Journal Pre-proof

The microbiota of dairy milk: a review

Eugenio Parente, Annamaria Ricciardi, Teresa Zotta

PII: S0958-6946(20)30084-4

DOI: https://doi.org/10.1016/j.idairyj.2020.104714

Reference: INDA 104714

To appear in: *International Dairy Journal*

Received Date: 20 December 2019

Revised Date: 16 March 2020

Accepted Date: 16 March 2020

Please cite this article as: Parente, E., Ricciardi, A., Zotta, T., The microbiota of dairy milk: a review, *International Dairy Journal*, https://doi.org/10.1016/j.idairyj.2020.104714.

The microbiota of dairy milk: a review

Eugenio Parente*, Annamaria Ricciardi, Teresa Zotta

Scuola di Scienze Agrarie, Forestali, Alimentari ed Ambientali, Università degli Studi della Basilicata, Potenza, Italy

*Corresponding author. Tel.: +39 0971205561

E-mail address: eugenio.parente@unibas.it (E. Parente)

ABSTRACT

The number of studies using high throughput sequencing (HTS) methods for the characterisation of microbial communities of milk and dairy products has increased markedly in the last decade. Besides confirming what was previously known from low sensitivity and throughput cultivation based and cultivation independent techniques, HTS studies have provided deeper insights into the structure and function of microbial communities. While the comparison of raw data from different studies is still difficult due the lack of standard operating procedures, the availability of well-structured databases with raw and processed sequences has boosted our ability to get quantitative insights into the factors that shape the microbial communities of milk and dairy products. Here we critically review metataxonomic and metagenomic studies on milk from cows and other dairy species, and discuss potential sources and dynamics of microbiota during storage, transport and processing in liquid pasteurised milk, using the FoodMicrobionet database to carry out meta-analyses.

42 **Table of contents**

43

59

1. Introduction

In the last ten years more than 165 papers in which the microbiota of dairy products has been characterised using high throughput sequencing (HTS) approaches have been published (Supplementary material Table S1). Milk and dairy products are therefore the foods for which most data on the structure and functions of microbial communities involved in animal health, safety, fermentation and spoilage are available (De Filippis, Parente, & Ercolini, 2017). Literature on the microbiome of the bovine udder and its role in the health and well-being of dairy cows (Derakhshani et al., 2018a; Rainard, 2017), on the microbiome of milk (Addis et al., 2016; Oikomonou et al., 2020; Quigley et al., 2013; Tilocca et al., 2020), and cheese (Afshari, Pillidge, Dias, Osborn, & Gill, 2018; Yeluri Jonnala, McSweeney, Sheehan, & Cotter, 2018) has been recently reviewed. However, the most recent reviews on milk microbiota focus on aspects related to the role of udder and milk microbiota on the health of dairy animals (Addis et al., 2016; Oikomonou et al., 2020) and on methods (Tilocca et al., 2020), rather than on aspects relevant to the safety and quality of milk for the production of dairy products. While the teat surface and interior are certainly a source of bacteria which may play a role in safety and quality of dairy products, other factors, including further contamination from milking, farm and processing plant equipment environments, growth during storage and destruction of microorganisms by pasteurisation will give a significant contribution in shaping the microbiota of dairy and cheese milk.

In addition, the availability of raw or processed sequencing data in large databases (QIITA, https://qiita.ucsd.edu/; MGnify, https://www.ebi.ac.uk/metagenomics/; FoodMicrobionet, Parente et al., 2016a; Parente, De Filippis, Ercolini, Ricciardi, & Zotta, 2019) offers unprecedented opportunities for exploring the food microbiome in meta-studies. Recently,

we have released version 3.1 of FoodMicrobionet (Parente et al., 2019), a database of studies on the bacterial microbiota of foods, which is easily accessible through an interactive app. The database is steadily growing and its latest version includes, at the time of writing this article, 42 studies on the bacterial microbiome of dairy products (see supplementary material and supplementary data).

The objective of this work is to critically review the recent knowledge on methods used in the characterisation of the microbiota of milk from cows and other dairy species, on its potential sources and on its dynamics during storage, transport and processing in liquid pasteurised milk, and make use of the data available in FoodMicrobionet to carry out meta-analyses on studies on the dairy milk microbiota.

2. Milk and dairy products illuminated: the evolution of methods.

All three HTS approaches (amplicon sequencing, shotgun metagenomics and metatranscriptomics; Yeluri Jonnala et al., 2018) have been used, alone or in combination, to characterise the microbiome of milk. Methods for nucleic acid extraction, wet laboratory stages, sequencing and bioinformatic analyses vary greatly among studies (Supplementary Table 1), and current approaches have been outlined in a recent review (Tilocca et al., 2020). We will briefly discuss some aspects which shape our ability to understand dairy microbiota by nucleic acid targeted omics approaches.

2.1. Amplicon sequencing

Amplicon sequencing is by far the most used approach for metataxonomic studies. For bacteria, the 16S RNA gene has been most frequently targeted, but a few studies have used 16S RNA as target (Supplementary material Table S1), to focus on the "active" fraction of the microbiota. The variable regions targeted also vary, with V1–V3 being most frequently used before the demise of the Roche 454 platforms, and V3–V4 and V4 being most frequently used with the Illumina platforms. As to fungi, Internal Transcribed Spacers ITS1 and ITS2 have been most frequently used as a target, while the use of regions of 18S and 28S is far less common. Due to the relatively short length of the regions targeted, the taxonomic resolution is often limited to the genus or, less frequently, to the species level, depending on the length and quality of sequences. However, most recently the use of single molecule sequencing (Jin et al., 2018; Li et al., 2017; Mo et al., 2019; Yang et al., 2019; Yu et al., 2018), the ability to detect oligotypes (Kamimura, De Filippis, Sant'Ana, & Ercolini, 2019) and Amplicon Sequence Variants (ASV; Callahan et al., 2016), or the availability of optimised databases (Meola et al., 2019) has been claimed to increase taxonomic resolution.

Resolution at the species level or below is critical in the analysis of the microbiota of dairy products, because species belonging to the same genus may have a very different significance for the quality and safety of dairy products. For example, several genera, like *Streptococcus, Staphylococcus and Corynebacterium*, include both pathogenic and starter and non-starter microorganisms, while others, like *Lactobacillus* include species that are either starter or non-starter bacteria. The use of the methods with the highest taxonomic resolution should therefore be encouraged.

A few other coding or non-coding regions of the genome of selected bacteria have been used as a target to study the micro-diversity of selected species, with a variable degree

of success in terms of accuracy and sensitivity: *Streptococcus thermophilus* (*lacSZ*, De Filippis, La Storia, Stellato, Gatti, & Ercolini, 2014; *serB*, Parente et al., 2016b; Ricciardi et al., 2016), amine producing lactic acid bacteria (*tdc* and *hdc*, O'Sullivan et al., 2015), *Lactococcus lactis* (*purR, epsD*, Frantzen, Kleppen, & Holo, 2018), genus *Lactobacillus* (*groEL*, Jiang et al., 2019; ITS, Milani et al., 2018), *L. casei* group (*spxB*, Levante et al., 2017), genus *Bifidobacterium* (Milani et al., 2019), and members of the *Bacillus cereus* group (*panC, glpT*, Porcellato, Aspholm, Skeie, & Mellegård, 2019). Proteolysis related genes of LAB (*prtP, pepN, pepX, bcaT*) have also been used as target for metatranscriptomic studies (Pangallo et al., 2019). The usefulness of protein coding genes might be limited to selected species, but it is a cheaper alternative to shotgun whole genome sequencing, which might detect only the dominating strains in microbial communities.

To date, no Standard Operating Procedures (SOPs) for AT studies in foods exist. A discussion on the best approach for AT studies for milk and dairy foods is beyond the scope of this review, and the factors affecting the results have been reviewed recently (Pollock, Glendinning, Wisedchanwet, & Watson, 2018). However, it is quite clear that the development of SOPs, and the use of negative controls and mock communities or internal standards would be highly desirable and should be requested by editors and reviewers of scientific journals.

2.2. Shotgun approaches

To date, shotgun metagenomic studies of dairy products are comparatively rare (only 11% of studies listed in Supplementary material Table S1) and meta-transcriptomic studies are even less frequent (only 4%). Due to the decreasing costs in sequencing and the

progress in the development of bioinformatic pipelines for taxonomic annotation and genome reconstruction, as well as the staggering amount of information they can provide (Tilocca et al., 2020; Yeluri Jonnala et al., 2018) they are likely to become more frequent. Pipelines for metagenomic annotation and data visualisation have been reviewed recently (Breitwieser, Lu, & Salzberg, 2017; Quince, Walker, Simpson, Loman, & Segata, 2017; Sudarikov, Tyakht, & Alexeev, 2017), and the choice of the pipeline has been shown to affect the results, especially in low diversity samples which are typical of cheese and fermented milks (Walsh et al., 2018). An alternative to the use of shotgun metagenomics is the inference of metagenomes using bioinformatic tools such as PICRUSt (and its most recent iteration, PICRUSt2; Douglas et al., 2019). Although this tool has performed relatively well in benchmarking (Douglas et al., 2019) it has been used relatively rarely for microbial communities of milk and dairy products (Cremonesi et al., 2018; Li et al., 2018; Ramezani, Hosseini, Ferrocino, Amoozegar, & Cocolin, 2017; Stellato, De Filippis, La Storia, & Ercolini, 2015; Yang et al., 2019).

2.3. A question of life and death

A further issue related to the experimental approach is the inability of methods targeting DNA to distinguish active/viable members of microbial communities from those which are dead/inactive and contribute little or nothing to fermentation or spoilage. Studies using both DNA and RNA as a target are relatively rare (see Supplementary material Table S1; De Filippis, Genovese, Ferranti, Gilbert, & Ercolini, 2016; Kastman et al., 2016; Sattin et al., 2016b) and the use of dyes which prevent the PCR amplification of DNA from dead cells (or more properly, cells with a damaged membrane), such as PMA (propidium monoazide,

Emerson et al., 2017), is only marginally more frequent (Erkus et al., 2016; Kable, Srisengfa, Xue, Coates, & Marco, 2019; Mo et al., 2019; Porcellato & Skeie, 2016). Extensive benchmarking it still needed to rule out biases due to differential ability of PMA to penetrate cell membranes (Emerson et al., 2017). At any rate, whenever both the "active" and "inactive" fraction of the microbiota have been targeted, significant differences have been found between the two, usually with lower diversity in the "active" microbiota.

2.4. Of experimental design (or lack thereof)

Regrettably, the overwhelming majority of the studies listed in Supplementary material Table S1 are descriptive in nature, and even when inferential methods are used, their effectiveness in detecting significant differences (because of high natural variability, and potentially high type I and/or type II errors) is dubious. In fact, only in a very few cases experimental designs have been used (De Filippis et al., 2016; Doyle, Gleeson, O'Toole, & Cotter, 2017b; Ganda et al., 2016, 2017; Guzzon et al., 2017; Porcellato & Skeie, 2016), and for most studies the approach is quasi-experimental in nature, with insufficient randomisation, blocking and control of confounding factors. The issue of sampling effort is also critical: the range for the number of samples analysed in studies shown in Supplementary material Table S1 is 1 to 1674, but 50% of the studies have used 24 samples or less. Because of the very high variability of the microbiota of raw milk, due to seasonal, geographical and technological factors (Kable et al., 2016; Skeie, Håland, Thorsen, Narvhus, & Porcellato, 2019) one really wonders if, especially for raw milk fermented milks or cheeses produced in artisanal plants, low (<50) sample numbers and low numbers of sampling locations (farms, cheesemaking plants) are adequate to cover the expected diversity, and,

even in larger studies, utmost care should be dedicated to the design of the experiments and to the analysis of the data using appropriate inferential methods.

The issue of microbial interactions in dairy ecosystems is of great interest for both scientific and practical reasons, but it has been addressed only infrequently (Frétin et al., 2018; Murugesan et al., 2018; Parente et al., 2016a; Parente, Zotta, Faust, De Filippis, & Ercolini, 2018; Wolfe, Button, Santarelli, & Dutton, 2014). Detecting true interactions among species presents several challenges (Layeghifard, Hwang, & Guttman, 2017). Unfortunately, most studies are cross-sectional in nature, and, even when they are longitudinal, the number and distribution of sampling times is insufficient for model-based methods for detection of microbial interactions (Faust & Raes, 2012). More research is definitely needed in this area and combinations of culture independent and dependent approaches (Wolfe et al., 2014) are needed to validate the nature of the microbial interactions and evaluate their significance for the quality of dairy foods.

3. The microbiota of milk: from the teat to the carton

A large amount of data on the microbiota of raw or pasteurised milk composition or milk contact surfaces (teat and udder surface, tanks and silos at the dairy farm or at the processing plant, etc.) are available for milk from practically all dairy animals (mostly cow, but also ewes, goats, water buffaloes, yaks, camels), either as a part of studies specifically focusing on milk quality or as a part of studies on cheese microbiota (see Supplementary material Table S1). Milk microbiota is undoubtedly complex and highly variable and in most studies the sampling effort is limited or the approach is merely descriptive, thus obscuring causal relationships. However, a few large, designed or quasi-experimental studies

addressing one of more aspects (effect of cows' health, feeding, farming, breed, season, geographical source of milk, effect of contamination during the production and distribution chain, effect of storage temperature) are available, and combination of data from different studies in meta-analyses may help in identifying a core microbiota or detecting wider geographical or temporal trends. In the following sections, we will review the composition of the microbiota of milk as it travels from the udder to the storage tank in processing plants and, finally to the carton of pasteurised milk.

3.1. Raw milk

3.1.1. Inside and outside the udder

The first sources of microorganisms in raw milk are, quite obviously, the udder and the teat surface (Derakhshani, Plaizier, De Buck, Barkema, & Khafipour, 2018b). The composition of the mammary microbiota in ruminants has been recently reviewed (Derakhshani et al., 2018b; Rainard, 2017), and the mechanisms that determine its composition and dynamics are outside the scope of this review. While it is still somewhat controversial if a microbiota of the healthy mammary gland exists or if it is the result of contamination during sampling (Derakhshani et al., 2018a; Rainard, 2017), it is clear that the teat canal and apex may be colonised by bacteria and that these bacteria may contribute to the homeostasis of this niche or cause infection of the mammary gland. In fact, most of the studies using milk from individual quarters or individual animals have focused on the effect of disease (mastitis, either clinical or subclinical, subclinical acidosis) on the microbiota of milk from cows (see below and Supplementary material Table S4 for a list of studies), while only a few studies are available on ewe (Castro et al., 2019; Esteban-Blanco et al., 2019),

goat (McInnis, Kalanetra, Mills, & Maga, 2015) or on water buffalo milk (Catozzi et al., 2017; Patel et al., 2016; Patel, Kunjadia, Koringa, Joshi, & Kunjadiya, 2019). This is justified by the economical and practical importance of mastitis, which is the most important disease in dairy animals in terms of both impact on milk production and quality and in terms of animal well-being (Ruegg, 2017).

A few studies have analysed the milk from individual healthy cows and the teat surface to investigate the sources of microorganisms, beneficial or not, and their potential effect on cows' health (Cremonesi et al., 2018; Falentin et al., 2016; Frétin et al., 2018). It is important to remember that sampling and disinfection may dramatically affect the composition of the microbiota of individual milk samples (Metzger et al., 2018a), that milk obtained aseptically or by abiding to hygienic practices has usually low counts (often less than 1×10^4 cfu mL^{-1}), and that contamination might significantly affect the results of AT studies for low count samples (Dahlberg et al., 2019). In addition, a high number of amplification cycles may be necessary for teat milk obtained aseptically (Metzger et al., 2018a) and success rate of amplification may be relatively low for milk obtained aseptically from healthy quarters. The results of some early studies on low counts milk which did not include negative control or proper treatment for removing contamination might be therefore slightly biased.

The teat interior and surface are among the most significant sources of microorganisms for individual milk samples, and microorganisms from these sources may persist during transport and transformation of milk. On the other hand, mastitis is likely to have a larger impact compared with external contamination from the teat interior and surface. In a carefully controlled experiment (Andrews, Neher, Weicht, & Barlow, 2019), intramammary infection was found to dramatically affect the composition of the microbiota

of teat cistern milk, which, compared with the milk of healthy animals, had a lower bacterial diversity, was more variable among different cows and was often enriched in pathogenic bacteria belonging to the same genus of those isolated by culturing from the affected quarters, supporting the hypothesis that mastitis is correlated with dysbiosis of the mammary gland. In addition, the microbiota of cistern milk and teat apex of infected quarters was more similar compared with healthy quarters, while the microbiota of cistern milk and teat apex in healthy animals was also more variable in time, suggesting that it was more affected by external contamination.

The effect of mastitis on the microbiota of the teat apex may be detectable even long after the demise of symptoms. Falentin et al. (2016) examined the microbiota of the teat apex (foremilk + teat apex swabs) for healthy Holstein cows from a single experimental farm. The cows had different previous histories of mastitis. The cows' teat canal microbiota were highly variable (even within the same cow) and teat microbiota from cows without and with a previous history of mastitis could be clearly differentiated, while microbiota of cows with an uncertain status tended to cluster with those of cows with a previous history of mastitis (cluster 1). Discrimination between the two clusters was due to a higher abundance of members of class *Bacilli* (with *Staphylococcus aureus* and *Staphylococcus equorum* as the most prevalent and abundant species) in cluster 1 and higher relative abundance of a diverse array of genera belonging to the phylum *Actinobacteria* (including *Bifidobacterium*), class *Clostridia*, phylum *Bacteroidetes*, including several genera associated with the gastrointestinal (GI) tract, in cluster 2. The origin of these microorganisms may be therefore the teat canal itself, in the case of mastitis agents like *Staph. aureus,* while the potential origin of bacteria associated with the GI tract is uncertain. The procedure used in this study included thorough washing and sanitation before sampling of the teat canal, and may have

299 reduced contamination from loosely attached bacterial cells originating from faeces or from
300 the environment, but it might not have prevented it completely.

301 Data on the microbiota of the cow teat skin surface in animals showing no signs of
302 clinical mastitis are available for three more studies (Doyle et al., 2017b; Falardeau, Keeney,
303 Trmčić, Kitts, & Wang, 2019; Frétin et al., 2018). Teat skin microbiota was highly diverse and
304 variable: a prevalence and abundance plot with data extracted from FoodMicrobionet is
305 shown in Supplementary material Fig. S1, and a table showing the top 50 taxa in terms of
306 prevalence and relative abundance is provided as Supplementary material Table S2. Frétin
307 et al. (2018) sampled the teat skin of cows belonging to two breeds (Holstein and
308 Montbeliarde) under two different farming regimes (extensive EXT, with cows feeding
309 exclusively on pasture, and semi-extensive, SEMI, with cows feeding on pasture and
310 concentrate) prior to evening milking (i.e., prior to washing). As a consequence, the results
311 were probably affected by both autochthonous species and by contaminants from faeces
312 and the farm/pasture environment. More than 300 operational taxonomic units (OTUs) and
313 98 genera were identified, including both *Actinobacteria* and *Clostridia* as abundant
314 members. Some OTUs (twelve, including members of the genera *Brevibacterium,*
315 *Lactococcus, Lactobacillus, Streptococcus, Staphylococcus, Macrococcus, Escherichia*)
316 persisted throughout the process, from teat skin to ripened cheese, while 201 were specific
317 to teat skin. The microbiota of teat skin was most affected by the grazing system and by the
318 season of sampling (July versus September).

319 Falardeau et al. (2019), in a large source tracking study, confirmed that teat skin
320 (sampled prior to washing and milking) had a high microbial diversity, which, however, was
321 comparable with that of teat milk and tank milk. *Clostridiales* were the most abundant
322 members of the microbiota (17–41%), but *Actinobacteria* (*Corynebacterium* and

323 *Brevibacterium*), *Bacteroidetes* (*Bacteroides* and *Alistipes*), and *Proteobacteria* (including
324 *Pseudomonas* and *Acinetobacter*) were all found in the subdominant microbiota.
325 The main difference between these two studies is the higher relative proportion of
326 *Clostridia* and *Bacteroidia* in Falardeau et al. (2019) and the higher proportion of *Bacilli* and
327 *Erysipelotrichia* in Frétin et al. (2018). The results of Falardeau et al. (2019) are similar to
328 those of Doyle et al. (2017b) who analysed, in a systematic study, the effect of farming
329 (indoor versus outdoor), and cleaning procedure on the microbiota of teat surface,
330 individual and bulk milk samples, and confirmed that teat swab microbiota is highly diverse
331 and significantly affected by both farming practices and cleaning.
332
333 *3.1.2. The microbiota of teat milk.*
334 The microbiota of samples obtained by milking individual quarters or individual
335 animals has been analysed in several studies, focusing on the effect of disease and/or
336 disease treatment, such as mastitis (Angelopoulou et al., 2019; Bhatt et al., 2012; Ganda et
337 al., 2016, 2017; Hoque et al., 2019; Kuehn et al., 2013; Metzger et al., 2018b; Oikonomou,
338 Machado, Santisteban, Schukken, & Bicalho, 2012; Oikonomou et al., 2014; Oultram, Ganda,
339 Boulding, Bicalho, & Oikonomou, 2017; Pang et al., 2018; Taponen et al., 2019; Vasquez et
340 al., 2019) and subclinical acidosis (Zhang, Huo, Zhu, & Mao, 2015), on sampling (Metzger et
341 al., 2018a) or on the tracking of sources of contaminations (Cremonesi et al., 2018; Dahlberg
342 et al., 2019; Doyle et al., 2017b; Falardeau et al., 2019; Metzger et al., 2018a). Comparing
343 these studies is difficult, because of differences in practically all the factors that are known
344 to affect the composition of microbiota (breed, health status, farming, bedding, feeding,
345 lactation stage, etc.), in sampling, in methods used for the analysis of the microbiota. In
346 particular, the composition of the microbiota of individual milk samples has been proven to

be strongly dependent on the sampling procedure used (Metzger et al., 2018a) and sampling procedures must be carefully documented to allow the interpretation of results from different studies.

However, several findings have been confirmed by multiple studies: (i) the microbiota of milk of healthy animals is highly diverse and variable; (ii) mastitis and other clinical and subclinical conditions strongly affect the composition and diversity of the microbiota; (iii) a large number of other factors, including breed, parity, farming systems, feeding, bedding, season of the year, and days in milking significantly affect the composition of milk microbiota.

To illustrate the variability of the composition of teat milk samples we have compared the results for milk from healthy cows from two studies [one illustrating the effect of breed and days in milking (Cremonesi et al., 2018), the other on contamination sources from farm to fork (Falardeau et al., 2019)] for which sequences are publicly available and which are included in FoodMicrobionet. The assembly of taxa for the two studies include 1184 genera (most of which with very low prevalence and abundance) belonging to 115 classes of 45 phyla. A bar plot of the relative abundance of the 20 most abundant and prevalent taxa is shown in Fig. 2, while a prevalence and abundance plot and the data on prevalence on the top 50 most abundant taxa are shown in Supplementary material Fig. S2 and Supplementary material Table S3, respectively.

In both studies the most prevalent and abundant taxa belong to phyla *Firmicutes*, *Actinobacteria* and *Bacteroidetes*, but large differences are evident both within and between studies. Cremonesi et al. (2018) compared Holstein Friesians with an Italian breed (Rendena, which shows lower prevalence of mastitis) from the same farm, from drying off, to colostrum stage and to late lactation. The number of samples in this study was relatively

low, but variation over time was observed for both breeds and beginning of lactation had a significant impact on the microbiota. The composition of the microbiota of Rendena cows was more stable, and significantly different from that of Holstein cows. *Streptococcus* was the most prevalent and abundant genus in both breeds and together with *Lactobacillus* was the only genus shared by all samples. On the other hand, Falardeau et al. (2019) found that *Actinobacteria* (with genera *Kocuria, Dermatococcus* and *Dietzia*) were by far the dominating phylum in teat milk, while *Firmicutes* (with *Lactococcus* and *Clostridium* XI as most abundant genera) and *Proteobacteria* (with *Enhydrobacter* and *Psychrobacter* as most abundant genera) were less abundant. Notably, the microbiota of teat milk for this study was quite different from that of teat skin (see above).

A seasonal effect on the composition of teat milk was also found by Metzger et al. (2018b) who monitored for over 150 days the composition of microbiota of teat milk obtained from healthy cow quarters, newly infected quarters and quarters with chronic inflammation or clinical mastitis. They found a strong effect of season of the year and of time of lactation, which resulted in increased richness from winter to summer in all cohorts, and in significant changes in the relative abundance of 20 OTUs (including *Fibrobacter*, *Corynebacterium*, *Arthrobacter*, *Bacteroidetes*), which they attributed to contamination from sand bedding and/or to physiological changes during lactation (*Bacteroidetes*). Interestingly, milk from quarter with chronic inflammation showed the greatest seasonal changes.

Much emphasis has been given to the comparison of the microbiota of individual milk samples from healthy cows and cows with subclinical or clinical mastitis, and in the latter, for culture positive and negative samples. Differences between healthy and diseased quarters are almost always significant (even for culture negative quarters; Kuehn et al.,

2013; Oikonomou et al., 2012) and analysis of the microbiota may contribute to the diagnosis in quarters with subclinical mastitis. Dominating bacteria in mastitic milk change quite substantially in different studies, depending on the number of samples tested and on the causative agents in individual cows (Supplementary material Table S4). Mastitis agents are usually abundant components of the microbiota in mastitic milk, but not necessarily the most abundant (Oikonomou et al., 2012) and their abundance may change over time (Ganda et al., 2016, 2017).

In several cases, association of two or more agents of mastitis are found (Angelopoulou et al., 2019; Bhatt et al., 2012; Oikonomou et al., 2012). Potential bacterial pathogens as *Streptococcus uberis* and *Staph. aureus* might also be found with high prevalence in milk from healthy quarters (Oikonomou et al., 2014) while, on the other hand, some mastitis agents like *Escherichia coli* or *Klebsiella* were never found in samples from healthy quarters: this has led to speculate that bacterial mastitis can be considered as a dysbiosis rather than a primary clinical infection. This hypothesis may also be supported by the fact that species which most contribute to the discrimination between mastitic and non mastitic milk are not necessarily mastitis pathogens, although they might have been occasionally associated with mastitis (Kuehn et al., 2013). A recent study (Angelopoulou et al., 2019) has confirmed the complex, polymicrobial nature of mastitis and showed that culture based approaches and AT metagenomics complement each other. In an attempt to evaluate if significant associations could be detected in this data set between known mastitis pathogens and other bacteria, we carried out inferences of microbial association networks as described in Parente et al. (2018) (Supplementary material Fig. S3). Only two modules were detected, one including *Escherichia/Shigella* and the other *Staphylococcus*, two genera which include species identified by culturing as potential agents of mastitis. The

two modules did not overlap but, due to the low number of samples, it is not clear if this reflect true differences in the microbiota of milk connected to infection from either *Escherichia* or *Staphylococcus*.

Another frequently observed consequence of inflammation due to mastitis is a decreased diversity in the microbiota (Andrews et al., 2019; Bonsaglia et al., 2017; Ganda et al., 2016, 2017; Kuehn et al., 2013; Metzger et al., 2018b; Taponen et al., 2019; Vasquez et al., 2019), although the lower number of OTUs in mastitic milk may be simply due to the compositional nature of AT data (when the relative abundance of one OTU increases the relative abundance of the others decreases and may fall below detection limits). Decrease in alpha diversity is a clear indication of a dysbiosis and in at least one study the largest reduction was associated with the largest decrease in milk production (Vasquez et al., 2019). However, distinguishing samples with subclinical mastitis from samples from healthy cows might not be straightforward using descriptive techniques when quasi-experimental designs are used (Pang et al., 2018), and in some non-severe cases it might be difficult to associate the mastitis condition with significant changes in abundance of bacterial families (Ganda et al., 2016; Vasquez et al., 2019). The potential impact of treatments used to control mastitis at dry off is of both scientific and practical importance. Ganda et al. (2016) compared two groups of Holstein cows that had been randomly allocated to a control group or to a group treated with 5 day intramammary treatment with Ceftiofur (a third generation cephalosporin). Antibiotic treatment did not affect the clinical or bacteriological cure rates, nor bacterial clearance or bacterial load, but did reduce over time the relative abundance of *Enterobacteriaceae* in quarters with *E. coli* mastitis. However, no clear effect was observed on quarters without a culture diagnosis.

In a subsequent study, the same group (Bonsaglia et al., 2017) showed that treatment with the same antibiotic at dry-off did not change the incidence of mastitis in the first 60 days post-partum for healthy cows, nor did it significantly affect the composition of the milk microbiota after 7 days. The authors concluded that this type of treatment does not cause dysbiosis but it does not have any therapeutic value for healthy cows. However, large individual variability may have increased type II error and prevented the detection of significant differences. Using the same antibiotic in a challenge study with *E. coli,* Ganda et al. (2017) observed that, independently of antibiotic use, normal microbiota re-established itself over a 216 h sampling time. This study is an excellent example of the dynamic changes of the composition of milk microbiota prior and during infection with a mastitis pathogen and shows how, after a disturbance, the microbiota may shortly revert to its initial status (or to a status that is not statistically different from the initial one).

Studies on the effect of mastitis on milk microbiota for other species are relatively rare and somewhat limited in scope. Catozzi et al. (2017) investigated the teat milk microbiota in 137 samples of water buffalo milk obtained from healthy quarters and from quarters with evidences of clinical and subclinical mastitis from 88 farms of limited area in Southern Italy. They identified a core microbiota of fifteen genera, including genera commonly found in cows' milk (see Fig. 3). Both subclinical and clinical mastitis significantly changed the composition of the microbiota, usually with a relative decrease in psychrotrophic microorganism (*Pseudomonas, Psychrobacter*), a decrease in *Actinobacteria* and *Firmicutes* and increase in *Proteobacteria* and *Bacteroidetes*, and clinical mastitis resulted in a decrease in alpha diversity. In general, the strongest differences were found between samples with low ($< 1 \times 10^5$ mL^{-1}) and high (0.5×10^6 to $> 1 \times 10^6$ mL^{-1}) somatic cell counts (SCC). Culture results for quarters with clinical mastitis confirmed the occurrence of

common mastitis pathogens, such as *Staph. aureus, Turicibacter pyogenes, Streptococcus agalactiae* (alone or in combination with *Staph. aureus*), *Pseudomonas aeruginosa* and some coagulase negative staphylococci. Similar results (reduced diversity, ability to discriminate samples from healthy animals from those with subclinical and clinical mastitis) have been found for water-buffalo milk in India (Patel et al., 2016, 2019).

Results on the effect of mastitis on ewes ' milk microbiota are somewhat contradictory. Esteban-Blanco et al. (2019) investigated the microbiota of teat milk obtained from a relatively low number (50) of healthy Assaf ewes from a single flock in Spain. Only 5 genera (*Staphylococcus, Lactobacillus, Corynebacterium, Streptococcus* and *Escherichia/Shigella*) were shared among all samples, and a high diversity was observed. Evidence of sub-clinical mastitis was associated to a reduced diversity. Using inference of microbial association networks, the authors identified two modules of ASVs, and observed that the relative abundance of the species in the two modules in samples without or with subclinical mastitis was different: this further confirms the hypothesis that subclinical mastitis causes global changes in the microbiota, which affect not only the potential causative agent but a number of other taxa.

Castro et al. (2019) analysed teat milk from 36 healthy Manchega ewes with or without a previous history of mastitis from two farms in Spain. They found significant differences between the microbiota in the two farms (with significant differences between the relative abundances *Staphylococcus, Paenibacillus* and *Geobacillus*) but, contrary to what had been reported for teat microbiota of cows by Falentin et al. (2016) did not find any significant difference due to history of mastitis: again, it is unclear if this reflects true lack of differences or is a consequence of high variability.

Other clinical or subclinical conditions may affect milk microbiota and susceptibility to mastitis. Zhang et al. (2015) using a crossover experiment analysed pooled teat milk samples from Holstein cows with or without an induced subclinical acidosis condition. Although a high concentrate (HC) diet, which resulted in subclinical acidosis, did not affect microbial diversity, significant differences were found between cows with or without clinical acidosis with the former having a significantly higher abundance of *Proteobacteria,* and lower abundance of *Armatimonadetes, Spirochaetes, Planctomycetes, Fibrobacteres, Chloroflexi, Tenericutes, Lentisphaerae, Synergistetes, Elusimicrobia, Cyanobacteria, Verrucomicrobia* and *Firmicutes*. The authors claimed that potential mastitis agents (including *Stenotrophomonas maltophila*, *Brevundimonas diminuta*, *Streptococcus parauberis* and *Enterococcus faecalis*) were significantly more abundant in the milk of cows fed the HC diet, which also resulted in an increase in the abundance of psychrotrophic organisms and this may support the idea that mastitis is related to dysbiosis. On the other hand, other mastitis agents, like *Str. agalactiae* were significantly more abundant in the milk of cows fed a low concentrate diet.

Finally, there are some limited data (Zhong, Xue, & Liu, 2018) that may support the idea that udder health status may be related to the microbiota of other body sites. In fact, diversity of the microbiota of rumen in cows with low ($< 2 \times 10^5$ mL^{-1}) or high ($> 1 \times 10^6$ mL^{-1}) SCC has been found to be significantly different, and although no evidence of separation of the composition of the microbiota of rumen in four groups of cows with different SCC was found by beta diversity analysis, significant differences were found in the relative abundance of a few taxa (phyla SR1, *Actinobacteria*, unclassified family *Clostridiales*, genus *Butyrivibrio*, *Proteobacteria* and family *Succinivibrionaceae*). However, the authors did not

present an evidence of a cause effect relationships nor analysed the composition of the microbiota of milk.

Overall, these data support the idea that clinical and sub-clinical conditions significantly affect the composition and diversity of teat milk microbiota, prior to any further contamination from environmental sources, and that these changes may result in dysbiosis, compared with the "normal" situation characterised by a highly diverse microbiota. The dysbiosis status may be more (Falentin et al., 2016) or less (Ganda et al., 2016) persistent, and more complex and controlled longitudinal studies are clearly needed to clarify how the homeostasis of the milk microbiota is maintained or recovered in different conditions.

Due to limited availability of data, to high variability, and to differences in methodologies, a direct comparison of the composition of microbiota of teat milk from different dairy species is difficult. The distribution of genera in teat milk obtained from cows, ewes or water buffaloes is shown in Fig. 3, while a NMDS (Non-metric MultiDimensional Scaling) plot is shown in Supplementary material Fig. S4. Due to the low number of studies and samples shown here it is difficult to generalise, but it is clear that several of the most abundant genera, although varying in abundance, appear in the milk of the 3 species. A shared core microbiome may exist for these three species, at least when results are aggregated at the genus level. This is confirmed by the partial overlap of the confidence ellipses of the samples from different species in Supplementary material Fig. S4.

3.1.3. Further down the line: the microbiota of bulk tank milk.

The vast majority of studies on milk microbiota focus on composite samples obtained from bulk tanks at the dairy farm, from tanker trucks or from silos at the dairy processing plant. Apart from disease, a large number of interrelated factors has been shown

536 to affect the composition of the microbiota for bulk milk: season of the year (Doyle,
537 Gleeson, O'Toole, & Cotter, 2017a; Doyle et al., 2017b; Kable et al., 2016, 2019; Li et al.,
538 2018; Porcellato et al., 2018; Zhang, Palmer, Teh, Biggs, & Flint, 2019), lactation stage (Doyle
539 et al., 2017a), type of farming (indoor/outdoor: Doyle et al., 2017a,b), geographic location
540 within a country/region (Kable et al., 2016, 2019; Porcellato et al., 2018; Skeie et al., 2019;
541 Zhang et al., 2019), processing environment (Kable et al., 2016), teat preparation (Doyle et
542 al., 2017b), storage conditions (Doyle et al., 2017a).

543 The issue of sources of contamination of bulk tank milk is of great practical
544 importance, since preventing and controlling contamination by selected pathogenic or
545 spoilage organisms may contribute to improve the safety and quality of raw milk and raw
546 milk products. Since not all studies fully document all potential sources of variation, it is
547 difficult to track unambiguously sources of contamination and to separate the effect of
548 contamination from that of storage (combination of time temperature), except for a few
549 large and structured studies (Doyle et al., 2017a,b; Falardeau et al., 2019; Kable et al., 2016;
550 Porcellato et al., 2018).

551 Using a source tracking approach, Doyle et al. (2017b) clearly identified teat surface
552 and faeces as two of the major sources of microorganisms in bulk tank milk and were able
553 to identify the contribution of other major sources of contamination (including grass for
554 cows grazed on pasture, bedding and silage for cows housed indoor). However, the effect of
555 housing was confounded with that of season and lactation stage since experiments were
556 carried out on the same herd. For sources of contamination which were common to both
557 housing regimes (faeces, teat) the composition of microbiota was affected more by the
558 housing than by the nature of the sample (i.e., samples from the outdoor regime tended to
559 cluster together in beta diversity analysis). In addition, the relative importance of a given

source of contamination changed for the two housing regimes (with faeces giving a higher relative contribution for bulk tank milk from cows housed indoors). The effect of teat treatment (which compared no treatment with a treatment including washing with water, disinfectant and thorough washing) clearly showed an interaction with housing (indoor versus outdoor) perhaps due to the different ability of main teat contaminants to adhere to teat surface or to survive to the treatment.

These findings were confirmed by a large recent study (Falardeau et al., 2019) in which the microbiota in both environmental (soil, faeces, pasture, hay, bedding, cow environment) and food samples (individual and pooled milk samples at both the dairy farm and at the cheesemaking plant) in an artisanal cheese making facility was analysed. The microbiota of teat milk was significantly different from the microbiota of bulk tank milk (pooled pre- and post-transport milk), possibly because of both contamination from equipment and growth at refrigeration temperature. In fact, several anaerobic *Firmicutes* and *Bacteroidetes* genera were the most abundant in the bulk tank milk, but not in the teat milk, and the authors hypothesised that their source was the milking machine environment, although this was not formally proven. At any rate, 78 out of 93 core OTUs present in milk environments were also present in the dairy farm environment, supporting the idea that the dairy farm is an important source of microorganisms in bulk tank milk. An even larger number of taxa (at the genus level or above) were shared by pasture and feed, farm environments, teat skin, teat milk and bulk tank milk (Supplementary material Fig. S5). However, the relative abundance of the top 25 most abundant genera varied greatly among and within different sample sources (Fig. 4).

In general, source tracking studies should be taken with caution, even when longer sequences are used. Falardeau et al. (2019) used relatively short fragments (V3 region) and

584 Doyle et al. (2017b) targeted the V3–V4 region but used OTUs inferred using a 97%
585 similarity level, while in an amplicon targeted study with higher taxonomic resolution, Skeie
586 et al. (2019) showed that even within the same species different ASVs may have a different
587 distribution (see below). Metagenomic studies may reveal the composition of populations at
588 the strain level and allow the tracking of the sources of the most prevalent and abundant
589 strains, but the cost of studies with sufficient sample sizes would probably be unjustified.

590 The effect of season on the composition of bulk tank milk has been studied by
591 several authors (Doyle et al., 2017a; Li et al., 2018; Zhang et al., 2019): all have found a
592 significant effect, which, however, may be confounded with many other factors (days in
593 lactation, farming system, feeding, etc.). Doyle et al. (2017a) compared samples collected in
594 spring and October, but, due to farming practices in Ireland, this was completely
595 confounded with feed, lactation stage and housing. Mid-lactation samples (collected in
596 spring with cows feeding on pasture outdoors) were significantly different from late
597 lactation samples (collected in autumn when at least for part of the sampling period the
598 cows were housed inside and fed a diet containing concentrate and silage): they had a
599 higher diversity, while for individual milk samples of cows fed on pasture a slightly lower
600 diversity was found (Doyle et al., 2017b), and 85 taxa showed significant differences in
601 abundance between mid- and late-lactation samples.

602 Two further studies (Li et al., 2018; Zhang et al., 2019) have confirmed that a
603 significant seasonal variation exists in the composition of milk microbiota. In both studies
604 members of the genera *Acinetobacter*, *Lactococcus* and *Pseudomonas* were found to be
605 both abundant and highly prevalent, and several genera showed changes of abundance in
606 different seasons (*Pseudomonas*, *Propionibacterium, Flavobacterium*: Li et al., 2018;
607 *Pseudomonas*, two *Lactococcus*, one *Serratia* and one *Acinetobacter:* Zhang et al., 2019).

608 Unfortunately, both studies used a descriptive approach and little or no details were
609 provided on critical factors which are likely to affect microbiota and which may be
610 confounded with the effect of season (farming, breeds, feeding, lactation stage, etc.) and
611 the causes of the observed patterns remain unclear.

612 Using a high-resolution method based on ASV inference, Skeie et al., (2019)
613 confirmed the high variability, even over short time scales, of the composition of bulk tank
614 milk microbiota. The authors analysed 135 milk samples in three samplings from 45 farms in
615 Norway over three months in winter. The farms were located in two geographically distant
616 areas sharing similar climatic conditions, but had three different milking systems. Although
617 milk was collected on average every 3 days, bacterial counts were reasonably low, with a
618 median value of 4.25 log cfu mL^{-1} and only two samples exceeding 10^5 cfu mL^{-1}, with higher
619 values associated systematically with parlour farms with automatic milking systems. Beta-
620 diversity (weighted UniFrac) was significantly affected by all the variables used in this quasi-
621 experimental study (areas, sampling, farms, housing/milking systems) and the abundance of
622 several bacterial genera (including *Pseudomonas, Bacillus, Staphylococcus, Paenibacillus,*
623 *Psychrobacter, Chryseobacterium, Aerococcus,* and *Rhizobium*) was significantly different
624 among geographic areas, farms or sampling dates. The microdiversity for selected taxa, as
625 measured by the number and types of ASVs, was also affected by several factors.

626 Interestingly, *Corynebacterium*, which had the highest number of ASVs but a low
627 average abundance, had the lowest number of ASVs per farm. The composition of
628 populations of *Pseudomonas* and *Lactococcus* were significantly affected by collection day,
629 area and housing/milking system possibly because of the potential variety of environmental
630 sources and growth during refrigerated storage. *Bacillus* and *Streptococcus* populations
631 changed between collection days from the same farm and between farms and geographical

areas. Levels and composition of populations of *Bacillus* and *Paenibacillus*, two aerobic spore formers genera that are of particular concern due to their ability to survive pasteurisation and grow in long shelf life pasteurised milk (Porcellato et al., 2018, 2019), were different between the 2 geographical areas. Fluctuations and diversity in the composition of *Streptococcus* populations were attributed to variability in the contamination of the udder of individual cows, with higher variability in larger farms.

The temperature and duration of storage are clearly two factors that may have a dramatic effect on the composition of the microbiota, with higher refrigeration temperatures favouring the growth of psychrotrophic microorganisms. Doyle et al. (2017a) investigated the effect of temperature (2, 4 or 6 °C) on the composition of microbiota of mid- and late-lactation milk stored for 5 days. There was very little increase in bacterial numbers during storage, as judged by qPCR, except for late-lactation samples stored at 6 °C. No significant differences in composition of the microbiota was noted at the end of storage at 2 °C. However, at both 4 and 6 °C a significant increase was found in the proportion of *Streptococcus* and *Pseudomonas* for samples stored at 4 °C and of *Acinetobacter* and *Pseudomonas* for samples stored at 6 °C. The authors did not attempt to use the qPCR data in conjunction with the relative abundance data obtained by AT metagenomics and it is not clear if the strong decrease in abundance for some taxa (*Staphylococcus*, *Rhodanococcus* and uncultured *Ruminococcaceae*) is due to the compositional nature of abundance tables or if any other taxa increased in number during refrigerated storage.

In an another study (Zhang et al., 2019) prolonged storage at 7 °C also resulted in high relative abundance of several psychrotrophic bacteria belonging to the genera *Pseudomonas, Acinetobacter, Carnobacterium, Chryseobacterium, Erwinia, Hafnia, Flavobacterium, Kluyvera* and *Lactococcus*, with some species (*Pseudomonas fluorescens*

656 and *Pseudomonas psychrophila*) reaching almost 80% of the sequences. A similar trend for
657 increase in the relative abundance of psychrotrophic bacteria was also found in the source
658 tracking study of Falardeau et al. (2019), where bulk and transport tanks contamination may
659 have significantly contributed, together with storage, to the increase of psychrotrophs, with
660 *Pseudomonas*, *Psychrobacter* and *Acinetobacter* among the most abundant genera.

661 Variations in simple milk quality parameters may be associated with changes in the
662 microbiota composition. Rodrigues, Lima, Canniatti-Brazaca, and Bicalho (2017) analysed
663 472 low counts ($<10^4$ cfu mL^{-1}) bulk tank milk samples from dairy farms in New York state
664 (USA) over 2 months and evaluate the correlation between SCC, and standard plate counts
665 (SPC) with composition of the microbiota. As usual, a high diversity and variability was found
666 but a core microbiome was identified across the 19 farms. Significant association was found
667 between the relative abundance of several genera and low or high SCC and SPC counts. In
668 fact, high SCC milk was significantly associated with increased abundance of
669 *Corynebacterium*, *Streptococcus*, *Lactobacillus*, *Coxiella*, *Arthrobacter*, and *Lactococcus*
670 (noticeably only some of this may be potentially associated with mastitis), while high SPC
671 milk (>3.6 log cfu mL^{-1}) had increased abundances of *Acinetobacter, Enterobacteriaceae,*
672 *Corynebacterium* and *Streptococcus* and usually lower diversity: this suggests that shifts in
673 community composition were due to bacterial growth.

674 Even with this very large variability, a core microbiota of bulk tank milk does
675 apparently exist. Using data in FoodMicrobionet, we were able to combine the results, at
676 the genus level or above, for five studies including 199 samples of bulk tank milk from
677 different geographic regions: Ireland (Doyle et al., 2017a), France (Frétin et al., 2018), New
678 Zealand (Li et al., 2018), Norway (Skeie et al., 2019), and Canada (Falardeau et al., 2019).
679 Data on prevalence and abundance of the top most prevalent and abundant taxa are shown

in Supplementary material Fig. S6 and Supplementary material Table 5, while the distribution of abundance for the 25 most prevalent and abundant genera is shown in Fig. 5.

The combined diversity in these five studies is impressive, with almost 2,000 taxa identified at the genus level or above. However, four phyla (*Firmicutes*, *Proteobacteria*, *Bacteroidetes* and *Actinobacteria*) include the majority of the most abundant and prevalent taxa. The top 25 genera include psychrotrophs (*Acinetobacter, Chryseobacterium, Pseudomonas, Psychrobacter*), bacteria associated with gut (*Atopostipes, Bacteroides, Christensenellaceae, Clostridium, Rikenellaceae, Romboutsia, Ruminococcaceae*), or with teat skin (*Staphylococcus, Aerococcus, Turicibacter, Streptococcus, Facklamia, Corynebacterium, Bacillus*), including genera with potentially beneficial microorganisms (*Lactobacillus, Streptococcus, Lactococcus, Staphylococcus, Corynebacterium*). The intra-study variability was sometimes substantial but, surprisingly enough, the composition of the microbiota for some studies, even from different countries, was similar, as assessed by non-metric multidimensional scaling (Supplementary material Fig. S7). Not surprisingly, the diversity was lower for samples obtained from a single farm (ST37: Doyle et al., 2017a; ST39: Frétin et al., 2018; ST74: Falardeau et al., 2019), even when samples were obtained in different seasons (Doyle et al., 2017a; Frétin et al., 2018) or from cows belonging to different breeds (Frétin et al., 2018) or with different feeding regimes (Doyle et al., 2017a; Frétin et al., 2018).

There are very limited data on the composition of bulk tank milk at the dairy farm for species other than cows and the results are difficult to generalise. In two small studies on jennies (de los Dolores Soto del Rio, Dalmasso, Civera, & Bottero, 2017) and on goats (McInnis et al., 2015) the usual high diversity and variability was found, but any inference on the potential sources of variation is impossible due to the limited scope of these studies.

However, in both studies a potential effect of lysozyme, which was naturally abundant in jennies' milk or whose secretion was engineered in goats might have affected the composition of the microbiota.

3.1.4. From the farm to the processing plant

Transfer of bulk tank milk from the dairy farm to the processing plant and further storage and processing steps may alter significantly the composition of milk microbiota as a result of contamination and growth. A number of recent well-structured studies have offered significant insight in this area for both milk processed to become liquid milk products or cheese (Falardeau et al., 2019; Kable et al., 2016, 2019; Porcellato et al., 2018, 2019).

All evidence confirms that contamination form transport trucks and from tanks and equipment at the dairy plant, cleaning routines, heat treatments, and duration and conditions of storage has a significant impact on the structure of microbial communities of milk, and thus potentially affect the quality of cheese and milk due to variations in the abundance of potential starter and non-starter species of the genera (*Streptococcus, Staphylococcus, Macrococcus, Corynebacterium*, etc.) and of spoilage bacteria (*Acinetobacter*, *Pseudomonas,* psychrotrophic spore-formers).

Kable et al. (2016) analysed the microbiota of milk from 899 tankers delivering raw cow milk to two processing farms in California (USA) over three seasons. They confirmed the occurrence of a high variability in the composition of the microbiota, the occurrence of seasonal variations, and found a core microbiota dominated by members of the genera *Streptococcus*, *Staphylococcus* and *Clostridiales*. Notably, only some of the taxa (*Staphylococcus, Bacillus, Enterococcus, Streptococcus, Clostridium, Ruminococcus,*

728 *Corynebacterium, Acinetobacter*) they identified as members of the core cow milk
729 microbiota match those found as being highly prevalent and abundant in studies on bulk
730 tank milk (see Fig. 5 and Supplementary material Table 5), but this might have been due to
731 differences in criteria used to define the core microbiota. Significant differences were found
732 between samples collected in the three seasons. Relative abundance of *Firmicutes* was
733 significantly smaller in spring samples, while those of *Actinobacteria* was higher, while the
734 relative abundance of *Bacteroidetes* was higher in autumn.

735 The authors hypothesised that differences in composition might be due to
736 differential exposure to sources of contamination in different seasons (with possibly more
737 contact with soil, a potential source of *Actinobacteria*, in spring), but clear-cut evidence for
738 this is lacking. While tanker milk samples had low counts (median ~5×10^3 cfu mL^{-1}, as
739 measured by qPCR), growth between refilling cycles of the tankers cannot be excluded as a
740 contributing factor to the observed differences (Kuhn, Meunier-Goddik, & Waite-Cusic,
741 2018). This was confirmed by a later work (Kable et al., 2019; see below) and by the
742 observation that the silo used for milk storage may significantly affect the composition of
743 the microbiota of milk transferred from tanker trucks, with *Pseudomonadales* and
744 *Lactobacillales* being more abundant in the silos. Growth in residual milk in the silo may be a
745 contributing factor (higher cell counts in silos compared with tankers) but stochastic
746 patterns of contamination may contribute. Two groups of silos were identified. In one the
747 microbiota was similar to those of the tankers used to fill them, and had significantly higher
748 proportions of *Streptococcus*, *Corynebacterium, Macrococcus* and *Clostridium*. In the other
749 the microbiota of the silos was distant from those of the tankers (weighted UniFrac
750 distance), and *Acinetobacter* was significantly more abundant.

In a follow-up study Kable et al. (2019) carried out an in depth investigation of the quantitative and qualitative variations of the microbiota in the processing plant. In addition, PMA treatment was used to enrich viable cells after lethal treatments (pasteurisation). They confirmed that OTUs belonging to the genus *Streptococcus* (and tentatively identified as *S. thermophilus/salivarius*) were most abundant and prevalent in milk at all stages, and confirmed the occurrence of a seasonal effect, with some genera (including *Acinetobacter* and *Lactococcus*) more abundant in late summer compared with spring. Several taxa, whose overall abundance was relatively low, showed interesting time-dependent or spatial patterns and were occasionally more abundant. An effect of growth on the composition of bacterial communities was clearly related to the length of time a piece of equipment was operated after cleaning and to heat treatments. Psychrotrophic species clearly increased as storage duration increased: the relative abundance of *Acinetobacter* and *Lactococcus* significantly increased over time in the raw milk silo, while a single *Pseudomonas* OTU was enriched in summer after a post-pasteurisation concentration step. On the other hand, heat treatments caused a decrease in non spore-formers and an increase of thermoduric species and spore-formers, especially when the active fraction of the microbiota was targeted using PMA treatments. *Anoxybacillus*, whose overall abundance was relatively low, seemed to be enriched after long operation time and both *Anoxybacillus* and *Thermus* were enriched after pasteuriSation.

Composition of viable and total microbiota after pasteuriSation was dramatically different for some steps, as shown by weighted UniFrac distance. *Turicibacter*, was significantly enriched in the viable fraction, while the abundance of *Staphylococcus* was significantly lower. Other spore-formers or thermotolerant genera (including *Bacillus*, *Clostridium*, *Anoxybacillus* and *Thermus*) were also more abundant in the viable fraction

after pasteurisation (the difference was not statistically significant) while several other non spore-formers (*Bacilli*, *Clostridia*, and *Actinobacteria*) showed a lower abundance in the viable fraction. This confirms that obtaining a realistic picture of the active fraction of the microbial community by eliminating the contribution from dead or membrane damaged cells is of utmost importance (Erkus et al., 2016; Porcellato & Skeie, 2016).

Kable et al. (2019) also evaluated the impact of time of operation of individual pieces of equipment after cleaning. Within 19 h from cleaning-in-place (CIP) the numbers of viable cells were low (<3200 cells mL^{-1}) and the microbiota composition was diverse, although different species prevailed in different pieces of equipment, mostly depending on whether they contained raw or pasteurised milk: spore-formers, including *Bacillus* and *Anoxybacillus*, were abundant in equipment containing pasteurised milk. After 19 h of operation since the last cleaning, the equipment was divided in two groups. In some samples there was little increase in bacterial numbers and the dominating taxa were close to those appearing in raw milk. In the others, which mostly included milk feeds which had not been pasteurised, were dominated by *Acinetobacter* and *Lactococcus*, while in the concentration step silos the dominating genus was *Anoxybacillus*. Unfortunately, in this study relatively short sequences were used (V4 region of the 16S RNA gene) and proper source tracking at or below the species level is almost impossible.

Porcellato et al. (2018, 2019), using a similar approach, evaluated the dynamics of the microbiota in a Norwegian plant producing liquid pasteurised milk as a function of season of production, milk source, and time and temperature of storage of the pasteurised milk. As in many other papers, a very high diversity and significant differences in the composition of the microbiota between different seasons was found in silo milk, although no significant difference was found for dairies located in two different areas of the country.

Pasteurisation reduced counts by up to 2 log cycles, and only after prolonged (13–14 d, end of shelf life) incubation of the pasteurised milk cartons at abuse temperature (8 °C) counts exceeding 10^7 cfu mL^{-1} (with high presumptive *B. cereus* counts) were obtained, with significantly higher abundances of *Bacillus, Paenibacillus, Solibacillus, Anoxybacillus, Geobacillus* and *Jeotgalicococcus*. Although no treatment was used to separate viable from total bacteria, the composition of the microbiota immediately after pasteurisation was significantly different from that of raw milk: 21 order level OTUs (15 of which *Clostridiales*) were more abundant in raw milk and 27 (including *Lactobacillales, Clostridiales* and *Pseudomonadales*) more abundant in pasteurised milk. After incubation at 4 °C a seasonal variation of the abundance of *Bacillus* was found in pasteurised milk, and this correlated well with the seasonal variation of this genus in raw milk.

In a follow up work, Porcellato et al. (2019) used a more sensitive AT strategy, an in depth analysis of OTUs identified on the basis of 16S RNA gene sequence and assigned to the genus *Bacillus* and amplifying conserved regions of three genes (pantothenate synthase, *panC*; glycerol-3 phosphate transporter *glpT*; pyruvate carboxylase, *pyrC*). A single OTU belonging to the *B. cereus* group made up 99.6% of the sequences of the genus *Bacillus*, while another occasionally dominated the microbiota of spoiled cartons at 8 °C. Use of *panC* as gene target resulted in the highest diversity. A seasonal variation of the composition of the population of the *B. cereus* group was found, and storage at 8 °C resulted in a higher diversity compared with raw milk and milk stored at 4 °C, perhaps simply because more strains had the opportunity to grow over the detection limit. Two *panC* sequence types (ST) were found at relatively high abundance in all samples.

Several other papers have demonstrated that the pasteurisation treatment results in an enrichment of spore-formers (*Bacillus, Anoxybacillus, Turicibacter*) or thermotolerant

bacteria (*Thermus*) in both the viable fraction and in the total microbiota. These genera have been frequently found in raw milk (see Figs. 4 and 5 and Supplementary material Tables 3 and 5) and some have been associated to spoilage, e.g., *Anoxybacillus* (Kable et al., 2019), *Thermus* (Quigley et al., 2016), *Bacillus* (Sattin et al., 2016a; Porcellato et al., 2018).

Turicibacter is universally present in milk, from teat milk to pasteurised milk, but its significance and its potential to grow and spoil milk are not known. *Anoxybacillus* may show some potential to survive and grow in milk (Kable et al., 2019) and has been found in high numbers in Ricotta cheese (Sattin et al., 2016a) where it might contribute to spoilage.

Another recent study (Falardeau et al., 2019) has confirmed that contamination with microorganisms from storage tanks and storage at low temperatures are the main drivers of changes in the composition of microbiota of bulk tank milk. In this study, the composition of the microbiota of milk at the farm bulk tank was dramatically different from that of the storage tank at the cheesemaking plant (with a duration of transportation of about 20 min), and further incubation in a chilled room resulted in the microbiota being dominated by *Pseudomonas*. On the other hand, the composition of the microbiota of the bulk tank and that of the transport tank were very similar and significantly different from that of the pooled tank milk pre- and post-transport.

When data from these three studies (Falardeau et al., 2019; Kable et al., 2019; Porcellato et al., 2018) are compared, a more general picture emerges. The 50 most prevalent and abundant taxa are listed in Supplementary material Table 6. Several of these taxa match those found in bulk tank milk, and, although many taxa whose likely origin is the GI tract are still present at low abundance, the most abundant and prevalent taxa include *Streptococcus, Lactococcus, Pseudomonas, Acinetobacter, Lactobacillus, Staphylococcus, Psychrobacter* and *Escherichia/Shigella*. However, the prevalence and maximum relative

abundance of many genera including psychrotrophs is increased compared with bulk tank milk at the farm, confirming that storage at low temperature is the main driver of the change in composition of the microbiota. Several thermoduric taxa, including spore-formers (*Turicibacter, Anoxybacillus, Bacillus, Clostridium*) are also prevalent and sometimes abundant.

The distribution of abundance of the top 25 genera in raw and HTST (high temperature short time) treated milk at the processing plant is compared in Fig. 6. While there is no guarantee that the nucleic acid target was from viable cells [only data for samples not treated with PMA from Kable et al. (2019) are shown], some major differences are evident between studies (some taxa that were relatively abundant in ST74 and ST87 are almost absent in ST38) and, within study, between HTST and raw milk. In the latter, the relative proportion of psychrotrophs tends to be higher, while that of thermoduric species, including both non spore-formers and spore-formers is higher.

Several papers describing the evolution of the composition of cheese microbiota report the composition of raw milk prior to starter addition. These include studies on cows' milk cheeses (Alessandria et al., 2016; Bokulich & Mills, 2013; Calasso et al., 2016; Carafa et al., 2019; De Filippis et al., 2016; De Pasquale, Di Cagno, Buchin, De Angelis, & Gobbetti, 2014b; Dolci, De Filippis, La Storia, Ercolini, & Cocolin, 2014; Falardeau et al., 2019; Frétin et al., 2018; Giello et al., 2017; Masoud et al., 2011), water-buffalo cheeses (Ercolini, De Filippis, La Storia, & Iacono, 2012), ewes' milk cheeses (De Pasquale et al., 2014a), fermented yak milk products (Jiang et al., 2019) and camel milk (Amrouche, Mounier, Pawtowski, Thomas, & Picot, 2020; Zhao et al., 2019). Although these studies invariably show a high diversity in milk microbiota, the description of the storage conditions and

duration, and of the source of the milk is generally insufficient to allow any comparison with studies focusing on milk.

4. Conclusions

The microbiota of milk is probably one of the most complex food microbial communities, because of the multiplicity of sources of contamination (Addis et al., 2016; Derakshani et al., 2018a). The availability of HTS methods, with the ability to study in detail both the taxonomic structure and the functionality of microbial communities has revolutionised our ability to study the structure and function of food microbiomes and has undoubtedly greatly contributed to our understanding of the factors affecting contamination patterns and successions in milk and dairy foods. The availability of raw and processed sequence data (see Supplementary material and data) greatly enhances our ability of combining and analysing results from different studies, thus, facilitating metastudies. At almost 9 years from the publication of the first paper on the microbiota of dairy products, sequencing platforms, methods and bioinformatic approaches have evolved greatly and, although most recent papers use similar pipelines, there is still a need for a consensus of dairy microbiologists on SOPs, which would improve confidence in the results and comparability of studies. Careful documentation of all potential factors affecting the composition of milk microbiota is of uttermost importance for the interpretation of the results.

Even with these caveats, our understanding of the relationships between udder health and milk quality has significantly improved. Mastitis and dysbiosis of microbial communities of the udder are strongly related, although the cause-effect relationship is not

completely clear. In addition, HTS approaches allow to shed light on the potential causative agents in culture negative and sub-clinical cases and to clarify the polymicrobial nature of some mastitis cases.

Practically all conceivable factors related to farming and storage of raw milk have been shown to affect the composition of milk microbiota, and several sources (faeces, pasture, feed, milking equipment, storage tanks, etc.) might contribute microorganisms relevant for the quality of dairy milk and fermented dairy products. However, the low taxonomic resolution of some studies (that track genera, or at best, species, not strains) still obscures the potential contribution of each source and more effort should be probably devoted to disentangle the relative contribution of contamination, growth at low temperature and ability to survive pasteurisation treatments in determining the potential for microorganisms from different sources to affect the quality of dairy products. Recent findings on the origin and survival of spore-formers and other thermoduric microorganisms in the dairy plant (Kable et al., 2019; Porcellato et al., 2018, 2019) have indeed provided insights which might contribute to improving practices in cleaning, sanitation and heat treatment of milk.

However, still much remains to be done on the potential contribution of milk in terms of starter and non-starter microorganisms relevant to cheese production: although HTS methods are indeed much more sensitive than other cultivation dependent and independent approaches, the low levels of contamination of hygienically produced milk complicate the tracking of species and strains. The decreasing costs of shotgun approaches and the availability of powerful bioinformatic pipelines might in the near future open new avenues to the study of sources of milk contamination.

Acknowledgements

This work was partially funded by Ministero delle Politiche Agricole, Alimentari e Forestali, Rome, Italy, Project MILKBIOACTINCAPS n. ID9

References

Addis, M. F., Tanca, A., Uzzau, S., Oikonomou, G., Bicalho, R. C., & Moroni, P. (2016). The bovine milk microbiota: insights and perspectives from -omics studies. *Molecular bioSystems*, *12*, 2359–2372.

Afshari, R., Pillidge, C., Dias, D., Osborn, A., & Gill, H. (2018). Cheesomics: the future pathway to understanding cheese flavour and quality. *Critical Reviews in Food Science and Nutrition*, *60*, 1–15.

Alessandria, V., Ferrocino, I., De Filippis, F., Fontana, M., Rantsiou, K., Ercolini, D., et al. (2016). Microbiota of an Italian Grana like cheese during manufacture and ripening unraveled by 16S rRNA-based approaches. *Applied and Environmental Microbiology*, *82*, 3988–3995.

Amrouche, T., Mounier, J., Pawtowski, A., Thomas, F., & Picot, A. (2020). Microbiota associated with dromedary camel milk from Algerian Sahara. *Current Microbiology*, *77*, 24–31.

Andrews, T., Neher, D., Weicht, T., & Barlow, J. (2019). Mammary microbiome of lactating organic dairy cows varies by time, tissue site, and infection status. *PLoS One*, *14*, Article e0225001.

Angelopoulou, A., Holohan, R., Rea, M. C., Warda, A. K., Hill, C., & Ross, R. P. (2019). Bovine

mastitis is a polymicrobial disease requiring a polydiagnostic approach. *International Dairy Journal*, *99*, Article 104539.

Bhatt, V. D., Ahir, V. B., Koringa, P. G., Jakhesara, S. J., Rank, D. N., Nauriyal, D. S., et al. (2012). Milk microbiome signatures of subclinical mastitis-affected cattle analysed by shotgun sequencing. *Journal of Applied Microbiology*, *112*, 639–650.

Bokulich, N. A., & Mills, D. A. (2013). Facility-specific "house" microbiome drives microbial landscapes of artisan cheesemaking plants. *Applied and Environmental Microbiology*, *79*, 5214–5223.

Bonsaglia, E. C. R., Gomes, M. S., Canisso, I. F., Zhou, Z., Lima, S. F., Rall, V. L. M., et al. (2017). Milk microbiome and bacterial load following dry cow therapy without antibiotics in dairy cows with healthy mammary gland. *Scientific Reports*, *7*, 8067–10.

Breitwieser, F. P., Lu, J., & Salzberg, S. L. (2017). A review of methods and databases for metagenomic classification and assembly. *Briefings in Bioinformatics*, *3*, Article 31.

Calasso, M., Ercolini, D., Mancini, L., Stellato, G., Minervini, F., Di Cagno, R., et al. (2016). Relationships among house, rind and core microbiotas during manufacture of traditional Italian cheeses at the same dairy plant. *Food Microbiology*, *54*, 115–126.

Callahan, B., McMurdie, P., Rosen, M., Han, A., Johnson, A., & Holmes, S. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods*, *13*, 581–583.

Carafa, I., Stocco, G., Nardin, T., Larcher, R., Bittante, G., Tuohy, K., et al. (2019). Production of naturally γ-aminobutyric acid-enriched cheese using the dairy strains *Streptococcus thermophilus* 84C and *Lactobacillus brevis* DSM 32386. *Frontiers in Microbiology*, *10*, Article 93.

Castro, I., Alba, C., Aparicio, M., Arroyo, R., Jiménez, L., Fernández, L., et al. (2019).

966 Metataxonomic and immunological analysis of milk from ewes with or without a
967 history of mastitis. *Journal of Dairy Science*, *102*, 9298–9311.
968 Catozzi, C., Sanchez Bonastre, A., Francino, O., Lecchi, C., De Carlo, E., Vecchio, D., et al.
969 (2017). The microbiota of water buffalo milk during mastitis. *PLoS One*, *12,* Article
970 e0184710.
971 Cremonesi, P., Ceccarani, C., Curone, G., Severgnini, M., Pollera, C., Bronzo, V., et al. (2018).
972 Milk microbiome diversity and bacterial group prevalence in a comparison between
973 healthy Holstein Friesian and Rendena cows. *PLoS One*, *13*, Article e0205054.
974 Dahlberg, J., Sun, L., Waller, K. P., Östensson, K., McGuire, M., Agenäs, S., et al. (2019).
975 Microbiota data from low biomass milk samples is markedly affected by laboratory
976 and reagent contamination. *PLoS One*, *14*, Article e0218257.
977 De Filippis, F., Genovese, A., Ferranti, P., Gilbert, J. A., & Ercolini, D. (2016).
978 Metatranscriptomics reveals temperature-driven functional changes in microbiome
979 impacting cheese maturation rate. *Scientific Reports*, *6*, Article 21871.
980 De Filippis, F., La Storia, A., Stellato, G., Gatti, M., & Ercolini, D. (2014). A selected core
981 microbiome drives the early stages of three popular Italian cheese manufactures.
982 *PLoS One*, *9*(, Article e89680.
983 De Filippis, F., Parente, E., & Ercolini, D. (2017). Metagenomics insights into food
984 fermentations. *Microbial Biotechnology*, *10*, 91–102.
985 de los Dolores Soto del Rio, M., Dalmasso, A., Civera, T., & Bottero, M. T. (2017).
986 Characterization of bacterial communities of donkey milk by high-throughput
987 sequencing. *International Journal of Food Microbiology*, *251*, 67–72.
988 De Pasquale, I., Calasso, M., Mancini, L., Ercolini, D., La Storia, A., De Angelis, M., et al.
989 (2014a). Causal relationship between microbial ecology dynamics and proteolysis

during manufacture and ripening of Canestrato Pugliese PDO cheese. *Applied and Environmental Microbiology*, *80*, 4085–4094.

De Pasquale, I., Di Cagno, R., Buchin, S., De Angelis, M., & Gobbetti, M. (2014b). Microbial ecology dynamics reveal a succession in the core microbiota involved in the ripening of pasta filata caciocavallo pugliese cheese. *Applied and Environmental Microbiology*, *80*, 6243–6255.

Derakhshani, H., Fehr, K. B., Sepehri, S., Francoz, D., De Buck, J., Barkema, H. W., et al. (2018a). Invited review: Microbiota of the bovine udder: Contributing factors and potential implications for udder health and mastitis susceptibility. *Journal of Dairy Science*, *101*, 10605–10625.

Derakhshani, H., Plaizier, J. C., De Buck, J., Barkema, H. W., & Khafipour, E. (2018b). Composition of the teat canal and intramammary microbiota of dairy cows subjected to antimicrobial dry cow therapy and internal teat sealant. *Journal of Dairy Science*, *101*, 10191–10205.

Dolci, P., De Filippis, F., La Storia, A., Ercolini, D., & Cocolin, L. (2014). rRNA-based monitoring of the microbiota involved in Fontina PDO cheese production in relation to different stages of cow lactation. *International Journal of Food Microbiology*, *185*, 127–135.

Douglas, G. M., Maffei, V. J., Zaneveld, J., Yurgel, S. N., Brown, J. R., Taylor, C. M., et al. (2019), PICRUSt2: An improved and extensible approach for metagenome inference. *bioRxiv*, *8*, Article 672295.

Doyle, C. J., Gleeson, D., O'Toole, P. W., & Cotter, P. D. (2017a). High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions. *International Journal of Food*

Microbiology, *255*, 1–6.

Doyle, C. J., Gleeson, D., O'Toole, P. W., & Cotter, P. D. (2017b). Impacts of seasonal housing and teat preparation on raw milk microbiota: a high-throughput sequencing study. *Applied and Environmental Microbiology*, *83*, Article e02694-16.

Emerson, J. B., Adams, R. I., Román, C. M. B., Brooks, B., Coil, D. A., Dahlhausen, K., et al. (2017). Schrödinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome*, *5*, Article 86.

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. *Applied and Environmental Microbiology*, *78*, 8142–8145.

Erkus, O., de Jager, V. C. L., Geene, R. T. C. M., van Alen-Boerrigter, I., Hazelwood, L., van Hijum, S. A. F. T., et al. (2016). Use of propidium monoazide for selective profiling of viable microbial cells during Gouda cheese ripening. *International Journal of Food Microbiology*, *228*, 1–9.

Esteban-Blanco, C., Gutiérrez-Gil, B., Puente-Sánchez, F., Marina, H., Tamames, J., Acedo, A., et al. (2019). Microbiota characterization of sheep milk and its association with somatic cell count using 16s rRNA gene sequencing. *Journal of Animal Breeding and Genetics = Zeitschrift Fur Tierzuchtung Und Zuchtungsbiologie*, *81*, Article 411.

Falardeau, J., Keeney, K., Trmčić, A., Kitts, D., & Wang, S. (2019). Farm-to-fork profiling of bacterial communities associated with an artisan cheese production facility. *Food Microbiology*, *83*, 48–58.

Falentin, H., Rault, L., Nicolas, A., Bouchard, D. S., Lassalas, J., Lamberton, P., et al. (2016). Bovine teat microbiome analysis revealed reduced alpha diversity and significant changes in taxonomic profiles in quarters with a history of mastitis. *Frontiers in*

Microbiology, *7*, Article 954.

Faust, K., & Raes, J. (2012). Microbial interactions: from networks to models. *Nature Reviews Microbiology*, *10*, 538–550.

Frantzen, C. A., Kleppen, H. P., & Holo, H. (2018). *Lactococcus lactis* diversity in undefined mixed dairy starter cultures as revealed by comparative genome analyses and targeted amplicon sequencing of *epsD*. *Applied and Environmental Microbiology*, *84*, Article 2725.

Frétin, M., Martin, B., Rifa, E., Isabelle, V.-M., Pomiès, D., Ferlay, A., et al. (2018). Bacterial community assembly from cow teat skin to ripened cheeses is influenced by grazing systems. *Scientific Reports*, *8*, Article 200.

Ganda, E. K., Bisinotto, R. S., Lima, S. F., Kronauer, K., Decter, D. H., Oikonomou, G., et al. (2016). Longitudinal metagenomic profiling of bovine milk to assess the impact of intramammary treatment using a third-generation cephalosporin. *Scientific Reports*, *6*, Article 37565.

Ganda, E. K., Gaeta, N., Sipka, A., Pomeroy, B., Oikonomou, G., Schukken, Y. H., et al. (2017). Normal milk microbiome is reestablished following experimental infection with Escherichia coli independent of intramammary antibiotic treatment with a third-generation cephalosporin in bovines. *Microbiome*, *5*, 74–17.

Giello, M., La Storia, A., Masucci, F., Di Francia, A., Ercolini, D., & Villani, F. (2017). Dynamics of bacterial communities during manufacture and ripening of traditional Caciocavallo of Castelfranco cheese in relation to cows' feeding. *Food Microbiology*, *63*, 170–177.

Guzzon, R., Carafa, I., Tuohy, K., Cervantes, G., Vernetti, L., Barmaz, A., et al. (2017). Exploring the microbiota of the red-brown defect in smear-ripened cheese by 454-pyrosequencing and its prevention using different cleaning systems. *Food*

Microbiology, *62*, 160–168.

Hoque, M. N., Istiaq, A., Clement, R. A., Sultana, M., Crandall, K. A., Siddiki, A. Z., et al. (2019). Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis. *Scientific Reports*, *9*, Article 13536.

Jiang, Y., Li, N., Wang, Q., Liu, Z., Lee, Y., Liu, X., et al. (2019). Microbial diversity and volatile profile of traditional fermented yak milk. *Journal of Dairy Science*, *103*, 87–97.

Jin, H., Mo, L., Pan, L., Hou, Q., Li, C., Darima, I., et al. (2018). Using PacBio sequencing to investigate the bacterial microbiota of traditional Buryatian cottage cheese and comparison with Italian and Kazakhstan artisanal cheeses. *Journal of Dairy Science*, *101*, 6885–6896.

Kable, M. E., Srisengfa, Y., Laird, M., Zaragoza, J., McLeod, J., Heidenreich, J., et al. (2016). The core and seasonal microbiota of raw bovine milk in tanker trucks and the impact of transfer to a milk processing facility. *mBio*, *7*, Article e00836–16.

Kable, M. E., Srisengfa, Y., Xue, Z., Coates, L. C., & Marco, M. L. (2019). Viable and total bacterial populations undergo equipment- and time-dependent shifts during milk processing. *Applied and Environmental Microbiology*, *85*, Article 190.

Kamimura, B. A., De Filippis, F., Sant'Ana, A. S., & Ercolini, D. (2019). Large-scale mapping of microbial diversity in artisanal Brazilian cheeses. *Food Microbiology*, *80*, 40–49.

Kastman, E. K., Kamelamela, N., Norville, J. W., Cosetta, C. M., Dutton, R. J., & Wolfe, B. E. (2016). Biotic interactions shape the ecological distributions of *Staphylococcus* species. *mBio*, *7*, Article e01157–16.

Kuehn, J. S., Gorden, P. J., Munro, D., Rong, R., Dong, Q., Plummer, P. J., et al. (2013). Bacterial community profiling of milk samples as a means to understand culture-negative bovine clinical mastitis. *PLoS One*, *8*, Article e61959.

1086 Kuhn, E., Meunier-Goddik, L., & Waite-Cusic, J. (2018). Effect of leaving milk trucks empty
1087 and idle for 6 h between raw milk loads. *Journal of Dairy Science*, *101*, 1767–1776.
1088 Layeghifard, M., Hwang, D. M., & Guttman, D. S. (2017). Disentangling interactions in the
1089 microbiome: a network perspective. *Trends in Microbiology*, *25*, 217–228.
1090 Levante, A., Filippis, F., Storia, A., Gatti, M., Neviani, E., Ercolini, D., et al. (2017). Metabolic
1091 gene-targeted monitoring of non-starter lactic acid bacteria during cheese ripening.
1092 *International Journal of Food Microbiology*, *257*, 276–284.
1093 Li, J., Zheng, Y., Xu, H., Xi, X., Hou, Q., Feng, S., et al. (2017). Bacterial microbiota of
1094 Kazakhstan cheese revealed by single molecule real time (SMRT) sequencing and its
1095 comparison with Belgian, Kalmykian and Italian artisanal cheeses. *BMC Microbiology*,
1096 *17*, Article 13.
1097 Li, N., Wang, Y., You, C., Ren, J., Chen, W., Zheng, H., et al. (2018). Variation in raw milk
1098 microbiota throughout 12 months and the impact of weather conditions. *Scientific*
1099 *Reports*, *8*, Article 2371.
1100 Masoud, W., Takamiya, M., Vogensen, F. K., Lillevang, S., Al-Soud, W. A., Sørensen, S. J., et
1101 al. (2011). Characterization of bacterial populations in Danish raw milk cheeses made
1102 with different starter cultures by denaturating gradient gel electrophoresis and
1103 pyrosequencing. *International Dairy Journal*, *21*, 142–148.
1104 McInnis, E. A., Kalanetra, K. M., Mills, D. A., & Maga, E. A. (2015). Analysis of raw goat milk
1105 microbiota: impact of stage of lactation and lysozyme on microbial diversity. *Food*
1106 *Microbiology*, *46*, 121–131.
1107 Meola, M., Rifa, E., Shani, N., Delbès, C., Berthoud, H., & Chassard, C. (2019). DAIRYdb: a
1108 manually curated reference database for improved taxonomy annotation of 16S
1109 rRNA gene sequences from dairy products. *BMC Genomics*, *20*, Article 560.

Metzger, S. A., Hernandez, L. L., Skarlupka, J. H., Suen, G., Walker, T. M., & Ruegg, P. L. (2018a). Influence of sampling technique and bedding type on the milk microbiota: Results of a pilot study. *Journal of Dairy Science*, *101*, 6346–6356.

Metzger, S. A., Hernandez, L. L., Skarlupka, J. H., Walker, T. M., Suen, G., & Ruegg, P. L. (2018b). A Cohort study of the milk microbiota of healthy and inflamed bovine mammary glands from dryoff through 150 days in milk. *Frontiers in Veterinary Science*, *5*, Article 247.

Milani, C., Alessandri, G., Mancabelli, L., Lugli, G., Longhi, G., Anzalone, R., et al. (2019). Bifidobacterial distribution across Italian cheeses produced from raw milk. *Microorganisms*, *7*, Article 599.

Milani, C., Duranti, S., Mangifesta, M., Lugli, G., Turroni, F., Mancabelli, L., et al. (2018). Phylotype-level profiling of Lactobacilli in highly complex environments by means of an internal transcribed spacer-based metagenomic approach. A*pplied and Environmental Microbiology*, *84*, Article e00706-18.

Mo, L., Yu, J., Jin, H., Hou, Q., Yao, C., Ren, D., et al. (2019). Investigating the bacterial microbiota of traditional fermented dairy products using propidium monoazide with single-molecule real-time sequencing. *Journal of Dairy Science*, *102*, 3912–3923.

Murugesan, S., Reyes-Mata, M. P., Nirmalkar, K., Chavez-Carbajal, A., Juárez-Hernández, J. I., Torres-Gómez, R. E., et al. (2018). Profiling of bacterial and fungal communities of Mexican cheeses by high throughput DNA sequencing. *Food Research International*, *113*, 371–381.

Oikonomou, G., Addis, M., Chassard, C., Nader-Macias, M., Grant, I., Delbès, C., et al. (2020). Milk microbiota: what are we exactly talking about? *Frontiers in Microbiology*, *11*, Article 60.

Oikonomou, G., Bicalho, M. L., Meira, E., Rossi, R. E., Foditsch, C., Machado, V. S., et al. (2014). Microbiota of cow's milk; distinguishing healthy, sub-clinically and clinically diseased quarters. *PLoS One*, *9*, Article e85904.

Oikonomou, G., Machado, V. S., Santisteban, C., Schukken, Y. H., & Bicalho, R. C. (2012). Microbial diversity of bovine mastitic milk as described by pyrosequencing of metagenomic 16s rDNA. *PLoS One*, *7*, Article e47671.

O'Sullivan, D. J., Fallico, V., O'Sullivan, O., McSweeney, P. L. H., Sheehan, J. J., Cotter, P. D., et al. (2015). High-throughput DNA sequencing to survey bacterial histidine and tyrosine decarboxylases in raw milk cheeses. *BMC Microbiology*, *15*, Article 266.

Oultram, J. W. H., Ganda, E. K., Boulding, S. C., Bicalho, R. C., & Oikonomou, G. (2017). A Metataxonomic approach could be considered for cattle clinical mastitis diagnostics. *Frontiers in Veterinary Science*, *4*, Article 36.

Pang, M., Xie, X., Bao, H., Sun, L., He, T., Zhao, H., et al. (2018). Insights into the bovine milk microbiota in dairy farms with different incidence rates of subclinical mastitis. *Frontiers in Microbiology*, *9*, Article 2359.

Pangallo, D., Kraková, L., Puškárová, A., Šoltys, K., Bučková, M., Koreňová, J., et al. (2019). Transcription activity of lactic acid bacterial proteolysis-related genes during cheese maturation. *Food Microbiology*, *82*, 416–425.

Parente, E., Cocolin, L., De Filippis, F., Zotta, T., Ferrocino, I., O'Sullivan, O., et al. (2016a). FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis. *International Journal of Food Microbiology*, *219*, 28–37.

Parente, E., De Filippis, F., Ercolini, D., Ricciardi, A., & Zotta, T. (2019). Advancing integration of data on food microbiome studies: FoodMicrobionet 3.1, a major upgrade of the

FoodMicrobionet database. *International Journal of Food Microbiology*, *305*, Article 108249.

Parente, E., Guidone, A., Matera, A., De Filippis, F., Mauriello, G., & Ricciardi, A. (2016b). Microbial community dynamics in thermophilic undefined milk starter cultures. *International Journal of Food Microbiology*, *217*, 59–67.

Parente, E., Zotta, T., Faust, K., De Filippis, F., & Ercolini, D. (2018). Structure of association networks in food bacterial communities. *Food Microbiology*, *73*, 49–60.

Patel, R., Kunjadia, P., Koringa, P., Joshi, C., & Kunjadiya, A. (2019). Microbiological profiles in clinical and subclinical cases of mastitis in milking Jafarabadi buffalo. *Research in Veterinary Science*, *125*, 94–99.

Patel, R., Pandit, R., Bhatt, V., Kunjadia, P., Nauriyal, D., Koringa, P., et al. (2016). Metagenomic approach to study the bacterial community in clinical and subclinical mastitis in buffalo. *Meta Gene*, *12*, 4–12.

Pollock, J., Glendinning, L., Wisedchanwet, T., & Watson, M. (2018). The madness of microbiome: attempting to find consensus “best practice” for 16S microbiome studies. *Applied and Environmental Microbiology*, *84*, Article 3225.

Porcellato, D., & Skeie, S. B. (2016). Bacterial dynamics and functional analysis of microbial metagenomes during ripening of Dutch-type cheese. *International Dairy Journal*, *61*, 182–188.

Porcellato, D., Aspholm, M., Skeie, S. B., & Mellegård, H. (2019). Application of a novel amplicon-based sequencing approach reveals the diversity of the *Bacillus cereus* group in stored raw and pasteurized milk. *Food Microbiology*, *81*, 32–39.

Porcellato, D., Aspholm, M., Skeie, S. B., Monshaugen, M., Brendehaug, J., & Mellegård, H. (2018). Microbial diversity of consumption milk during processing and storage.

International Journal of Food Microbiology, *266*, 21–30.

Quigley, L., O'Sullivan, D. J., Daly, D., O'Sullivan, O., Burdikova, Z., Vana, R., et al. (2016). *Thermus* and the pink discoloration defect in cheese. *mSystems*, *1,* Article e00023-16

Quigley, L., O'Sullivan, O., Stanton, C., Beresford, T. P., Ross, R. P., Fitzgerald, G. F., et al. (2013). The complex microbiota of raw milk. *FEMS Microbiology Reviews*, *37*, 664–698.

Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nature Biotechnology*, *35*, 833–844.

Rainard, P. (2017). Mammary microbiota of dairy ruminants: fact or fiction? *Veterinary Research*, *48*, 25–10.

Ramezani, M., Hosseini, S. M., Ferrocino, I., Amoozegar, M. A., & Cocolin, L. (2017). Molecular investigation of bacterial communities during the manufacturing and ripening of semi-hard Iranian Liqvan cheese. *Food Microbiology*, *66*, 64–71.

Ricciardi, A., De Filippis, F., Zotta, T., Facchiano, A., Ercolini, D., & Parente, E. (2016). Polymorphism of the phosphoserine phosphatase gene in *Streptococcus thermophilus* and its potential use for typing and monitoring of population diversity. *International Journal of Food Microbiology*, *236*, 138–147.

Rodrigues, M. X., Lima, S. F., Canniatti-Brazaca, S. G., & Bicalho, R. C. (2017). The microbiome of bulk tank milk: Characterization and associations with somatic cell count and bacterial count. *Journal of Dairy Science*, *100*, 2536–2552.

Ruegg, P. (2017). A 100-Year Review: Mastitis detection, management, and prevention. *Journal of Dairy Science*, *100*, 10381–10397.

Sattin, E., Andreani, N. A., Carraro, L., Fasolato, L., Balzan, S., Novelli, E., et al. (2016a). Microbial dynamics during shelf-life of industrial Ricotta cheese and identification of

a *Bacillus* strain as a cause of a pink discolouration. *Food Microbiology*, *57*, 8–15.

Sattin, E., Andreani, N. A., Carraro, L., Lucchini, R., Fasolato, L., Telatin, A., et al. (2016b). A multi-omics approach to evaluate the quality of milk whey used in Ricotta cheese production. *Frontiers in Microbiology*, *7*, Article 1272.

Skeie, S. B., Håland, M., Thorsen, I. M., Narvhus, J., & Porcellato, D. (2019). Bulk tank raw milk microbiota differs within and between farms: A moving goalpost challenging quality control. *Journal of Dairy Science*, *102*, 1959–1971.

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. *Applied and Environmental Microbiology*, *81*, 7893–7904.

Sudarikov, K., Tyakht, A., & Alexeev, D. (2017). Methods for the metagenomic data visualization and analysis. *Current Issues in Molecular Biology*, *24*, 37–58.

Taponen, S., McGuinness, D., Hiitiö, H., Simojoki, H., Zadoks, R., & Pyörälä, S. (2019). Bovine milk microbiome: a more complex issue than expected. *Veterinary Research*, *50,* 44–15.

Tilocca, B., Costanzo, N., Morittu, V., Spina, A., Soggiu, A., Britti, D., et al. (2020). Milk microbiota: Characterization methods and role in cheese production. *Journal of Proteomics*, *210*, Article 103534.

Vasquez, A. K., Ganda, E. K., Capel, M. B., Eicker, S., Virkler, P. D., Bicalho, R. C., et al. (2019). The microbiome of *Escherichia coli* and culture-negative nonsevere clinical mastitis: characterization and associations with linear score and milk production. *Journal of Dairy Science*, *102*, 578–594.

Walsh, A. M., Crispie, F., O'Sullivan, O., Finnegan, L., Claesson, M. J., & Cotter, P. D. (2018). Species classifier choice is a key consideration when analysing low-complexity food

microbiome data. *Microbiome*, *6*, Article 50.

Wolfe, B. E., Button, J. E., Santarelli, M., & Dutton, R. J. (2014). Cheese rind communities provide tractable systems for in situ and in vitro studies of microbial diversity. *Cell*, *158*, 422–433.

Yang, C., Zhao, F., Hou, Q., Wang, J., Li, M., & Sun, Z. (2019). PacBio sequencing reveals bacterial community diversity in cheeses collected from different regions. *Journal of Dairy Science*, *103*, 1238–1249.

Yeluri Jonnala, B. R., McSweeney, P. L. H., Sheehan, J. J., & Cotter, P. D. (2018). Sequencing of the cheese microbiome and Its relevance to industry. *Frontiers in Microbiology*, *9*, 1890–1912.

Yu, J., Mo, L., Pan, L., Yao, C., Ren, D., An, X., et al. (2018). Bacterial microbiota and metabolic character of traditional sour cream and butter in Buryatia, Russia. *Frontiers in Microbiology*, *9*, Article 2496.

Zhang, D., Palmer, J., Teh, K. H., Biggs, P., & Flint, S. (2019). 16S rDNA high-throughput sequencing and MALDI-TOF MS are complementary when studying psychrotrophic bacterial diversity of raw cows' milk. *International Dairy Journal*, *97*, 86–91.

Zhang, R., Huo, W., Zhu, W., & Mao, S. (2015). Characterization of bacterial community of raw milk from dairy cows during subacute ruminal acidosis challenge by high-throughput sequencing. *Journal of the Science of Food and Agriculture*, *95*, 1072–1079.

Zhao, J., Fan, H., Kwok, L., Guo, F., Ji, R., Ya, M., et al. (2019). Analyses of physicochemical properties, bacterial microbiota, and lactic acid bacteria of fresh camel milk collected in Inner Mongolia. *Journal of Dairy Science*, *103*, 106–116.

Zhong, Y., Xue, M., & Liu, J. (2018). Composition of rumen bacterial community in dairy cows

1254 with different levels of somatic cell counts. *Frontiers in Microbiology*, *9*, Article 1551.

Legends to figures

Fig. 1. Stacked bar plot showing the relative abundance of bacterial classes on cows' teats in two studies (Falardeau et al., 2019; Frétin et al., 2018). Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019).

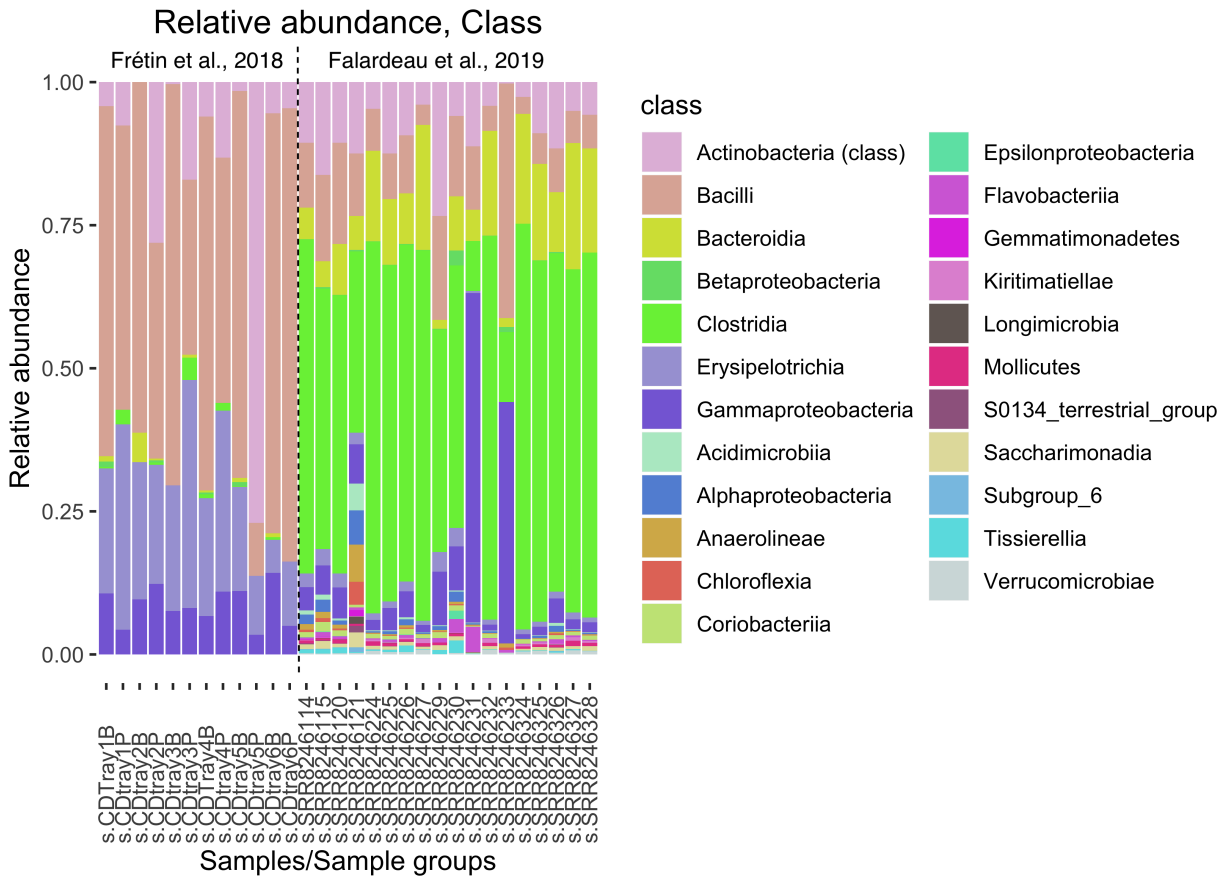
Fig. 2. Stacked bar plot showing the relative abundance of the 20 most abundant taxa (at the genus level or above are shown) in individual samples from teat milk from Cremonesi et al. (2018; colostrum samples were removed) and Falardeau et al. (2019). Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019). For the Cremonesi et al. (2018) samples, HF and REN in sample names indicate samples from Holstein Fresians and Rendena cows, respectively.

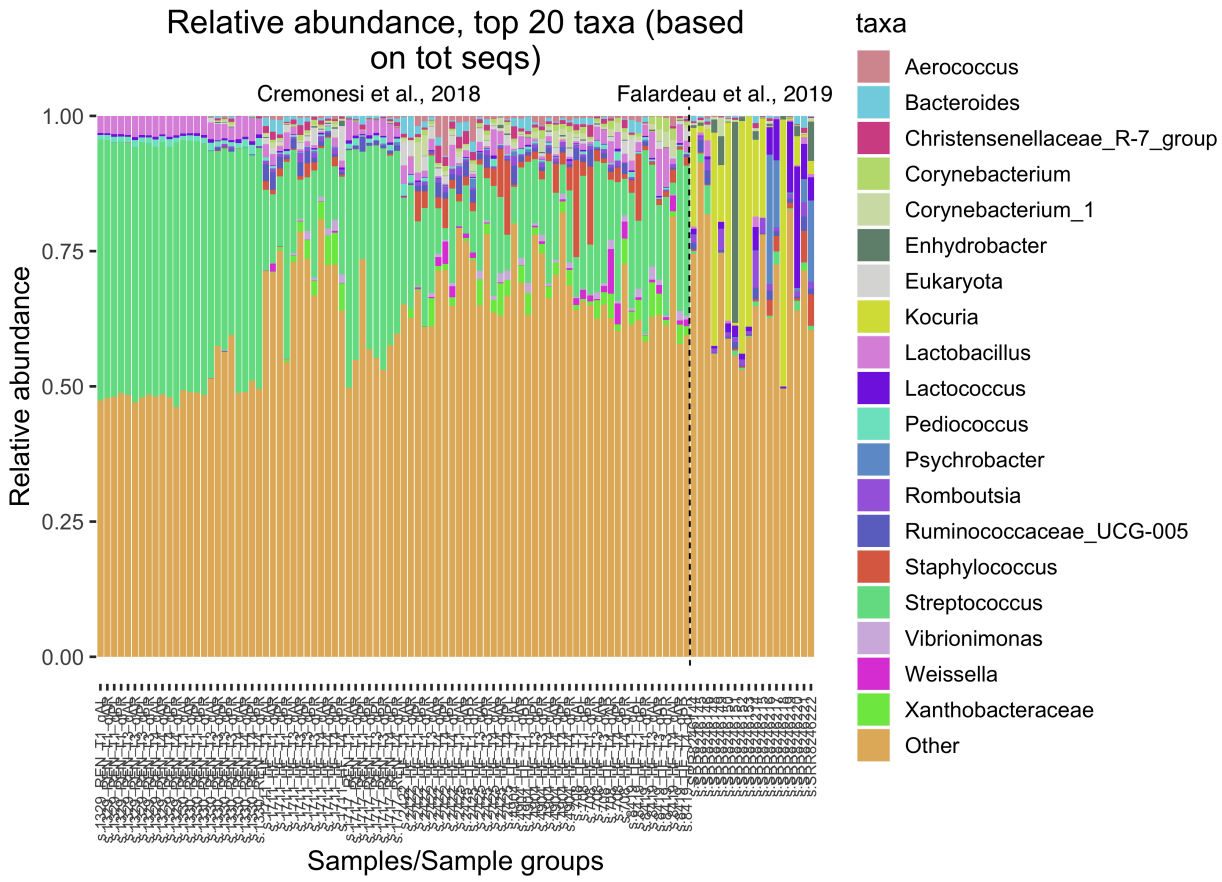
Fig. 3. Stacked bar plot showing the relative abundance of the 24 most abundant taxa (at the genus level or above are shown) for teat milk from cows [ST47, Cremonesi et al., 2018 (colostrum samples were removed); ST74, Falardeau et al., 2019], ewes (Castro et al., 2019) and water-buffaloes (Catozzi et al., 2017; only samples from healthy quarters are shown). Samples from each study have been pooled. Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019).

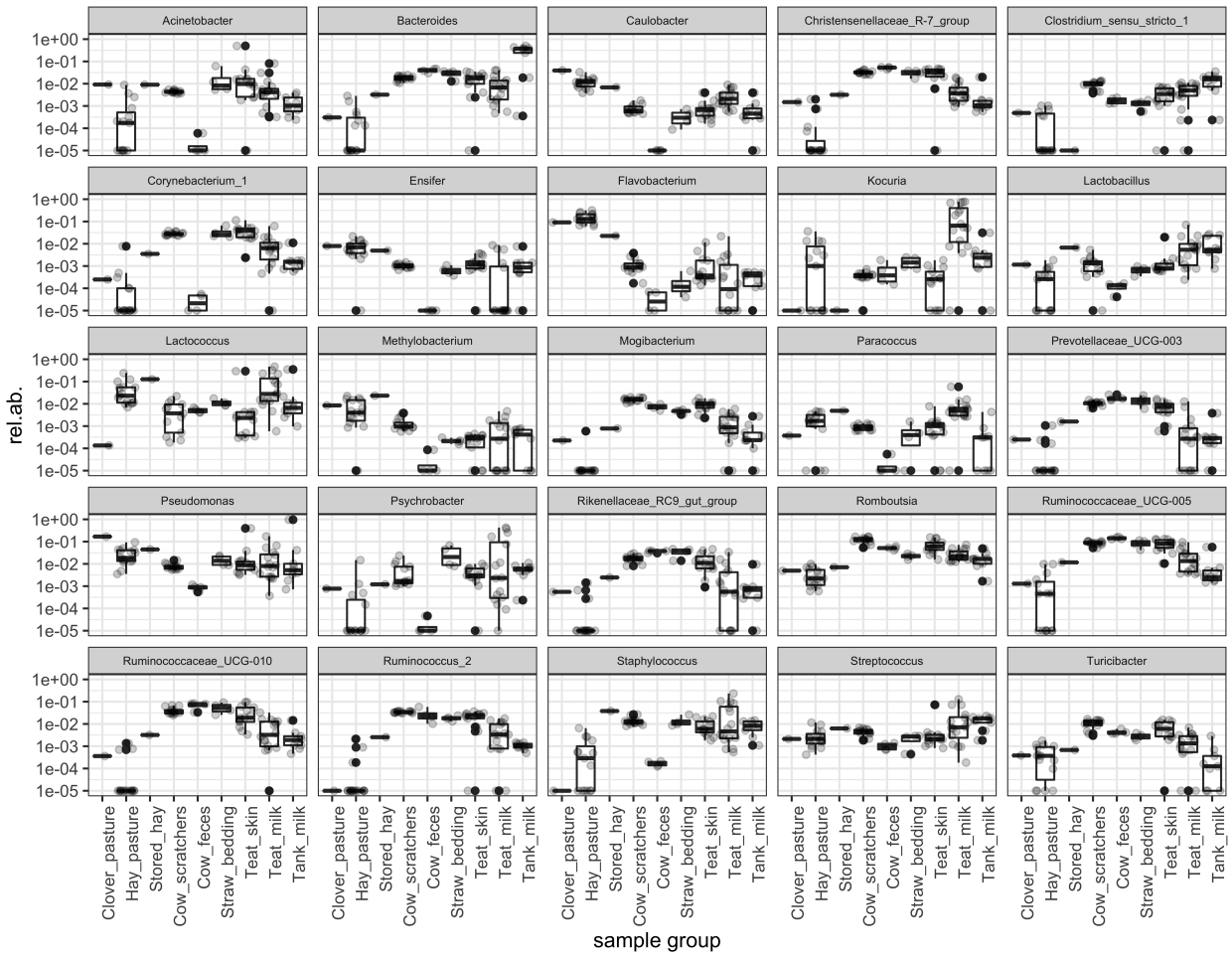
Fig. 4. Boxplots for the distribution of relative abundance for the 25 most abundant and prevalent genera in pasture and feed, farm environments, teat skin, teat and bulk tank milk in Falardeau et al. (2019). Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019).

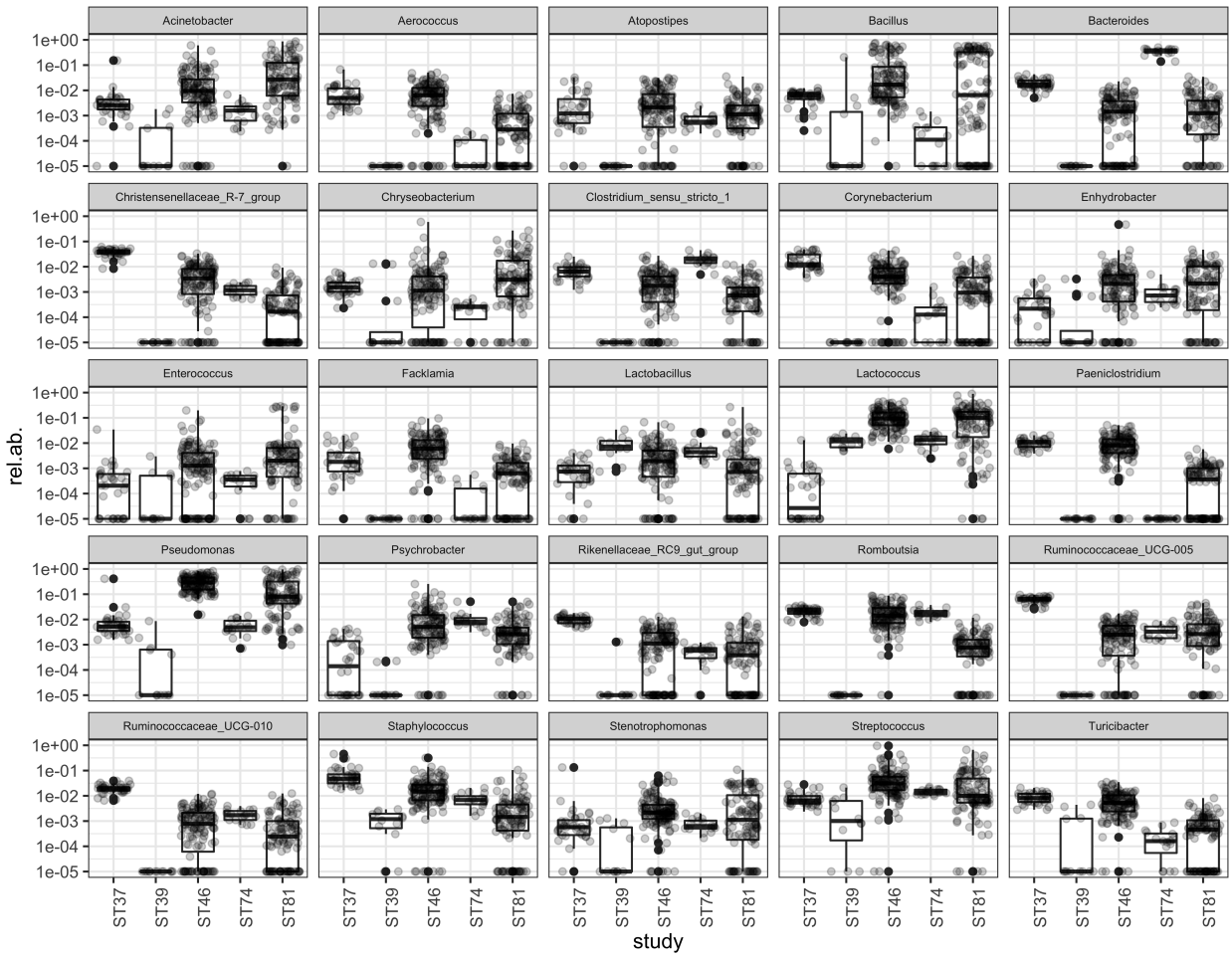
Fig. 5. Boxplots for the distribution of relative abundance for the 25 most abundant and prevalent genera in cow bulk tank milk from five studies: ST37, Doyle et al., 2017a; ST39, Frétin et al., 2018; ST81, Li et al., 2018; ST46, Skeie et al., 2019; ST74, Falardeau et al., 2019. Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019).

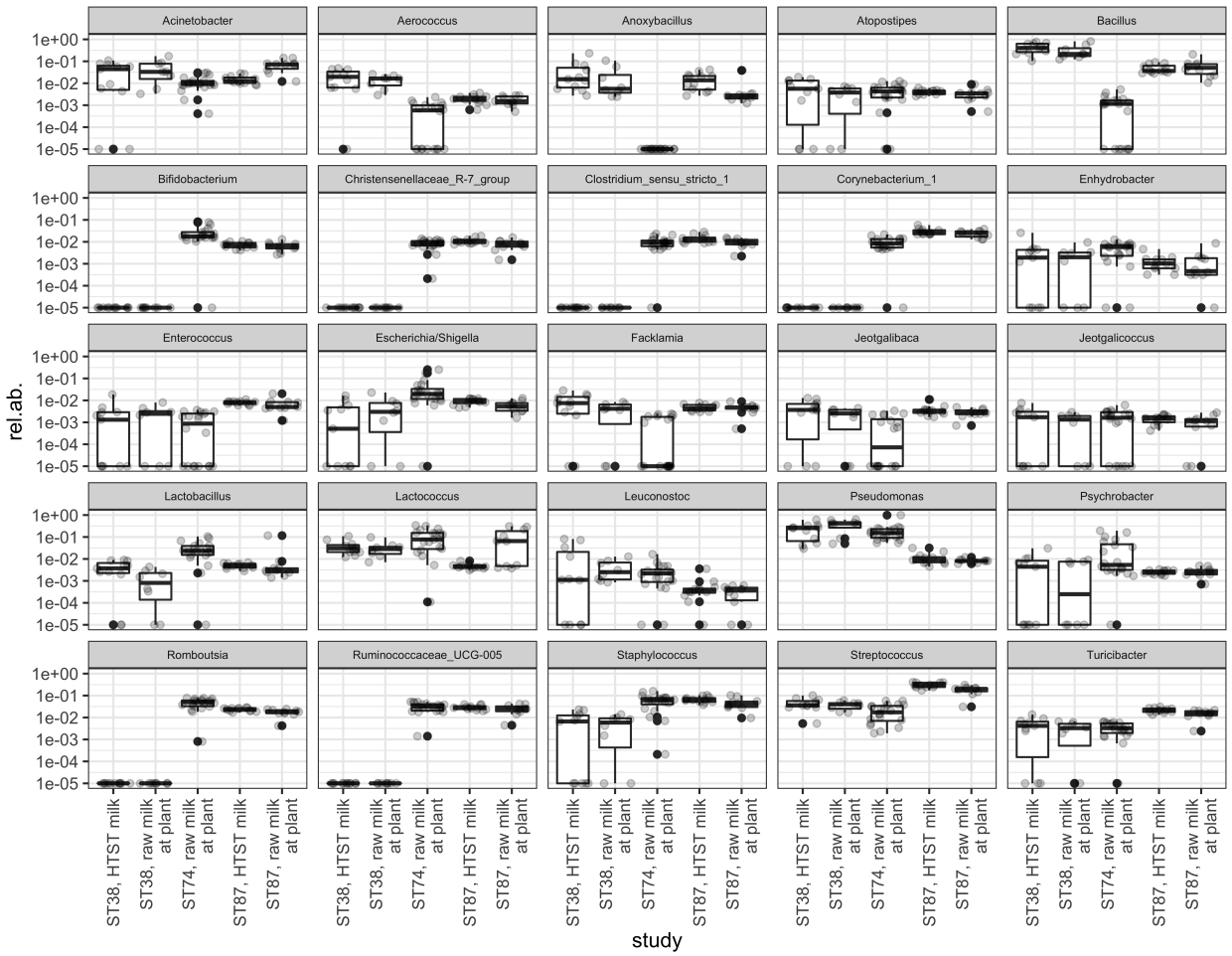
Fig. 6. Boxplots for the distribution of relative abundance for the 25 most abundant and prevalent genera in raw and HTST treated milk at the processing plant for three studies: ST38, Porcellato et al., 2018; ST74, Falardeau et al., 2019; ST87, Kable et al., 2019. Sequences were downloaded from NCBI/SRA and processed as described in Parente et al. (2019).

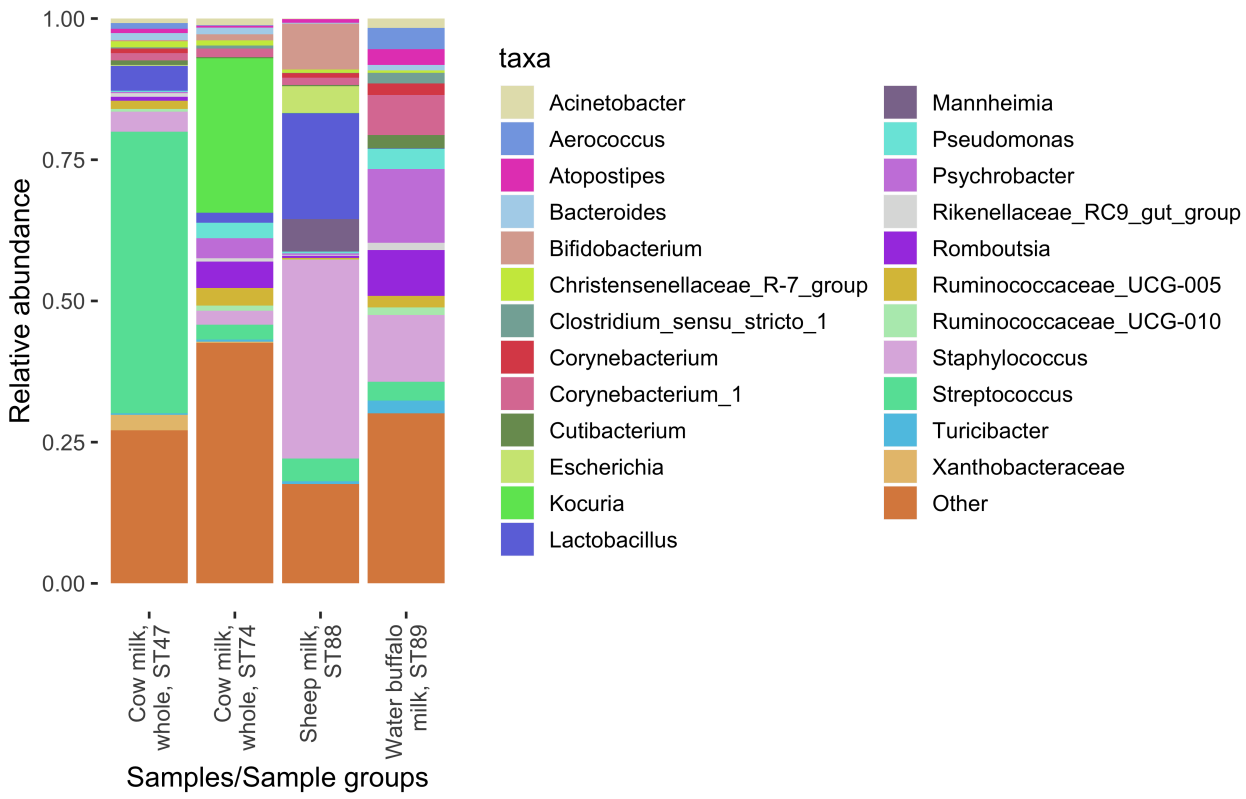

Relative abundance, Class
Frétin et al., 2018
Falardeau et al., 2019
Relative abundance
1.00
0.75
0.50
0.25
0.00
Samples/Sample groups
class
Actinobacteria (class)
Bacilli
Bacteroidia
Betaproteobacteria
Clostridia
Erysipelotrichia
Gammaproteobacteria
Acidimicrobiia
Alphaproteobacteria
Anaerolineae
Chloroflexia
Coriobacteriia
Epsilonproteobacteria
Flavobacteriia
Gemmatimonadetes
Kiritimatiellae
Longimicrobia
Mollicutes
S0134_terrestrial_group
Saccharimonadia
Subgroup_6
Tissierellia
Verrucomicrobiae


Relative abundance, top 20 taxa (based on tot seqs)
Cremonesi et al., 2018
Falardeau et al., 2019
Relative abundance
1.00
0.75
0.50
0.25
0.00
Samples/Sample groups
taxa
Aerococcus
Bacteroides
Christensenellaceae_R-7_group
Corynebacterium
Corynebacterium_1
Enhydrobacter
Eukaryota
Kocuria
Lactobacillus
Lactococcus
Pediococcus
Psychrobacter
Romboutsia
Ruminococcaceae_UCG-005
Staphylococcus
Streptococcus
Vibrionimonas
Weissella
Xanthobacteraceae
Other


Acinetobacter
Bacteroides
Caulobacter
Christensenellaceae_R-7_group
Clostridium_sensu_stricto_1
Corynebacterium_1
Ensifer
Flavobacterium
Kocuria
Lactobacillus
Lactococcus
Methylobacterium
Mogibacterium
Paracoccus
Prevotellaceae_UCG-003
Pseudomonas
Psychrobacter
Rikenellaceae_RC9_gut_group
Romboutsia
Ruminococcaceae_UCG-005
Ruminococcaceae_UCG-010
Ruminococcus_2
Staphylococcus
Streptococcus
Turicibacter
rel.ab.
1e+00
1e-01
1e-02
1e-03
1e-04
1e-05
Clover_pasture
Hay_pasture
Stored_hay
Cow_scratchers
Cow_feces
Straw_bedding
Teat_skin
Teat_milk
Tank_milk
sample group


Acinetobacter
Aerococcus
Atopostipes
Bacillus
Bacteroides
Christensenellaceae_R-7_group
Chryseobacterium
Clostridium_sensu_stricto_1
Corynebacterium
Enhydrobacter
Enterococcus
Facklamia
Lactobacillus
Lactococcus
Paeniclostridium
Pseudomonas
Psychrobacter
Rikenellaceae_RC9_gut_group
Romboutsia
Ruminococcaceae_UCG-005
Ruminococcaceae_UCG-010
Staphylococcus
Stenotrophomonas
Streptococcus
Turicibacter
rel.ab.
1e+00
1e-01
1e-02
1e-03
1e-04
1e-05
ST37
ST39
ST46
ST74
ST81
study


Acinetobacter
Aerococcus
Anoxybacillus
Atopostipes
Bacillus
Bifidobacterium
Christensenellaceae_R-7_group
Clostridium_sensu_stricto_1
Corynebacterium_1
Enhydrobacter
Enterococcus
Escherichia/Shigella
Facklamia
Jeotgalibaca
Jeotgalicoccus
Lactobacillus
Lactococcus
Leuconostoc
Pseudomonas
Psychrobacter
Romboutsia
Ruminococcaceae_UCG-005
Staphylococcus
Streptococcus
Turicibacter
rel.ab.
1e+00
1e-01
1e-02
1e-03
1e-04
1e-05
ST38, HTST milk
ST38, raw milk at plant
ST74, raw milk at plant
ST87, HTST milk
ST87, raw milk at plant
study


1.00
0.75
0.50
0.25
0.00
Relative abundance
Cow milk, whole, ST47
Cow milk, whole, ST74
Sheep milk, ST88
Water buffalo milk, ST89
Samples/Sample groups
taxa
Acinetobacter
Aerococcus
Atopostipes
Bacteroides
Bifidobacterium
Christensenellaceae_R-7_group
Clostridium_sensu_stricto_1
Corynebacterium
Corynebacterium_1
Cutibacterium
Escherichia
Kocuria
Lactobacillus
Mannheimia
Pseudomonas
Psychrobacter
Rikenellaceae_RC9_gut_group
Romboutsia
Ruminococcaceae_UCG-005
Ruminococcaceae_UCG-010
Staphylococcus
Streptococcus
Turicibacter
Xanthobacteraceae
Other