1,142
BDNG_24	BM	PDO	M	NWC	99,252	1,039	1,745
BDNG_25	BM	PDO	M	NWC	47,251	510	960
BDNG_26	BM	PDO	M	NWC	88,912	778	1,310
BDNG_27	BM	PDO	M	NWC	132,776	960	1,682
BDNG_28	BM	PDO	M	NWC	75,006	819	1,384
BDNG_29	BM	PDO	M	NWC	176,617	878	1,571
BDNG_30	BM	PDO	M	NWC	189,981	1,193	2,410
BDNG_31	BM	PDO	M	NWC	125,197	873	1,660
BNNG_19	BM	None	M	NWC	108,266	956	1,754
BNNG_20	BM	None	M	NWC	87,127	567	904
BNNG_21	BM	None	M	NWC	110,637	949	1,729
BNCG_22	BM	None	M	CS	57,720	770	1,640
BNCG_23	BM	None	M	CS	108,657	745	1,176
CC_32	CM	None	M	CS	169,503	1,313	2,451
CC_33	CM	None	M	CS	77,159	883	1,757
CC_34	CM	None	M	CS	154,577	1,384	2,309
CC_35	CM	None	M	CS	157,778	1,269	2,085
CC_36	CM	None	M	CS	151,638	1,159	2,009
CC_37	CM	None	M	CS	92,067	958	1,794
CC_38	CM	None	M	CS	216,852	1,569	2,623
CC_39	CM	None	M	CS	109,899	1,102	1,975

2 §Legend of names prefixes: **BDN**: Buffalo mozzarella with PDO certification, acidified with
 3 Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with
 4 Natural Whey Culture and collected at mass retailers; **BNNG**: Buffalo mozzarella without

57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112

5 certification, acidified with Natural Whey Culture and collected at mass retailers; **BNCG**:
6 Buffalo mozzarella without certification, acidified with commercial starters and collected at mass
7 retailers; **CC**: Cow mozzarella acidified with commercial starters and collected at mass retailers;
8 *, BM=buffalo Mozzarella, CM=cow Mozzarella; #, presence of PDO certification (PDO) or not
9 (none); ¥, local (L) or mass (M) retailer; †acidification with NWC=natural whey culture, or
10 CS=commercial starter. ^, According to Chao (1984)

1 Table 2. Differentially abundant OTUs between buffalo and cow mozzarella cheese. Only OTUs
 2 with a False Discovery Rate (FDR) <0.05 and present in more than 0.1% of the reads in at least
 3 one sample are shown.

Taxa	Reads in BM*	Reads in CM§	FDR	log2ratio
<i>Anoxybacillus flavithermus</i>	262	29,533	0.0015	-6.81
<i>Brochothrix</i>	31	12,634	0.0002	-8.63
<i>Erwinia</i>	230	467	0.0271	-1.02
<i>Flavobacterium frigidarium</i>	53	11,530	0.0046	-7.74
<i>Gluconacetobacter</i>	22	1,593	0.0107	-6.11
<i>Lactobacillus crispatus</i>	7,895	43	0.0002	7.49
<i>Lactobacillus delbrueckii</i>	622,438	6,973	0.0002	6.48
<i>Lactobacillus fermentum</i>	14,007	30	0.0099	8.82
<i>Lactobacillus helveticus</i>	736,445	4,051	0.0004	7.51
<i>Lactobacillus</i>	1,320	84	0.0007	3.96
<i>Lactobacillus ultunensis</i>	2,235	20	0.0002	6.73
<i>Pseudomonas azotoformans</i>	6,468	28,603	0.0295	-2.14
<i>Pseudomonas fragi</i>	6,507	27,043	0.0098	-2.06
<i>Pseudomonas lundensis</i>	62	539	0.0002	-3.10
<i>Pseudomonas</i>	1,740	5,518	0.0181	-1.66
<i>Pseudomonas umsongensis</i>	428	5,150	0.0007	-3.59
<i>Ruminococcus</i>	63	363	0.0051	-2.51
<i>Shewanella baltica</i>	0	917	0.0107	-9.84
<i>Shewanella putrefaciens</i>	23	6,092	0.0470	-7.99
<i>Streptococcus equinus</i>	16,585	130	0.0237	6.98
<i>Thermus thermophilus</i>	16	557	0.0030	-5.04

4 *, Buffalo Mozzarella samples: BDN, BDNG, BNCG and BNNG; §, Cow Mozzarella samples:

5 CC

1 Table 3. Differential abundance of species between Cluster1 and Cluster2. Only OTUs with FDR
 2 < 0.05 and represented by more than 0.1% of the reads in at least one sample are shown are
 3 listed. “Unclassified” collects all OTUs that are not characterized at the genus level.

Taxa	Reads in Cluster1	Reads in Cluster2	FDR	log2ratio
<i>Acinetobacter baumannii</i>	10688	21264	0.0056	-0.99
<i>Acinetobacter guillouiae</i>	348	4449	0.0002	-3.67
<i>Acinetobacter ursingii</i>	117	395	0.0129	-1.75
<i>Anoxybacillus flavithermus</i>	0	260	0.0003	-8.03
<i>Chryseobacterium</i>	971	3564	0.0498	-1.87
<i>Chryseobacterium indologenes</i>	9	494	0.0002	-5.63
<i>Chryseobacterium joostei</i>	192	620	0.0179	-1.68
<i>Citrobacter</i>	17	379	0.0002	-4.40
<i>Citrobacter freundii</i>	512	7100	8.83E-06	-3.79
<i>Corynebacterium diphtheriae</i>	54	391	0.0134	-2.83
<i>Enterobacter hormaechei</i>	1195	2559	0.0004	-1.10
<i>Erwinia</i>	10	118	0.0025	-3.43
<i>Klebsiella</i>	91	8637	0.0222	-6.55
<i>Lactococcus</i>	11303	9326	0.0118	0.28
<i>Lactococcus raffinolactis</i>	355	2298	0.0004	-2.69
<i>Methylobacterium mesophilicum</i>	35	365	0.0498	-3.34
<i>Methylobacterium organophilum</i>	18	181	0.0201	-3.26
<i>Plesiomonas</i>	420	25746	2.09E-07	-5.93
<i>Pseudomonas</i>	222	767	0.0214	-1.78
<i>Pseudomonas aeruginosa</i>	142	513	0.0046	-1.84
<i>Pseudomonas azotoformans</i>	1037	3056	0.0039	-1.56
<i>Pseudomonas fragi</i>	931	2237	0.0092	-1.26
<i>Pseudomonas plecoglossicida</i>	151	326	0.0145	-1.10
<i>Pseudomonas rhodesiae</i>	807	1965	0.0160	-1.28
<i>Pseudomonas tolaasii</i>	97	249	0.0400	-1.35
<i>Pseudomonas umsongensis</i>	41	292	0.0193	-2.80
<i>Pseudomonas veronii</i>	65	1245	0.0018	-4.24
<i>Pseudomonas viridiflava</i>	263	724	0.0012	-1.46
<i>Streptococcus equinus</i>	682	15604	0.0005	-4.51
<i>Streptococcus parauberis</i>	161	604	0.0004	-1.90
<i>Streptococcus vestibularis</i>	2979	1593	0.0500	0.90
<i>Wautersiella falsenii</i>	21	187	0.0060	-3.10
<i>Weissella viridescens</i>	330	0	0.0012	8.37
Unclassified	261	31747	1.34E-05	-6.92