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 Dubsky 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_Mozzarelle_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]

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

Manuscript FM_2018_722 by Marino et al. – Replies to the referees' comments

Reviewer 1

The paper FM_2018_772 "Metagenomic profiles of different types of Italian high-moisture Mozzarella cheese" – evaluated the microbiota, by Illumina MiSeq approach, 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.

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

I have some advices in order to improve the manuscript:

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

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

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

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

Minor issues:

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

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

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

Done.

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

Done.

Reviewer 2

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

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

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

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

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

L87. Please, used biota instead of flora.

Done

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

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

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

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

Were the samples collected and analyzed in triplicate?

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

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

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

L233. It’s statistically significant those difference in the microbial diversity?

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

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 Choa1 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.

Reviewer 3

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

1. Abstract, page 2, line 25, *Corynebacterium* belongs to the psychrotrophic genera. Results show that buffalo Mozzarella was enriched with *Lactococcus*, *Streptococcus*, and *Weissella*, instead. Thank you, we corrected the abstract accordingly. We did not mention Weissella in the abstract because it is a rare genus and we decided not to report extensive results for rare genera.
2. The second highlight exceeds the maximum number of characters. Try dividing it into two sentences. Thank you, we followed your suggestion.
3. Page 4, line 87, page 9, line 232 and page 11, lines 280 and 287, use microbiota instead of microflora or flora. Thank you, we followed your suggestion.
4. Page 4, lines 90-93, a reference is missing for that paragraph. We added the missing reference.
5. Page 5, line 122, please give a more explicit description of the second amplification step. Which flow-cell binding domains and unique indices? We added the missing information.
6. Page 5, line 124, what does SRA stands for? SRA stands for Sequence Reads Archive (https://www.ncbi.nlm.nih.gov/sra). We added the full name in the methods section.
7. Table 1, consider changing the following column descriptions: Reads/sample, Identified OTUs/sample and Estimated OTUs/sample§ Then in the table notes § (Chao, 1984). We followed reviewer's suggestion.
8. I suggest that Figure S1 should not be supplementary, but part of the main manuscript. We followed reviewer's suggestion. Figure S1 is now Figure 1. Numbers of all other figures have been shifted accordingly.
9. Page 7, line 156, use approximation instead of proxy. We followed reviewer's suggestion.
10. Page 7, line 171, in table S1 there are 74 OTUs, instead of 75. We incorrectly counted the header as an OTU. We now report the correct number of OTUs (74).
11. In the caption of Figure 3, for intelligibility, use natural whey culture instead of NWC. Done.
12. Table 2, please state that BM is buffalo Mozzarella and CM is cow Mozzarella. In the table notes, explain what FDR means. We followed reviewer's suggestion.
13. Page 10, line 265, the paper by Martino et al. (2013) refers to a bacteriocin produced by *Pediococcus pentosaceus*. In the analyzed Mozzarella cheeses, *Pediococcus* was not identified. Instead, there are plenty of references related to bacteriocins produced by the 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

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

Marilena Marino[a,*], Giorgia Dubsky de Wittenau[b], Elena Saccà[a], Federica Cattonaro[b], Alessandro Spadotto[b], Nadia Innocente[a], Slobodanka Radovic[b], Edi Piasentier[a], Fabio Marroni[b,*]

[a] Dipartimento di Scienze Agroalimentari, Ambientali e Animali, Università di Udine, via Sondrio 2/A, 33100 Udine, Italy

[b] IGA Technology Services s.r.l., Via J. Linussio 51, 33100, Udine, Italy

Corresponding author:

Marilena Marino: Tel. +39 0432 558150; E-mail address: marilena.marino@uniud.it

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

High-moisture Mozzarella cheese

Microbiota

Next Generation Sequencing

Psychrotrophs

Metagenomics

1 Introduction

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

During the last decades, several methodologies have been applied to characterize Mozzarella cheese with the aim to ensure high quality and safety standards. Polymerase chain reaction (PCR) has been employed to detect species-specific DNA sequences in milk and cheese (Lopparelli et al., 2007), and isoelectric focusing, reversed-phase liquid chromatography, mass spectrometry and enzymatic assays to check the presence of specific buffalo and cow proteins in milk and cheese (Addeo et al., 2009; Hurley et al., 2006). Recently, a metabolomic approach based on gas-chromatography mass-spectrometry coupled with the analysis of the composition of predominant cultivable microbiota has been used to discriminate different types of Mozzarella cheese and to protect the authenticity of PDO Mozzarella di Bufala Campana cheese (Pisano et al., 2016). Due to the high water content and relatively high pH, microbial spoilage of Mozzarella cheese might occur, caused by proteolytic and/or lipolytic microorganisms that can cause unwanted modifications of the texture, off-odors or discolorations (Andreani et al., 2014; Segat et al., 2014). In the last decade, the food microbiology has been deeply revolutionized by the use of Next Generation Sequencing (NGS) technologies, which can provide a thorough analysis of microbial diversity present in a food sample, producing much deeper output than more commonly used culture-independent approaches (Chen et al., 2017; Marino et al., 2017). Currently, only two studies have been carried out to study the microbial diversity of Mozzarella cheese using an NGS approach (Ercolini et al., 2012; Guidone et al., 2016). However, the microbiota of the buffalo and the cow Mozzarella cheese has been studied in separate papers, which makes it difficult to understand the potential of NGS-based metagenomics in distinguishing products obtained with milk of different animal origins and different technologies. Moreover, the only study carried out on cow Mozzarella cheese analyzed the cheese microbiota after a 5-d refrigerated storage, which could have favored the growth of psychrotrophic microorganisms and hence modified to some extent the composition of the native microbiota of Mozzarella cheese (Guidone et al., 2016).

The objective of this study was to analyze the composition of the microbiota of different types of high-moisture Mozzarella cheese by using an NGS approach. In order to identify possible drivers of the bacterial diversity, samples with different characteristics in term of type of milk (water-buffalo and cow), acidification (by natural whey culture or selected starter), certification status (PDO or non-PDO), and sampling point (local or mass retailer) were included in the study.

2 Materials and Methods

2.1 Samples collection

Thirty-nine samples of high-moisture Mozzarella cheese were collected in local or mass retailers to maximize the variability of factors potentially affecting the cheese microbiota composition, namely type of milk, acidification system, certification status, and sampling point (Table 1). Three main groups of buffalo Mozzarella and cow Mozzarella samples were collected as follows: (i) 15 PDO Mozzarella cheese produced with buffalo milk and acidified with NWC, and purchased at local diaries in the main districts of the production area, (ii) 11 PDO Mozzarella cheese produced in the PDO area with buffalo milk and acidified with NWC, but collected in supermarkets, and (iii) 13 non-PDO Mozzarella cheese collected in supermarkets, including buffalo Mozzarella acidified with CS, buffalo Mozzarella acidified with NWC, and cow's milk Mozzarella acidified with CS.

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

1) BDN: Buffalo mozzarella with PDO certification and acidified with Natural Whey Culture (15 samples)
2) BDNG: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at mass retailers (11 samples)
3) BNNG: Buffalo mozzarella without certification, acidified with Natural Whey Culture and collected at mass retailers (3 samples)
4) BNCG: Buffalo mozzarella without certification, acidified with commercial starters and collected at mass retailers (2 samples)
5) CC: Cow mozzarella acidified with commercial starters and collected at mass retailers (8 samples)

2.2 DNA extraction and sequencing

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

The bacterial diversity was obtained by the library preparation and sequencing of the 16S rRNA gene. The following two amplification steps were performed: an initial PCR amplification using 16S locus specific PCR primers (16S-341F 5'-CCTACGGGNGGCWGCAG-3' and 16S-805R 5'-GACTACHVGGGTATCTAATCC-3') and a subsequent amplification integrating relevant flow-cell binding domains (5'-TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-3' for the For primer and 5'-GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-3' for the reverse overhang) and unique indices selected among those available Nextera XT Index Kits combined according to manufacturer's instructions (Illumina, CA). Libraries were sequenced in a MiSeq (Illumina, CA) in paired end with 300-bp read length. Raw reads are available on Sequence Reads Archive under the accession SRP156292.

2.3 Data analysis

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

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

addition to *Acinetobacter*, a variety of other psychrotrophic genera (including *Corynebacterium*, *Flavobacterium*, *Chryseobacterium*, *Pseudomonas*, *Shewanella*, *Escherichia*, and *Enterobacter*) were found in all samples, although with different abundances within the groups.

Figure 3 shows the heatmap of the Mozzarella samples clustered by Euclidean distance computed based on the 50 more abundant OTUs. The cluster clearly separated cow Mozzarella from buffalo Mozzarella. In addition, the samples of buffalo Mozzarella obtained using commercial starters (BNCG samples) were in an intermediate position between cow and buffalo milk Mozzarella samples.

Considering the strong separation between cow Mozzarella and buffalo Mozzarella observed in the cluster, an enrichment test contrasting the samples belonging to the two categories was carried out to identify the species responsible for the differentiation. Table 2 shows the list of differentially abundant species between buffalo and cow Mozzarella samples. Buffalo Mozzarella samples showed a higher prevalence of *Lactobacillus* species, in addition to the species *Streptococcus equinus*, while no significant difference in abundance of *Streptococcus thermophilus* was observed. Conversely, in cow Mozzarella samples a higher prevalence of several psychrophilic taxa, including *Brochothrix*, *Erwinia*, *Flavobacterium*, *Pseudomonas*, and *Shewanella*, as well as thermoduric and spore-forming genera, as *Anoxybacillus flavithermus* and *Thermus thermophilus*, was observed.

To further characterize differences in the microbiota of Mozzarella acidified with NWC (BDN, BDNG and BNNG samples), a cluster analysis was carried out after removing samples obtained with CS. The analysis produced two main clusters (Figure 4). Cluster 1 included most samples purchased from local market (12 out of 15 samples), and Cluster 2 comprises most samples (10 out of 14 samples) collected in supermarkets. A similar conclusion can be drawn from the score plot of PLS-DA (Figure S1), where along Factor 1, CC and BNCG are clearly discriminated from buffalo Mozzarella obtained by acidification with NWC. Moreover, along Factor 2, NWC samples show the tendency to split into two populations according to the market target.

To further investigate the differences between samples belonging to Cluster 1 and Cluster 2, we performed a test for differential enrichment in OTUs between the clusters. Results are listed in Table 3. Several psychrotrophic genera were overrepresented in Cluster 2, including *Acinetobacter*, *Chryseobacterium*, *Citrobacter*, *Corynebacterium*, and *Pseudomonas*.

Conversely, in Cluster 1 *Lactococcus* spp., *Streptococcus vestibularis* and *Weissella viridescens* were present with significantly higher prevalence than in Cluster2.

4 Discussion

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

The Chao1 diversity index in cow Mozzarella microbiota was higher than that of buffalo Mozzarella (Table S2) (two tailed t-test, $p<10^{-4}$), the same was true for the Good's coverage ($p=0.0255$) and for the Shannon index, although the latter difference was not statistically significant. These observations are in contrast with the data presented in the only recent report on the microbiological profile of Mozzarella cheese produced with buffalo and cow milk, in which the authors identified a larger number of species in buffalo mozzarella (Pisano et al., 2016). However, the authors isolated a very small number of strains and explored the Mozzarella diversity using only culture-based techniques, which are known to have low sensitivity and may lead to an underestimation of microbial diversity present in food environments. Cow Mozzarella samples were all acidified with commercial starters, which is known to reduce the diversity of microbiota in cheese (Coppola et al., 2001). However, in this study, a different observation was made. In fact, with the exception of the lactic starter microbiota, the samples of cow's milk Mozzarella were characterized by a higher microbial diversity than buffalo Mozzarella samples. This might be attributed to a different microbial composition of bovine milk compared to that of the buffalo. In fact, although at present there are no data comparing the composition of the milk microbiota of the two species obtained using NGS techniques, an overview derived from a number of separate studies allowed evidencing a greater number of bacterial genera present in cow's milk (Quigley et al., 2013).

The starter composition has a major effect on the microbiota of the final Mozzarella. In water buffalo samples acidified with NWC the species *L. delbrueckii* subsp. *bulgaricus*, *L. helveticus* and *S. thermophilus* were present with a similar prevalence. The presence of these species reflects the microbial composition of NWC used for acidification processes, where these LAB assure lactose fermentation, curd ripening and formation of a typical aroma profile. Mozzarella samples produced with cow milk were instead dominated by *S. thermophilus*, confirming previous findings (Guidone et al., 2016; Pisano et al., 2016), whereas the two thermophilic lactobacilli were less represented within the microbiota. This could be related to the use of commercial starters, which is quite common in cow Mozzarella, and usually consist of *S. thermophilus* alone or associated with *L. delbrueckii* in smaller concentrations, in order to avoid the risk of excessive secondary proteolysis that might take place in a high moisture environment (Pisano et al., 2016). Anyway, *L. helveticus* and *L. delbrueckii* subsp. *bulgaricus* were present as sub-dominant LAB. Galactose-fermenting *L. helveticus* could help to reduce the accumulation of galactose, which is not fermented by *S. thermophilus* and *L. delbrueckii* subsp. *bulgaricus*, thus reducing the risk of non-enzymatic browning on cooking (Ma et al., 2013). Similar considerations can be made for the BNCG samples (buffalo Mozzarella acidified with commercial starters). These samples, however, clustered in an intermediate position between buffalo and cow samples when the 50 more abundant OTUs are considered, suggesting that both the type of starter (natural or selected) and the type of milk are possible drivers of bacterial diversity in Mozzarella cheese.

Several species of non-starter lactic acid bacteria (NSLAB) belonging to the genera *Lactobacillus*, *Lactococcus*, *Streptococcus*, *Leuconostoc*, and *Weissella* were frequently detected in all groups of Mozzarella samples. NSLAB do not contribute to acidification during cheesemaking, but they can play a significant role during ripening by using residual lactose and other carbohydrates, citrate, peptide and aminoacids, giving rise to volatile aroma compounds. Moreover, they can exert protective effects by producing bacteriocins and other antimicrobial compounds (Ristagno et al., 2012). Recently, different amounts of some metabolites (namely threonine and lactic acid dimer) were linked to different levels of NSLAB in buffalo and cow Mozzarella cheese (Pisano et al., 2016). Several lactobacilli were more abundant in buffalo Mozzarella produced with natural cultures, probably coming from NWC (De Filippis et al., 2014). Another species more frequent in buffalo Mozzarella compared to cow mozzarella is *Streptococcus equinus*. It is a commensal inhabitant of the gastrointestinal tract of mammals, but also an opportunistic pathogen of humans and animals (Jans et al., 2014). Except for what is reported in this study, not much is known about the presence of *S. equinus* in dairy processing environments.

A large variety of psychrotrophic species belonging to different bacterial families were detected in all samples. However, most of the psychrotrophic genera (e.g. *Anoxybacillus*, *Brochothrix*, *Flavobacterium*, *Pseudomonas*, *Shewanella* and *Thermus*) were more abundant in cow Mozzarella than in buffalo Mozzarella samples. These genera have been evidenced in NGS separate studies in buffalo Mozzarella (Ercolini et al., 2012) and cow Mozzarella (Guidone et al., 2016), nevertheless this is the first report in which the prevalence of some microbial taxa has been differently associated to one type of cheese. Psychrotrophic populations are commonly present as minor components in raw milk from several species, including cows, sheep, and goats, but can become the most abundant genera in refrigerated milk. The higher prevalence of psychrotrophic bacteria in cow Mozzarella suggests a stronger application of refrigeration during the processing and/or storage of cow Mozzarella compared to buffalo's.

Usually, raw milk is not directly processed after milking and is stored under refrigerated conditions until it is delivered to the dairy plant, where an additional storage at low temperature for up to 48 h is possible. The excessive proliferation of psychrotolerant microorganisms during cold storage increases the risk of milk and cheese spoilage. Indeed, such species produce thermostable extracellular enzymes, with proteases and lipases being the most important. Proteases can degrade milk proteins (mainly casein) producing a grey discoloration, bittering, off-flavours, increase in viscosity and gelation, while lipases cause rancidity (Chen et al., 2003). Moreover, some psychrotrophic taxa, such as *Pseudomonas* spp. and *Thermus* spp., have been associated with cheese discoloration (Andreani et al., 2014; Quigley et al., 2016).

Thermoduric taxa, e.g. *Anoxybacillus flavithermus* and *Methylobacterium* spp., were detected in Mozzarella samples. *A. flavithermus*, frequently associated to cow Mozzarella samples, is a sporeformer that can attach to stainless steel and develop into biofilms, suggesting that an environmental contamination could be the source of this taxon in Mozzarella. In fact, spores can overcome the pasteurization and, being sticky, attach to the pasteurizer inside in the heat recovery portion, where the temperature is lower (Palmer et al., 2010). Thermoduric species can moreover easily survive through the high curd cooking and stretching temperatures, which might be the main reason for their presence in Mozzarella cheese.

The cluster analysis carried out on Mozzarella samples produced using NWC showed two distinct groups of samples, named Cluster 1 and Cluster 2. Cluster 1, which contained most buffalo Mozzarella purchased from local market, was characterized by an higher prevalence of lactic species belonging to the taxa *Lactococcus* spp., *Weissella viridescens* and *Streptococcus vestibularis*. Psychrotrophic taxa were overrepresented in Cluster 2, which comprises most samples collected in supermarkets. Some of these taxa are possibly involved in food and dairy spoilage (Innocente et al., 2009; Stellato et al., 2015). Moreover, several of the species overrepresented in Cluster 2 are potential pathogens. For example, *Plesiomonas* genus includes species that might be associated to foodborne disease (Janda et al., 2016). *Enterobacter hormaechei* is a pathogenic *Enterobacter* that has been previously isolated in cheese (Pangallo et al., 2014). In general, it has been observed that Cluster 2 is enriched in bacteria usually associated to lower quality compared to Cluster 1. Incidentally, Cluster 2 is the one containing the higher proportion of products marketed on mass distribution circuit, while in Cluster 1 the majority of products are marketed locally. One possible explanation of our findings is that products marketed in the mass distribution circuit might experience longer exposures to suboptimal temperatures, as well as longer production-to-consumption times, both potentially resulting in a relative increase of psychrotolerant, food spoilage-related organisms. It should be noted that, during processing of PDO buffalo Mozzarella, according to the procedural guidelines the milk must be delivered to the dairy within the sixteenth hour from the milking, and transformed into Mozzarella within the sixtieth hour from the first milking (Gobbetti et al., 2018). Thus, it is possible that the PDO Mozzarella cheese, also if locally sold, is produced using milk that has undergone more or less prolonged refrigeration. This may be the reason why three BDN samples are included in Cluster 2.

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

5 References

Addeo, F., Pizzano, R., Nicolai, M.A., Caira, S., Chianese, L., 2009. Fast isoelectric focusing and antipeptide antibodies for detecting bovine casein in adulterated water buffalo milk and derived Mozzarella cheese. J. Agric. Food Chem. 57, 10063–10066. doi:10.1021/jf9020009

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106. doi:10.1186/gb-2010-11-10-r106

Andreani, N.A., Martino, M.E., Fasolato, L., Carraro, L., Montemurro, F., Mioni, R., Bordin, P., Cardazzo, B., 2014. Tracking the blue: A MLST approach to characterise the *Pseudomonas fluorescens* group. Food Microbiol. 39, 116–126. doi:10.1016/j.fm.2013.11.012

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336. doi:10.1038/nmeth.f.303

Chao, A., 1984. Non-parametric estimation of the classes in a population. Scand. J. Stat. 11, 265–270. doi:10.2307/4615964

Chen, G., Chen, C., Lei, Z., 2017. Meta-omics insights in the microbial community profiling and functional characterization of fermented foods. Trends Food Sci. Technol. 65, 23–31. doi:10.1016/j.tifs.2017.05.002

Chen, L., Daniel, R.M., Coolbear, T., 2003. Detection and impact of protease and lipase activities in milk and milk powders. Int. Dairy J. 13, 255–275. doi:10.1016/S0958-6946(02)00171-1

Chevallier, S., Bertrand, D., Kohler, A., Courcoux, P., 2006. Application of PLS-DA in multivariate image analysis. J. Chemom. 20, 221–229. doi:10.1002/cem.994

Coppola, S., Blaiotta, G., Ercolini, D., Moschetti, G., 2001. Molecular evaluation of microbial diversity occurring in different types of Mozzarella cheese. J. Appl. Microbiol. 90, 414–420. doi:10.1046/j.1365-2672.2001.01262.x

de Candia, S., De Angelis, M., Dunlea, E., Minervini, F., McSweeney, P.L.H., Faccia, M., Gobbetti, M., 2007. Molecular identification and typing of natural whey starter cultures and microbiological and compositional properties of related traditional Mozzarella cheeses. Int. J. Food Microbiol. 119, 182–191. doi:10.1016/j.ijfoodmicro.2007.07.062

De Filippis, F., La Storia, A., Stellato, G., Gatti, M., Ercolini, D., 2014. A selected core microbiome drives the early stages of three popular Italian cheese manufactures. PLoS One 9, e89680. doi:10.1371/journal.pone.0089680

Edgar, R.C., 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat. Methods 10, 996–998. doi:10.1038/nmeth.2604

Ercolini, D., De Filippis, F., La Storia, A., Iacono, M., 2012. "Remake" by high-throughput sequencing of the microbiota involved in the production of water buffalo Mozzarella cheese. Appl. Environ. Microbiol. 78, 8142–8145. doi:10.1128/AEM.02218-12

Gobbetti, M., Neviani, E., Fox, P.F., 2018. The most traditional and popular Italian cheeses, in: Gobbetti, M., Neviani, E., Fox, P.F. (Eds.), The Cheeses of Italy: Science and Technology. Springer International Publishing, Cham, Switzerland, pp. 99–274.

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

Gorrasi, G., Bugatti, V., Tammaro, L., Vertuccio, L., Vigliotta, G., Vittoria, V., 2016. Active coating for storage of Mozzarella cheese packaged under thermal abuse. Food Control 64, 10–16. doi:10.1016/j.foodcont.2015.12.002

Guidone, A., Zotta, T., Matera, A., Ricciardi, A., De Filippis, F., Ercolini, D., Parente, E., 2016. The microbiota of high-moisture Mozzarella cheese produced with different acidification methods. Int. J. Food Microbiol. 216, 9–17. doi:10.1016/j.ijfoodmicro.2015.09.002

Hurley, I.P., Coleman, R.C., Ireland, H.E., Williams, J.H.H., 2006. Use of sandwich IgG ELISA for the detection and quantification of adulteration of milk and soft cheese. Int. Dairy J. 16, 805–812. doi:10.1016/j.idairyj.2005.07.009

Innocente, N., Marino, M., Marchesini, G., Biasutti, M., 2009. Presence of biogenic amines in a traditional salted Italian cheese. Int. J. Dairy Technol. 62, 154–160. doi:10.1111/j.1471-0307.2009.00479.x

Janda, M.J., Abbott, S.L., McIver, C.J., 2016. *Plesiomonas shigelloides* revisited. Clin. Microbiol. Rev. 29, 349–374. doi:10.1128/CMR.00103-15

Jans, C., Meile, L., Lacroix, C., Stevens, M.J.A., 2014. Genomics, evolution, and molecular epidemiology of the *Streptococcus bovis*/*Streptococcus equinus* complex (SBSEC). Infect. Genet. Evol. 33, 419–436. doi:10.1016/j.meegid.2014.09.017

Lopparelli, R.M., Cardazzo, B., Balzan, S., Giaccone, V., Novelli, E., 2007. Real-time TaqMan polymerase chain reaction detection and quantification of cow DNA in pure water buffalo

Mozzarella cheese: Method validation and its application on commercial samples. J. Agric. Food Chem. 55, 3429–3434. doi:10.1021/jf0637271

Ma, X., James, B., Balaban, M.O., Zhang, L., Emanuelsson-Patterson, E.A.C., 2013. Quantifying blistering and browning properties of Mozzarella cheese. Part I: Cheese made with different starter cultures. Food Res. Int. 54, 912–916. doi:10.1016/j.foodres.2013.06.007

Marino, M., Innocente, N., Maifreni, M., Mounier, J., Cobo-Díaz, J.F., Coton, E., Carraro, L., Cardazzo, B., 2017. Diversity within Italian cheesemaking brine-associated bacterial communities evidenced by massive parallel 16S rRNA gene tag sequencing. Front. Microbiol. 8, 2119. doi:10.3389/fmicb.2017.02119

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., Desantis, T.Z., Probst, A., Andersen, G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 6, 610–618. doi:10.1038/ismej.2011.139

Palmer, J.S., Flint, S.H., Schmid, J., Brooks, J.D., 2010. The role of surface charge and hydrophobicity in the attachment of *Anoxybacillus flavithermus* isolated from milk powder. J. Ind. Microbiol. Biotechnol. 37, 1111–1119. doi:10.1007/s10295-010-0758-x

Pangallo, D., Šaková, N., Koreňová, J., Puškárová, A., Kraková, L., Valík, L., Kuchta, T., 2014. Microbial diversity and dynamics during the production of May Bryndza cheese. Int. J. Food Microbiol. 170, 38–43. doi:10.1016/j.ijfoodmicro.2013.10.015

Paulson, J.N., Colin Stine, O., Bravo, H.C., Pop, M., 2013. Differential abundance analysis for microbial marker-gene surveys. Nat. Methods 10, 1200–1202. doi:10.1038/nmeth.2658

Pisano, M.B., Scano, P., Murgia, A., Cosentino, S., Caboni, P., 2016. Metabolomics and microbiological profile of Italian Mozzarella cheese produced with buffalo and cow milk. Food Chem. 192, 618–624. doi:10.1016/j.foodchem.2015.07.061

Quigley, L., O'Sullivan, O., Stanton, C., Beresford, T.P., Ross, R.P., Fitzgerald, G.F., Cotter, P.D., 2013. The complex microbiota of raw milk. FEMS Microbiol. Rev. 37, 664–698. doi:10.1111/1574-6976.12030

Quigley, L., Sullivan, D.J.O., Daly, D., Sullivan, O.O., Fitzgerald, G.F., Mcsweeney, P.L.H., Giblin, L., Sheehan, J.J., Cotter, D., 2016. *Thermus* and the pink discoloration defect in cheese. mSystems 1, e00023-16. doi:10.1128/mSystems.00023-16

R Core Team, 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

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

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

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

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

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

Figure captions

Figure 1. Alpha diversity measured as estimated number of OTUs in five different types of mozzarella cheese. Pairwise difference between groups was assessed using Wilcoxon test. Box-plots labeled with different letters are significantly different from each other.
BDN: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at mass retailers; **BNNG**: Buffalo mozzarella without certification, acidified with Natural Whey Culture and collected at mass retailers; **BNCG**: Buffalo mozzarella without certification, acidified with commercial starters and collected at mass retailers; **CC**: Cow mozzarella acidified with commercial starters and collected at mass retailers.

Figure 2. Abundance of bacterial families represented by at least 0.1% of reads in at least on sample.
BDN: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at mass retailers; **BNNG**: Buffalo mozzarella without certification, acidified with Natural Whey Culture and collected at mass retailers; **BNCG**: Buffalo mozzarella without certification, acidified with commercial starters and collected at mass retailers; **CC**: Cow mozzarella acidified with commercial starters and collected at mass retailers.

Figure 3. Clustering of samples based on the 50 most represented species.
BDN: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at mass retailers; **BNNG**: Buffalo mozzarella without certification, acidified with Natural Whey Culture and collected at mass retailers; **BNCG**: Buffalo mozzarella without certification, acidified with commercial starters and collected at mass retailers; **CC**: Cow mozzarella acidified with commercial starters and collected at mass retailers.

Figure 4. Clustering of samples obtained by the use of Natural Whey Culture.
BDN: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at

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

Figure 1

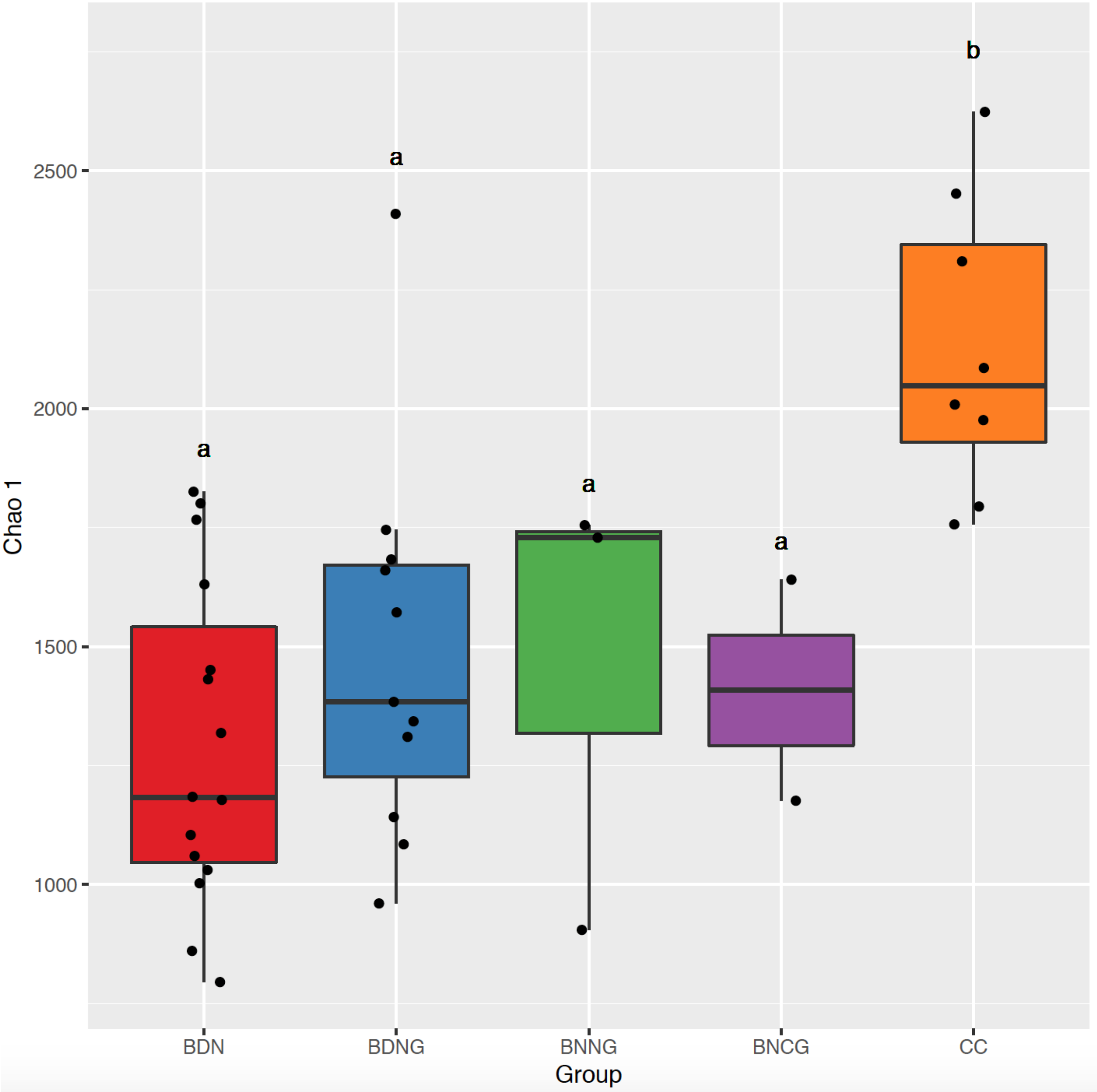
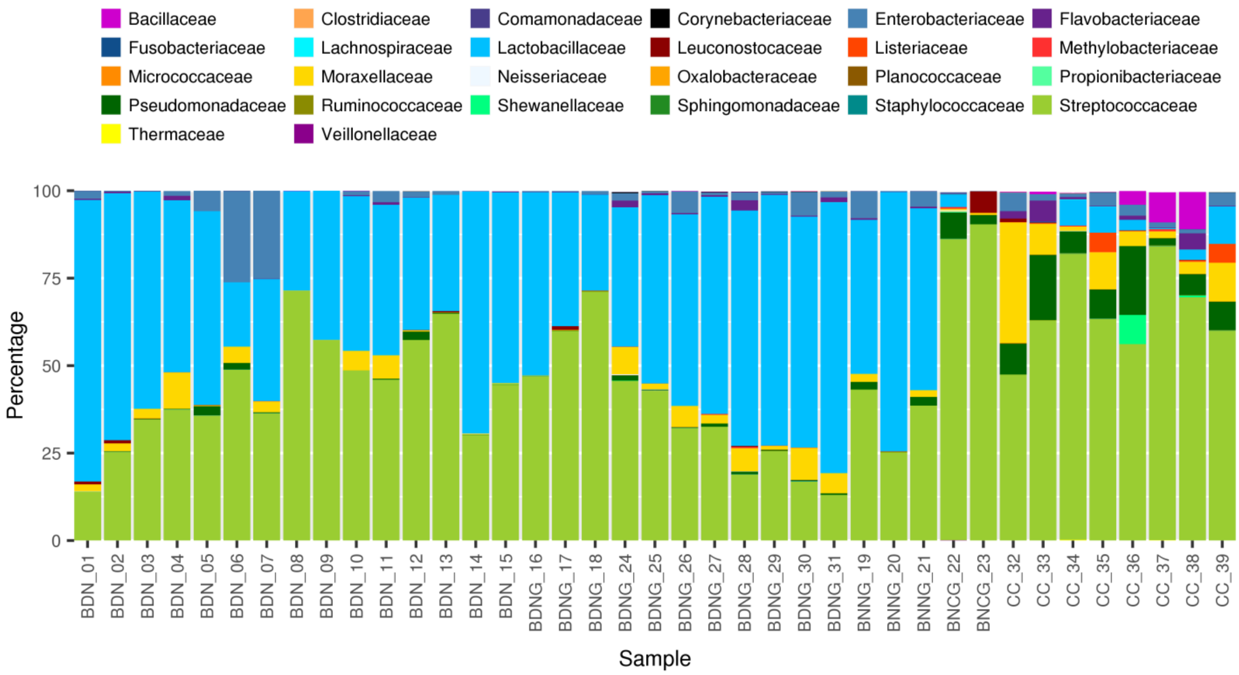


Figure 2



Bacillaceae
Clostridiaceae
Comamonadaceae
Corynebacteriaceae
Enterobacteriaceae
Flavobacteriaceae
Fusobacteriaceae
Lachnospiraceae
Lactobacillaceae
Leuconostocaceae
Listeriaceae
Methylobacteriaceae
Micrococcaceae
Moraxellaceae
Neisseriaceae
Oxalobacteraceae
Planococcaceae
Propionibacteriaceae
Pseudomonadaceae
Ruminococcaceae
Shewanellaceae
Sphingomonadaceae
Staphylococcaceae
Streptococcaceae
Thermaceae
Veillonellaceae
Percentage
100
75
50
25
0
BDN_01
BDN_02
BDN_03
BDN_04
BDN_05
BDN_06
BDN_07
BDN_08
BDN_09
BDN_10
BDN_11
BDN_12
BDN_13
BDN_14
BDN_15
BDNG_16
BDNG_17
BDNG_18
BDNG_24
BDNG_25
BDNG_26
BDNG_27
BDNG_28
BDNG_29
BDNG_30
BDNG_31
BNNG_19
BNNG_20
BNNG_21
BNCG_22
BNCG_23
CC_32
CC_33
CC_34
CC_35
CC_36
CC_37
CC_38
CC_39
Sample

Figure 3

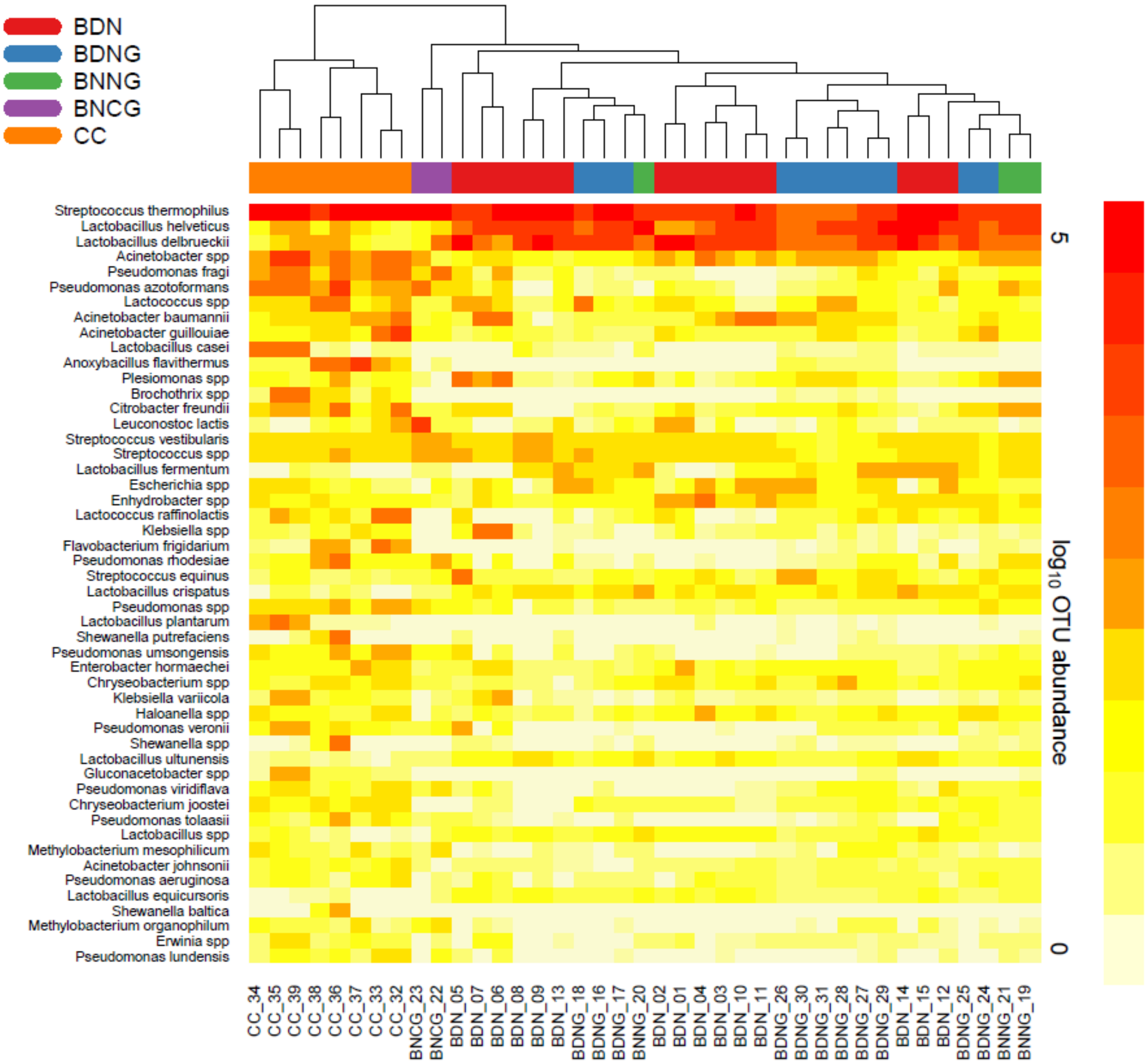
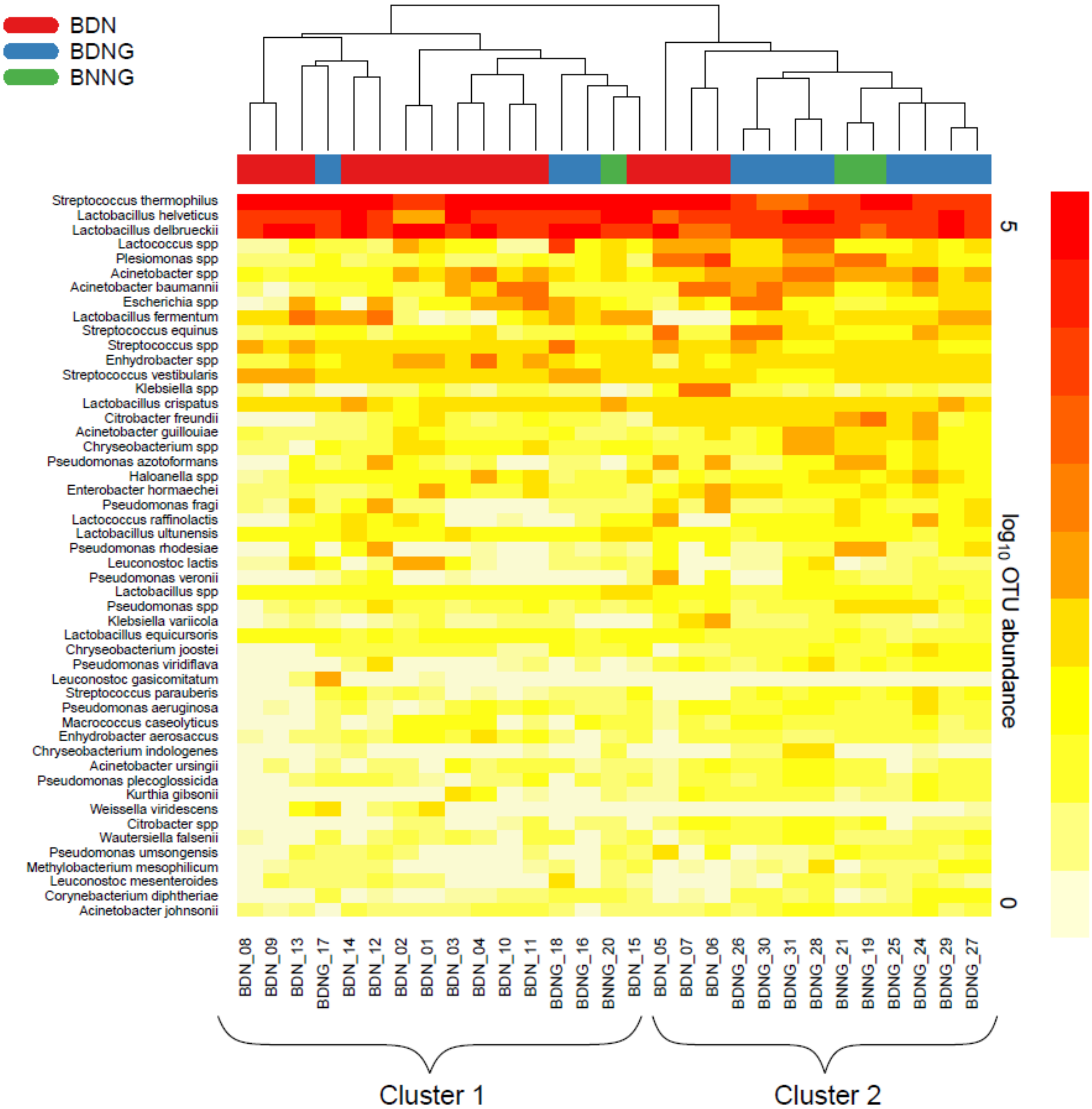

BDN
BDNG
BNNG
BNCG
CC
Streptococcus thermophilus
Lactobacillus helveticus
Lactobacillus delbrueckii
Acinetobacter spp
Pseudomonas fragi
Pseudomonas azotoformans
Lactococcus spp
Acinetobacter baumannii
Acinetobacter guillouiae
Lactobacillus casei
Anoxybacillus flavithermus
Plesiomonas spp
Brochothrix spp
Citrobacter freundii
Leuconostoc lactis
Streptococcus vestibularis
Streptococcus spp
Lactobacillus fermentum
Escherichia spp
Enhydrobacter spp
Lactococcus raffinolactis
Klebsiella spp
Flavobacterium frigidarium
Pseudomonas rhodesiae
Streptococcus equinus
Lactobacillus crispatus
Pseudomonas spp
Lactobacillus plantarum
Shewanella putrefaciens
Pseudomonas umsongensis
Enterobacter hormaechei
Chryseobacterium spp
Klebsiella variicola
Haloanella spp
Pseudomonas veronii
Shewanella spp
Lactobacillus ultunensis
Gluconacetobacter spp
Pseudomonas viridiflava
Chryseobacterium joostei
Pseudomonas tolaasii
Lactobacillus spp
Methylobacterium mesophilicum
Acinetobacter johnsonii
Pseudomonas aeruginosa
Lactobacillus equicursoris
Shewanella baltica
Methylobacterium organophilum
Erwinia spp
Pseudomonas lundensis
CC_34
CC_35
CC_39
CC_38
CC_36
CC_37
CC_33
CC_32
BNCG_23
BNCG_22
BDN_05
BDN_07
BDN_06
BDN_08
BDN_09
BDN_13
BDNG_18
BDNG_16
BDNG_17
BNNG_20
BDN_02
BDN_01
BDN_04
BDN_03
BDN_10
BDN_11
BDNG_26
BDNG_30
BDNG_31
BDNG_28
BDNG_27
BDNG_29
BDN_14
BDN_15
BDN_12
BDNG_25
BDNG_24
BNNG_21
BNNG_19
5
0
log10 OTU abundance

Figure 4


BDN
BDNG
BNNG
Streptococcus thermophilus
Lactobacillus helveticus
Lactobacillus delbrueckii
Lactococcus spp
Plesiomonas spp
Acinetobacter spp
Acinetobacter baumannii
Escherichia spp
Lactobacillus fermentum
Streptococcus equinus
Streptococcus spp
Enhydrobacter spp
Streptococcus vestibularis
Klebsiella spp
Lactobacillus crispatus
Citrobacter freundii
Acinetobacter guillouiae
Chryseobacterium spp
Pseudomonas azotoformans
Haloanella spp
Enterobacter hormaechei
Pseudomonas fragi
Lactococcus raffinolactis
Lactobacillus ultunensis
Pseudomonas rhodesiae
Leuconostoc lactis
Pseudomonas veronii
Lactobacillus spp
Pseudomonas spp
Klebsiella variicola
Lactobacillus equicursoris
Chryseobacterium joostei
Pseudomonas viridiflava
Leuconostoc gasicomitatum
Streptococcus parauberis
Pseudomonas aeruginosa
Macrococcus caseolyticus
Enhydrobacter aerosaccus
Chryseobacterium indologenes
Acinetobacter ursingii
Pseudomonas plecoglossicida
Kurthia gibsonii
Weissella viridescens
Citrobacter spp
Wautersiella falsenii
Pseudomonas umsongensis
Methylobacterium mesophilicum
Leuconostoc mesenteroides
Corynebacterium diphtheriae
Acinetobacter johnsonii
5
0
log10 OTU abundance
BDN_08
BDN_09
BDN_13
BDNG_17
BDN_14
BDN_12
BDN_02
BDN_01
BDN_03
BDN_04
BDN_10
BDN_11
BDNG_18
BDNG_16
BNNG_20
BDN_15
BDN_05
BDN_07
BDN_06
BDNG_26
BDNG_30
BDNG_31
BDNG_28
BNNG_21
BNNG_19
BDNG_25
BDNG_24
BDNG_29
BDNG_27
Cluster 1
Cluster 2

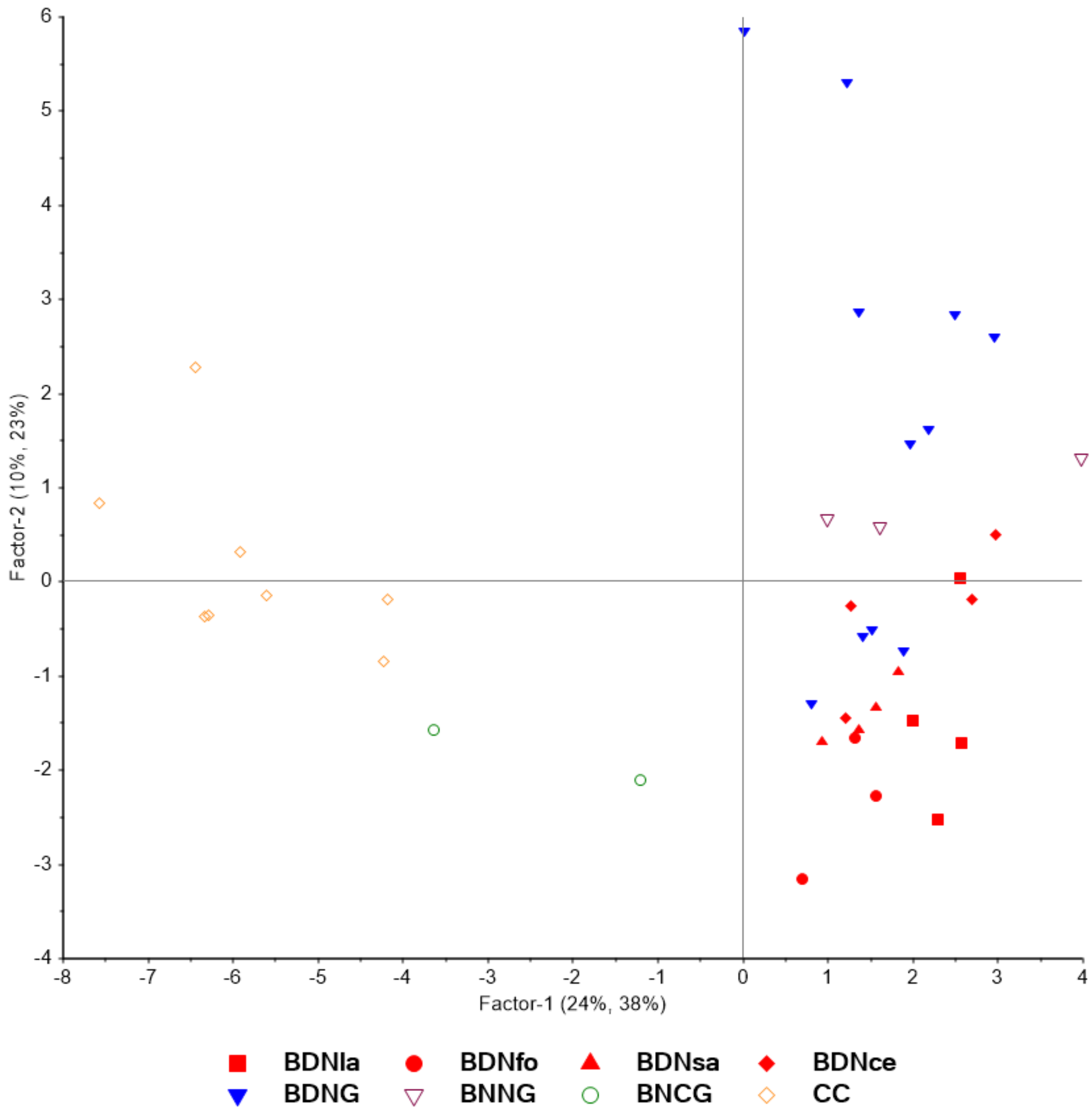


Figure S1: Score-plot of Mozzarella samples from a PLS-DA model of classification based on their OUTs profile. BDNla, BDNfo, BDNsa and BDNce were samples coming respectively from the provinces of Latina, Foggia, Salerno, and Caserta.

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

[§]Legend of names prefixes**:** **BDN**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture; **BDNG**: Buffalo mozzarella with PDO certification, acidified with Natural Whey Culture and collected at mass retailers; **BNNG**: Buffalo mozzarella without certification, acidified with Natural Whey Culture and collected at mass retailers; **BNCG**: Buffalo mozzarella without certification, acidified with commercial starters and collected at mass retailers; **CC**: Cow mozzarella acidified with commercial starters and collected at mass retailers; *, BM=buffalo Mozzarella, CM=cow Mozzarella; [#], presence of PDO certification (PDO) or not (none); [¥], local (L) or mass (M) retailer; [†]acidification with NWC=natural whey culture, or CS=commercial starter. [^], According to Chao (1984)

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

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

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

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

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