bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Revealing the microbial heritage of traditional Brazilian cheeses through metagenomics

Caroline Isabel Kothe[a]*, Nacer Mohellibi[a], Pierre Renault[a]*

[a] Université Paris-Saclay, INRAE, AgroParisTech, Micalis Institute, 78350 Jouy-en-Josas, France.

*Corresponding authors: C. I. Kothe. INRAE, Micalis Institute, Domaine de Vilvert, 78350 Jouy-en-Josas. E-mail: carolkothe@hotmail.com

P. Renault. INRAE, Micalis Institute, Domaine de Vilvert, 78350 Jouy-en-Josas. E-mail: Pierre.Renault.2@inrae.fr

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Brazilian artisanal cheeses date from the first Portuguese settlers and evolved via local factors, resulting in unique products that are now part of the patrimony and identity of different Brazilian regions. In this study, we combined several culture-independent approaches, including 16S/ITS metagenetics, assembly- and deep profiling-metagenomics to characterize the originality of the microbiota of five varieties of Brazilian artisanal cheeses from the South and Southeast regions of Brazil. Their core microbiota contained mainly lactic acid bacteria (LAB), of which *Lactococcus lactis* subsp. *lactis* was the most frequent, followed by *Streptococcus thermophilus* in the South region. Moreover, several samples from the Southeast region contained, as dominant LAB, two other food Streptococci belonging to a new species of the *salivarius* group and *S. infantarius*. Rinds of samples from the Southeast region were dominated by the halotolerant bacterium *Corynebacterium variabile* and the yeasts *Diutina catenulata* and, to a lesser extent, by *Debaryomyces hansenii* and *Kodamaea ohmeri*. Rinds from the South region contained mainly LAB due to their short ripening time, and the predominant yeast was *D. hansenii*. Phylogenomic analysis based on *L. lactis* metagenome-assembled genomes (MAGs) showed that most Brazilian strains are closely related and form a different clade from those whose genomes are available at this time, indicating that they belong to a specific group. Lastly, functional analysis showed that *S. infantarius* acquired a ~26 kb DNA fragment from *S. thermophilus* starter strains that carry the LacSZ system, allowing fast lactose assimilation, an adaptation advantage for growth in milk. Finally, our study identified several areas of concern, such as the presence of somatic cell DNA and high levels of antibiotic resistance genes in several cheese microbiota, implying that the milk used was from diseased herds. Overall, the data from this study highlight the potential value of the traditional and artisanal cheese production network in Brazil, and provide a metagenomic-based scheme to help manage this resource safely.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Keywords: microbial diversity, microbial ecology, fermented food, artisanal cheese, metagenome-assembled genomes, genomics, LAB.

1 Introduction

Beyond their historical, socio-economic and cultural aspects, artisanal fermented foods have recently received increased attention as potential sources of biological resources that can provide both technological and health benefits. Among them, traditional cheeses worldwide contain rich natural microbiota that improve the quality of cheese production in terms of flavor, texture and safety aspects (Montel et al., 2014; Walsh, Macori, Kilcawley, & Cotter, 2020). This is also true for emerging countries where dairies have more recently appeared.

There are few reports on the colonization of cheese in Brazil. Some studies have suggested that the potential origin of Brazilian dairy products traces back to the "discovery" of the country in 1581 when Portuguese settlers arrived with their cattle and passed on cheese production techniques to the local population (Borelli, Lacerda, Penido, & Rosa, 2016; Penna, Gigante, & Todorov, 2021). Although cheese production in Brazil was based on European practices, factors such as climate, farm organization, feed, breed of dairy animals, quality of raw milk and indigenous microbiota resulted in unique products that are now part of the patrimony and identity of different Brazilian regions (Kamimura, Magnani, et al., 2019; Penna et al., 2021). There are more than 30 varieties of artisanal cheeses produced in micro-regions of the Southeast (Araxá, Campo das Vertentes, Cerrado, Canastra and Serro), South (Colonial and Serrano), Center (Caipira), North (Marajó) and Northeast (Butter and Curd) regions of the country (Kamimura, Magnani, et al., 2019). Recent studies have shown that these local cheeses are potential sources of lactic acid bacteria (LAB) that can be used for the development of starter cultures (Cabral et al., 2016; Margalho et al., 2020). It could thus be

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interesting to more widely explore the biodiversity of these artisanal products whose microbiota may have evolved separately from the well-studied Western cheeses. These data would be valuable to more effectively develop them as biological resources, ensure their safety and facilitate their preservation, as well as to support the development of technical regulations and maintain their historical and cultural footprint.

Several studies have focused on the microbiological characterization of Brazilian artisanal cheeses using culture-dependent methods (Lima, Lima, Cerqueira, Ferreira, & Rosa, 2009; Luiz et al., 2016; Perin, Sardaro, Nero, Neviani, & Gatti, 2017; Pontarolo et al., 2017; Resende et al., 2011). However, these techniques do not provide a complete picture of the existing microbiota since several species are poorly cultivable or require appropriate culture media. Recently, Kamimura *et al.* (2019) (Kamimura, De Filippis, Sant'Ana, & Ercolini, 2019) conducted a pioneering study using culture-independent methods in Brazilian cheeses. Using 16S rRNA sequencing, they characterized the bacterial diversity of a wide number of artisanal cheeses from different regions of Brazil at the genus level. They mainly identified LAB of the genera *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Leuconostoc* and *Streptococcus*, as well as some probable contaminants such as *Enterobacteriaceae* and *Staphylococcus*. However, these studies seem to focus on the core microbiota, leaving the rind microbiota unknown, whereas it may harbor halotolerant and halophilic bacteria and yeasts, which are key actors in cheese ecology and technology (Banjara, Suhr, & Hallen-Adams, 2015; E. Dugat-Bony et al., 2016; Kothe et al., 2021; Wolfe, Button, Santarelli, & Dutton, 2014). Finally, the 16S gene analysis does not provide functional data or detailed taxonomic information that make it possible to uncover the original characteristics of their microbiota.

Consequently, there are still many unanswered questions concerning these artisanal Brazilian cheeses that would require deeper investigation, including species- and even strain-level analysis, for yeasts as well, and that target the specific microbiota of the rind and the

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

core. While metagenetics allows a rapid identification at the genus level, other meta-omic approaches such as shotgun metagenomics may provide a more accurate microbiota level of analysis. In this study, we focused our analyses on cheeses from two Brazilian regions (South and Southeast) and combined 16S rRNA and ITS metagenetic analysis with shotgun metagenomic analysis in order to reveal their specificities at the level of species composition and strain origin, and to highlight functional features of interest. These data could be a benefit to market expansion, ensuring the protection of historical and sanitary aspects, in addition to supporting the development of technical regulations, legislation and standardization parameters for the different types of traditional Brazilian cheeses.

2 Methodology

2.1 Sample collection and DNA extraction

A total of 23 artisanal cheeses were obtained from Southeast and South Brazil from five varieties: Araxá, Canastra, Serro, Colonial and Serrano. Samples were collected from local producers, artisan markets and fairs, and sent by post to France. The rinds were separated from the cores using sterile knives, and both fractions were analyzed to obtain a more detailed overview of the microbial diversity of those cheeses. The samples were diluted 1:1 (w/v) in guanidinium thiocyanate 4M solution (Sigma-Aldrich, USA) with Tris-HCl 0.1M, and mixed in an Ultra Turrax T25 (Labortechnik) at 8,000 rpm for 2 min. A 10% N-lauryl sarcosine solution (Sigma-Aldrich, USA) was added to the mixture, vortexed and centrifuged at 4°C and 14,000 rpm for 30 min. The fat and supernatant were eliminated and the pellet was used for DNA extraction using the protocol described by Almeida *et al.* (2014) (Almeida et al., 2014). DNA quality was visualized on 0.8% agarose gel and the quantification was measured with a Qubit 2.0 fluorometer (Life Technologies) using a Qubit dsDNA HS (High Sensitivity) Assay Kit.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

112

113 *2.2 Metagenetic analyses*

114 Bacterial diversity was analyzed by sequencing the amplified region V3-V4 of the 16S
115 rRNA gene using primers V3F (5'-ACGGRAGCWGCAGT-3') and V4R (5'-
116 TACCAGGGTATCTAATCCT-3'). Additionally, fungal diversity was evaluated in rinds
117 using ITS3 (5'-GCATCGATGAAGAACGCAGC-3') and ITS4 (5'-
118 TCCTCCGCTTWTGWTWTGC-3') primers. The PCR was performed with MTP Taq DNA
119 Polymerase (Sigma-Aldrich, USA), and the cycling conditions were: 94°C for 1 min,
120 followed by 30 cycles of amplification at 94°C for 1 min, 65°C for 1 min, and 72°C for 1 min,
121 with a final extension step of 10 min at 72°C. The sequencing was performed with a V3
122 Illumina MiSeq kit, as described in Poirier *et al.* (2018) (Poirier et al., 2018).

123 The quality of the raw data was evaluated with FastQC (Wingett & Andrews, 2018)
124 and the sequences were imported into the FROGS pipeline (Escudié et al., 2018) to obtain the
125 Operational Taxonomic Units (OTUs). The sequences were filtered by length (150–500 bp)
126 and then pooled into OTUs with SWARM (Mahe, Rognes, Quince, de Vargas, & Dunthorn,
127 2014) with the distance parameter of 3. Chimeras were removed with VSEARCH (Rognes,
128 Flouri, Nichols, Quince, & Mahe, 2016) and OTUs with at least 0.01% in the whole dataset
129 were retained. The OTUs were affiliated with SILVA 132 SSU databases (Quast et al., 2013)
130 for bacteria and UNITE 8.2 for fungi (https://unite.ut.ee/). Alpha-diversity and beta-diversity
131 analyses were performed in R Studio v.3.6.1 using the phyloseq and ggplot2 packages
132 (v1.30.0) (McMurdie & Holmes, 2013; Poirier et al., 2018).

133

134 *2.3 Taxonomic composition by shotgun metagenomics*

135 The DNA of 15 cheese samples was sequenced using Illumina HiSeq2500 technology
136 at GATC-Biotech (Konstanz, Germany), which yielded between six and eight million paired-

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

end reads of 150-nucleotide length. Metagenomic reads corresponding to the *Bos taurus* genome were filtered with Bowtie2 (Langmead & Salzberg, 2012) and visualized with Samtools flagstat (H. Li et al., 2009). From the generated fastq files, we first estimated microbial composition by mapping the samples reads against the representative clade-specific marker catalogue contained in the MetaPhlAn tool, v.3.0.4 (Truong et al., 2015).

Additionally, we performed taxonomic profiling using an assembly-based marker gene analysis, which allows non-supervised binning of metagenomes. To do this, the reads were trimmed for quality and *de novo* assembly was performed using metaSPAdes, v.3.9 (Bankevich et al., 2012). Genes were then predicted using Prodigal (v.2.6.3) and marker genes were extracted using fetchMG, v.1.0 (Ciccarelli et al., 2006; Sunagawa et al., 2013). We chose to perform our taxonomic assignations by using the *ychF* marker gene, whose closest homologue was assigned by a blastp search on all the available sequences from the NCBI protein database. Summary species composition plots were created in R (v.3.6.1) using the ggplot2 package, v.3.3.2. Finally, phylogenetic analyses were performed with the *ychF* proteins identified using ClustalX 2.1 (Thompson, Gibson, Plewniak, Jeanmougin, & Higgins, 1997) and MEGA7 (Kumar, Stecher, & Tamura, 2016) with the Neighbor-Joining method (Saitou & Nei, 1987) and 1,000 bootstrap replicates (Felsenstein, 1985).

2.4 Metagenome-assembled genome (MAG) analyses

Genome binning was performed using MetaBAT2-2.12.1 (Kang, Froula, Egan, & Wang, 2015), with a minimum contig size of 1,500 nucleotides and the default settings. The quality of the resulting prokaryotic bins was assessed with CheckM (Parks, Imelfort, Skennerton, Hugenholtz, & Tyson, 2015), and MAGs < 80% completeness and/or > 10% contamination were excluded. For eukaryotes, we considered those bins that aligned with > 60% of their sequence length to fungal reference genomes from GenBank as MAGs of quality.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

In addition, MAG assemblies were performed for Streptococcal species by additional binning, such as blastn analysis with reference genomes of related species. Contigs were then filtered by percentage of the length covered, identity and coverage levels. Finally, manual curing of questionable contigs (new, low coverage of homology on references, etc.) were performed by blastn and blastx analysis against nr/nt and nr NCBI databases, respectively.

2.5 Phylogenomic and functional analyses

The Automatic Multi-Locus Species Tree web server (https://automlst.ziemertlab.com/) (Alanjary, Steinke, & Ziemert, 2019) was used to determine closely related genomes based on core gene alignments of the recovered MAGs. The closest species were inferred based on the percentage of average nucleotide identity (ANI) calculated using FastANI, v.1.31 (Jain, Rodriguez-R, Phillippy, Konstantinidis, & Aluru, 2018).

kSNP (v.3.0) was used to perform the phylogenomic analysis for the *Lactococcus lactis* species (Gardner, Slezak, & Hall, 2015), with the maximum likelihood option and kmer size of 21. A total of 225 genomes of the *Lactococcus lactis* species were downloaded from Genbank and combined with our MAG to compute this analysis.

The search for antibiotic resistance (ABR) genes was performed by read mapping against the CARD database (Jia et al., 2017; McArthur et al., 2013) using the PATRIC web server (Antonopoulos et al., 2019). The presence of particular genes in reference genomes, such as virulence factors, ABR and genes of technological interest, was determined using the FoodMicrobiome tool (https://migale.jouy.inra.fr/foodMicrobiome/; (Kothe et al., 2021). This tool performs read mapping on given reference genomes and provides read counts for each annotated gene. Additionally, this tool was used to detect, with high sensitivity and reliability,

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the subdominant populations of given species by analyzing the distribution of metagenomic reads on reference genomes of these species.

2.6 Data availability

Raw sequences of amplification of 16S rRNA and ITS genes and raw metagenomic reads were deposited on the European Nucleotide Archive (ENA) under the BioProject ID PRJNA693797. The MAGs are available in: https://data.mendeley.com/datasets/4w8w9mkfjy/1

3 Results

3.1 Sample descriptions

A total of 23 artisanal cheeses were obtained from the Southeast and South regions of Brazil from five varieties: Araxá, Canastra, Serro, Colonial and Serrano (**Fig. 1**). The cheeses were produced with raw milk, except C-19, and the samples were collected from local producers, artisan markets and fairs (**Table S1**). These five cheese varieties are defined as semi-fat to fat (25-59.9% of fat) and semi-hard cheeses (humidity 36-45.9%) (de Medeiros Carvalho, de Fariña, Strongin, Ferreira, & Lindner, 2019; MAPA, 2020; RS, 2016). The differences in the production process and characteristics of each type of cheese are briefly described in **Supplementary Data 1**.

3.2 Taxonomic diversity of Brazilian cheeses using amplicon sequencing approaches

The bacterial diversity present in the core and the rind of the 23 Brazilian cheeses was assessed using amplicon sequencing targeting the 16S rRNA gene. The sequences were clustered in 100 bacterial OTUs whose taxonomic assignment was possible up to the species level in the majority of the cases (**Table S2**). The C-02 sample was discarded due to its low

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

read depth. For the majority of other samples, the rarefaction curves reached the saturation plateau, indicating that the sequencing depth was sufficiently recovered (**Fig. S1)**.

The alpha-diversity analysis for core samples showed greater observed richness of bacterial species ($p < 0.01$) in the South samples compared to the Southeast samples (**Fig. 2A)**. However, the bacterial evenness measured by Shannon and inverse Simpson indices showed that the diversity of species in both regions was quite similar ($p > 0.05$). The Principal Coordinates Analysis (PCoA) and clustering showed the diversity among the samples, and we observed that the majority of cheese cores from the two regions shared a common bacterial microbiota (**Fig. 2A**). Nevertheless, the samples C-08 and C-10 from the Araxá micro-region contain microbiota that is different from the other samples but similar between them, and C-22 belongs to a particular group. Indeed, the taxonomic composition showed that LAB constitute the major part of core sample microbiota with *Lactococcus lactis* and *Streptococcus thermophilus-salivarius* as the dominant species (over 50% of the reads) in 14 and five samples, respectively, while a *Streptococcus* from the *equinus-lutetiensis-infantarius* complex is dominant in the C-08 and C-10 samples (**Fig. 3A**). Additionally, in the core of the C-22 sample, we observed the predominance (~79.5%) of the enterobacteria *Kluyvera cryocrescens*, while *S. thermophilus* was present at only low levels (4.6%). Finally, several other LAB species were less abundant, such as *Levilactobacillus brevis*, *Lacticaseibacillus casei*, *Lb. delbrueckii, Lactiplantibacillus plantarum, L. piscium, Lc. mesenteroides,* as well as undefined species of *Streptococcus* and *Lactococus* (**Fig. 3A**).

Concerning bacterial richness and evenness of the rind samples, no statistical differences were observed based on the observed species and Shannon and Simpson diversity indices ($p > 0.05$, **Fig. 2B**). The PCoA and hierarchical clustering mainly showed two different groups of microbiota: one composed only of samples from the Southeast and another with cheeses from both regions. The taxonomic composition showed that the majority of

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cheese rinds from the South displayed the same dominant species as their core (seven out of nine), whereas those from the Southeast region are dominated by *Corynebacterium variabile* (eight out of 14) (**Fig. 3B**). The two divergent samples from the South, C-R20 and C-R22, were dominated by high levels of *Celerinatantimonas sp.* and *Cobetia marina*, respectively. Finally, four of the divergent rind samples from the Southeast had a composition closer to that of the core (C-R01, C-R05, C-R10 and C-R12), and the last two, C-R06 and C-R14, were dominated by *Enterobacteriaceae* species and an undefined *Brevibacterium* species, respectively (**Fig. 3B**).

Furthermore, *S. aureus* and *E. coli* species, which may have pathogenic strains, were detected in more than 45% of core samples and more than 35% of rinds at levels > 0.01%. The highest levels were detected in C-14 (which represented 2.86% of *S. aureus*) and C-20 samples (with 3.45% of *E. coli*) (**Fig. 3A**).

Concerning eukaryotic diversity, ITS amplification was successful for 17 rind samples, and the sequences were grouped into 34 yeast/fungal OTUs **(Table S2**). The alpha diversity analyses did not show statistical differences between the South and Southeast regions ($p > 0.05$). The beta-diversity analysis showed two groups, the first one consisting mainly of samples from the South, except C-R05, and the second one of samples from the Southeast (SE), except C-R16 (**Fig. 2C**). Taxonomical analysis of the ITS indicated the predominance of different fungal/yeast species between South and SE cheeses (**Fig. 3C**). Using a threshold of > 5% of read abundance, *Debaryomyces hansenii* was present in all South cheese rinds (n=6), but only in five from SE cheeses (45%). Conversely, the dominant yeast in SE cheeses was *Diutina catenulata*, which was present in nine samples at levels > 5%, and in only one South cheese. Moreover, several additional species were detected mainly in SE cheeses, such as *Kodamaea ohmeri*, *G. candidum*, *Trichosporon sp.* and *Moniliella sp.*,

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and in South cheeses, such as *Penicillium commune*, *Clavispora lusitaniae* and *Candida zeynaloides*.

3.3 Refined microbiota by metagenomics

In order to more precisely analyze the microbiota from our samples, we selected six and nine samples of cores and rinds, respectively, to perform shotgun metagenomic analyses. Such analyses allow a taxonomic identification up to strain level and the simultaneous assessment of the level of reads corresponding to bacteria and fungi. Moreover, it makes it possible to detect eventual virulence factors in pathogenic species, as well as antibiotic resistance genes and genes of technological interest.

3.3.1 Species level analyses

Metagenomic samples often contain animal reads, which may significantly bias the analysis when they are in great abundance. In the present samples, we detected *Bos taurus* reads at levels ranging from 0.1 to 76% (**Table S3**). We observed that cow reads were notably higher in the core samples, especially C-15, C-18 and C-20 from the South region, where they accounted for more than 30%.

The taxonomical composition of the 15 cheese metagenomes was first determined using MetaPhlAn, which yielded 128 bacterial and one fungal species (**Table S4**). Species composition obtained by this analysis is relatively in accordance with those obtained with 16S rRNA amplicon for species present at abundance levels > 1%. Interestingly, it resolved several taxonomical assignations such as that of *Streptococcus equinus-lutetiensis-infantarius*, which becomes *S. infantarius* and appears to be the dominant LAB in two Araxá cheeses (**Fig. 3A**). It also detected several species present at significant levels such as *Corynebacterium variabile*, *Staphylococcus saprophyticus*, *Brachybacterium alimentarium*, *Acidipropionibacterium acidipropionici*, *Brevibacterium linens* and *S. parauberis*.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Conversely, it did not detect two possibly abundant bacteria, an uncharacterized *Brevibacterium* sp. and *Providencia rettgeri* in C-R14 and C-R06, respectively. Moreover, it detected only one yeast (*Candida parapsilosis*) in our metagenome set, whereas ITS amplicon analysis showed large fungal diversity, indicating a probable lack of corresponding genome references in the database currently available from MetaPhlAn.

In order to characterize microbiota independently of fixed references, we applied a marker gene analysis from the assembled metagenomes, which yielded 57 bacterial and seven fungal species (**Table S4**); the overall composition is shown in **Fig. 4**. Although the results are consistent with the previous analysis made by read profiling, some differences merit highlighting. In particular, 14 species whose relative abundance ranges from 3 to 37% were detected uniquely by the marker analysis, probably due to their absence in the currently implemented MetaPhlAn database. Among bacteria of potential technological interest, it revealed *Lacticaseibacillus paracasei* and two potential new species of *Brevibacterium* and one of *Microbacterium*. Moreover, this analysis suggested that while the *S. thermophilus-salivarius* populations detected in the South samples belong to *S. thermophilus*, those detected in the Southeast samples are instead related to another species of the *salivarius* group (**Table S4**, **Fig. S2**). Indeed, their sequence of the well-conserved *ychF* gene displays more than 5% nucleotidic divergence with those of the other species of the *S. salivarius* group, suggesting that they belong to a new species. Finally, we detected seven species of yeasts, *Debaryomyces hansenii*, *Diutina catenulata*, *Kodamaea ohmeri*, *Geotrichum candidum*, *Kluyveromyces lactis* and two other Saccharomycetales. We also identified the proportions of prokaryotes and eukaryotes in the samples; yeasts appear to be absent in the core samples, and range from 0 to 27% in the rinds.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3.3.2 MAGs and strain-level analyses

A total of 54 prokaryotic and 10 eukaryotic metagenome-assembled genomes (MAGs) of high quality were reconstituted from the metagenome assemblies (**Table S5**). They include 27 prokaryotic and four eukaryotic species. The majority of bacterial MAGs correspond to *L. lactis* subsp. *lactis* (10) and *C. variabile* (8). Moreover, eight MAGs represent six potential novel species since they have less than 95% average nucleotide identity (ANI) with reference genomes (**Table S5, Fig. S3**). They also include species of the genera *Brevibacterium* (2), *Streptococcus*, *Corynebacterium*, *Lactobacillus* and *Microbacterium.* Additionally, one MAG that belongs to the *Micrococcaceae* family could not be identified at the genus level since it does not share ANI > 80% with any reference genome.

Further detailed analysis was performed with the Lactococcal and Streptococcal MAGs, which correspond to key species of starter bacteria. Phylogenomic analysis of the *L. lactis* MAGs was performed with 215 reference genomes of this species, representing strains from its two main subspecies (*lactis* and *cremoris*) and isolated from different environments (dairy, vegetable, animal, etc.) (**Fig. S4**). This showed that nine of the Brazilian cheese MAGs form a homogeneous group whose most closely related strains are *L. lactis* subsp. *lactis* isolated from various traditional dairy products around the world. Although several strains of this group are referenced at NCBI as belonging to the diacetylactis biovariant, they all lack the citrate lyase complex (*mae-citRCDEFXG*) involved in the production of diacetyl from citrate, and they are phylogenetically distant from the diacetylactis group. Additionally, the *L. lactis* MAG originating from the C-R02 sample is clustered with another *L. lactis* subsp. *lactis* group containing mostly industrial dairy starter strains. A search for the genes encoding the Lac-PTS system (ten genes) and PrtP that allow rapid lactose assimilation and casein breakdown in Lactococci, respectively, were found in all metagenomes at levels similar to or slightly higher than chromosomal Lactococcal genes (**Table S6**).

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Concerning *Streptococcus*, six MAGs were built, and ANI analyses showed that they may belong to three species (**Table S5**). A first group of three MAGs belong to *S. thermophilus* species (ANI > 98% with the genome of reference strains of this species), and detailed phylogenomic analysis clearly links them to strains isolated from cheese and milk in Europe. A second group of two MAGs (C-03 and C-13) belongs to the *salivarius* group, forming a cluster clearly separated from *S. thermophilus* and related to but different from *S. salivarius* (ANI~94%, **Fig. S3A**). This result indicates that they probably belong to a new species. Finally, MAG-C-10, belonging to the species *S. infantarius*, was compared to the genomes of the type strain isolated from feces and the CJ18 food strain. The three genomes share over 98.5% ANI and 1,576 orthologous proteins, confirming their close relationship. Moreover, MAG-C-10 displays a contig of 26.1 kb, sharing 99.6% identity over 25,870 nucleotides with *S. thermophilus* ATCC 19258. This region contains, in particular, the *lacZ* and *lacS* genes that allow the rapid assimilation of lactose of *S. thermophilus*. It is flanked at the 5' end by genes belonging to the IS3 mobile element and at its 3' end by a 55-nucleotide signature flanking the IS1182 family transposase in Streptococci. Complementary analysis by read mapping showed that this region is covered at the same average level as the 1,516 genes defined above as being common to the three *S. infantarius* strains.

Concerning species growing on the rinds, eight MAGs referring to *C. variabile* were recovered. They display a very close relatedness (ANI > 99%) with the only two genomes deposited at the NCBI (DSM 44702 and Mu292 strains isolated from smear-ripened cheeses in Ireland (Schröder, Maus, Trost, & Tauch, 2011) and France (E. Dugat-Bony et al., 2016), respectively. Furthermore, MAGs corresponding to known cheese halotolerant bacteria such as *S. saprophyticus* (three MAGs), *Bavaricoccus seileri* (4), *Brachybacterium alimentarium* (2) and several *Brevibacterium* species were recovered.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

359 The ten fungal MAG reconstituted in this study could be assigned to four species:
360 *Debaryomyces hansenii*, *Diutina catenulata*, *Geotrichum candidum* and *Kodamaeae ohmeri*
361 (**Table S5**). Genomic comparison with the reference strains available in the NCBI database
362 shows that *G. candidum* and *K. ohmeri* MAGs display ANI > 99% and alignment to reference
363 > 80% to CLIB918 and 148 strains isolated from Pont l'Evêque cheese in Normandy (France)
364 and from honeybee gut in the USA, respectively. Finally, *D. catenulata* and *D. hansenii*
365 MAGs present ANI = ~97.5% and alignment to reference of 60-90% (depending of the MAG
366 quality) and 95%, respectively.
367
368 3.3.3 Detection of pathogens, virulence factors and antibiotic resistances
369 Strains of 12 pathogen species commonly found in dairy products were searched for
370 by deep profiling analysis in the Brazilian cheese metagenomes (**Table S7**). *S. aureus* was
371 detected at significant levels in the Canastra cheese rind sample C-R02 (0.5% of reads
372 mapped), with an average gene detection rate of 14.6 reads/kb. Since this level of coverage
373 allows a reliable detection of genes in metagenomes, we searched for the presence of 27
374 staphylococcal enterotoxins responsible for foodborne outbreaks, and only enterotoxin A (*sea*,
375 *entA*) and X (*selx*) were found. Finally, among the other species, only *Escherichia coli* was
376 detected in two core samples from the South (C-15 and C-20), at average levels of 1.3 and 2.5
377 reads/kb, which is below the threshold required to perform a reliable detection of virulence
378 genes.
379 In addition, a mapping-based approach against a comprehensive collection of
380 antibiotic resistance genes was performed. A total of 30 ABR genes were identified,
381 belonging to nine different classes of antibiotics (**Fig. 5** and **Table S6**). The presence of ABR
382 genes was detected in all samples, except in three (C-R09, C-03 and C-R04). The C-R02
383 cheese sample displayed the highest number (nine ABR genes). Overall, the most abundantly

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

384 detected antibiotic class was tetracycline, comprising ten different genes of which the *tetK*
385 and *tetS* genes were present in five and four samples, respectively. While the *tetK* gene was
386 detected (exclusively) in cheese rind samples containing Staphyloccocal species (*S. aureus, S.*
387 *saprophyticus* and *S. xylosus*), the *tetS* gene was detected in four samples where streptococcal
388 species were dominant. In particular, *tetS* displays a coverage level similar to *S. thermophilus*
389 genes in C-15, C-18 and C-20, whereas it is 70 times lower than those of MAG-C-13
390 belonging to *S. salivarius*-like in C-13. Finally, the gene *qnrD1*, detected in the C-R06 and C-
391 R13 samples, was found in contigs with coverage similar to those of *Serratia* and *Providencia*
392 present in both samples, respectively.
393
394 **4 Discussion**
395 Artisanal and traditional cheese production has received increased attention as sources
396 of biological resources potentially capable of contributing to technological and health
397 benefits. However, there is an increasing trend, in artisanal production as well, of using starter
398 strains to ensure a rapid and safe lactic fermentation, and, often, the addition of secondary
399 microbiota to more effectively control ripening (Bintsis, 2018; García-Díez & Saraiva, 2021;
400 Vandera, Kakouri, Koukkou, & Samelis, 2019). On the other hand, proper selection of native
401 strains may offer the opportunity to preserve the typicity of the cheese, while offering an
402 excellent level of safety (Gaglio, Todaro, & Settanni, 2021). In order to optimize such a
403 process, cheeses used as original material for strain selection should be carefully selected in
404 order to maximize those with original microbiota and to absolutely avoid those “spoiled” by
405 industrial starter strains, making it difficult to isolate new strains with original properties. In
406 this respect, regions where legislation excludes the use of starter cultures or additives provide
407 the opportunity for the development of autochthonous microorganisms. In this study, we
408 decided to develop a method based on metagenomics to explore the biodiversity of cheese
409 microbiota in order to characterize their originality and potential value as a bio-resource. We

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

410 studied 23 cheese samples from two regions, South and Southeast Brazil, combining a
411 metagenetic approach to obtain a global view of their microbiota at the genus/family level
412 with metagenomic approaches to access functional and phylogenetic information (**Fig. S5**).
413
414 *4.1 Microbial patrimony of artisanal Brazilian cheeses*
415 As expected, analysis of the cheese cores showed the dominance of LAB such as *L.*
416 *lactis* and *S. thermophilus*, as already described at the genus level by Kamimura *et al.* (2019)
417 (Kamimura, De Filippis, et al., 2019), except in one sample, which was a production failure,
418 leading to the predominance of an enterobacterium, *Kluyvera cryocrescens*. The analysis of
419 the *L. lactis* MAGs indicated that *L. lactis* present in Brazilian cheese form a particular
420 phylogenomic group within the *lactis* subspecies that differentiate them from the traditional
421 starters (**Fig. S4**). Further analysis such as strain isolation and genome sequencing will be
422 valuable in providing more information about specific genes and should pave the way for
423 technological studies focused on the use of these strains, as biopreservatives, for example, as
424 was shown with the *L. lactis* strain QMF11 isolated from a Brazilian cheese, which has strong
425 anti-listerial activity (Costa et al., 2018).
426 Interestingly, our study also revealed the presence of two original streptococcal
427 species as the main LABs in several samples. Firstly, *S. infantarius* was found in cheeses
428 from the two regions, and it was even dominant in the core of two SE cheese samples. *S.*
429 *infantarius* has already been isolated from Brazilian cheeses (Brito et al., 2020; Medeiros et
430 al., 2016) and it is the main LAB isolated from traditional fermented camel and cow milk in
431 East and West Africa, respectively (Jans, Kaindi, et al., 2013). In these countries, the strains
432 isolated from dairy products were shown to contain *lacZ-lacS* genes that share high sequence
433 identity with those of *S. thermophilus*, which provide them with the ability to rapidly
434 assimilate lactose from milk (Jans, Follador, et al., 2013). Further analysis of these sequences

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

showed that East and West African strains independently acquired their *lacZ-lacS* genes from different donor species by horizontal gene transfer (HGT) (Almeida et al., 2014). Our data shows that the Brazilian dairy strains of *S. infantarius* also acquired these genes by HGT from *S. thermophilus* on a large DNA fragment (25,870 nucleotides) flanked by two mobile elements. Features such as the very high identity (99.6%) over the whole length of the fragment and the absence of a recombination event to remove the unnecessary genes indicate that this transfer occurred independently from those described in Africa, and very recently. Secondly, while metagenetic analysis highlighted the presence of Streptococci of the *thermophilus-vestibularis-salivarius* group, which are usually referred to as *S. thermophilus* in dairy products, the metagenomic analysis revealed that cheeses from the Southeast region contain a potentially novel species closely related to *S. salivarius* (**Fig. 4**; **Fig S3A**). The presence in food of these two streptococcal species that probably originated from the human gut, raises the question of their safety in the event of regular consumption. While Streptococci of the *salivarius* group are generally human commensal bacteria that dominate the oral microbiota (Carlsson, Grahnén, Jonsson, & Wikner, 1970; McCarthy, Snyder, & Parker, 1965) and may be used as probiotics (Wescombe, Hale, Heng, & Tagg, 2012), the role of *S. infantarius* in human microbiota is less well established. Nevertheless, its consumption at high levels in African fermented milk indicates that this bacterium is safe for humans (Jans et al., 2017), and a study has shown that it is not associated with colorectal cancer as is its relative in the SBSEC group, *S. gallolyticus* (Jans, Meile, Lacroix, & Stevens, 2015). Overall, these results indicate that cheese-making practices without the use of LAB starters favored the emergence of two novel streptococcal food strains whose properties should be further studied to ensure their safety and to reveal potentially interesting technological properties.

Concerning the cheese rinds, cheeses from the South region, which are not subject to long ripening, mainly contain core bacterial species (**Fig. 3B**), whereas cheeses from the

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Southeast region are generally dominated by *C. variabile*. Although less prevalent than *C. casei* (Kothe et al., 2021), this halotolerant species is well known for being part of the complex microbiota that develops on the surface of ripened cheeses (Bertuzzi et al., 2018; E. Dugat-Bony et al., 2016) and is also used as a ripening adjunct in several cheese production processes (Delbes, Monnet, & Irlinger, 2015). The *C. variabile*-type strain possesses specific genes associated with metabolic functions involved in the technology and adaptation to cheese habitats (Schröder et al., 2011). As shown by our MAG analysis, *C. variabile* present in SE Brazilian cheeses is closely related to strains isolated from smear ripened-cheeses in Ireland (Schröder et al., 2011) and in France (E. Dugat-Bony et al., 2016). Brazilian cheeses might thus be an interesting source to recover new strains for cheese technology. Moreover, *S. saprophyticus*, a coagulase-negative staphylococcus, was found to be the second most frequent and abundant halotolerant bacteria. This species is frequently detected on the surface of smear-ripened cheese and other fermented foods (Coton et al., 2010; Hammer, Jordan, Jacobs, & Klempt, 2019). While *S. equorum* appears to be the most frequent staphylococcal species in Western cheeses (Kothe et al., 2021), our analysis and an earlier study (Casaes Nunes, Pires de Souza, Pereira, Del Aguila, & Flosi Paschoalin, 2016) indicate that *S. saprophyticus* is more prevalent in Brazilian cheeses. Finally, several less frequent halotolerant species from the genera *Bavaricoccus* and *Brevibacterium* are commonly found in SE cheese, whereas more halophilic species belonging to genera such as *Psychrobacter* and *Halomonas* appear to be scarce compared to Western cheeses (Kothe et al., 2021).

Regarding eukaryotes, overall, we observed that the South samples are dominated by *Debaryomyces hansenii*, while Southeast cheeses are more diverse, notably with the presence of *Diutina catenulata, Kodamae ohmeri, Trichosporon sp.* and *Moniliella sp.* (**Fig. 3C**). Although the presence of these species has been reported worldwide in dairy products (Banjara et al., 2015; E. Dugat-Bony et al., 2016; Irlinger, Layec, Hélinck, & Dugat-Bony,

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2015; Wolfe et al., 2014), only *G. candidum* and *D. hansenii* species are known to play a role in cheese ripening (Irlinger et al., 2015; Irlinger & Monnet, 2021; Perkins et al., 2020; Pham, Landaud, Lieben, Bonnarme, & Monnet, 2019). *G. candidum*, in particular, has been found in significant amounts only in one cheese where it is added by the producer, suggesting that this yeast is not naturally of major importance in these artisanal products. Moreover, *Kodamaea ohmeri* and *Diutina catenulata* are poorly studied at the genomic level and no cheese strain genomes of these species are available at this time. The present MAGs thus provide a first insight into their relatedness to environmental yeast strains.

4.2 Potential safety concerns

In this study, we determined - through metagenomic reads - that several samples, especially C-13, C-15 and C-18, presented more than 30% of the reads mapped on the *Bos taurus* genome (**Table S3**). Almeida *et al.* (2014) also described similar findings in a blue-veined cheese (Almeida et al., 2014), and the presence at high levels of animal reads in dairy metagenomes could be associated with the presence of milk somatic cells, which are markers of mastitis in dairy cattle, indicating herd health problems (Moradi, Omer, Razavi, Valipour, & Guimarães, 2021; Petzer, Karzis, Donkin, Webb, & Etter, 2017). This may be due to the fact that in artisanal productions, especially in small farms, milk from an animal with mastitis may not be eliminated from the production chain because of the significant financial loss that this may represent.

Regarding microbiological risks, we detected Staphylococci in a number of samples, relatively in accordance with the previous large-scale study of microbial diversity in Brazilian cheeses (Kamimura, De Filippis, et al., 2019). Since this study was made at the genus level, it left open the possibility of a high prevalence in Brazilian cheeses of *S. aureus*, a common pathogen in dairy industries worldwide (Cretenet, Even, & Le Loir, 2011), including in Brazil

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Dittmann et al., 2017). We were able to show with a high degree of confidence that most species belong to coagulase-negative species such as *S. saprophyticus*, whereas a significant amount of *S. aureus* was detected in only one sample. However, in this sample, we identified the presence of the heat-stable enterotoxin A (*sea*), which is associated with illness, accounting for 77.8% of all staphylococcal foodborne disease outbreaks (Argudín, Mendoza, & Rodicio, 2010; Balaban & Rasooly, 2000; Kadariya, Smith, & Thapaliya, 2014), indicating a potential health risk to consumers.

Furthermore, we detected the presence of several transmissible antibiotic resistance genes that are considered to be a growing threat to public health, both in hospital and food industry environments (Y. Li et al., 2020; Yadav & Kapley, 2021). These notably include *tetS* and *tetK* genes, which are probably associated with the presence of streptococcal and staphylococcal species, respectively. Tetracycline resistance has been largely reported in fermented foods, for example, in cheese ripening (Ana Belén Flórez, Delgado, & Mayo, 2005; Ana B. Flórez, Vázquez, & Mayo, 2017; X. Li, Li, Alvarez, Harper, & Wang, 2011) and, more specifically, the occurrence of the *tetS* gene has been associated with *S. thermophilus* in cheeses (Ge et al., 2007; Rizzotti, La Gioia, Dellaglio, & Torriani, 2009), whereas that of *tetK* has been associated with *S. aureus* in dairy products (Jamali, Paydar, Radmehr, Ismail, & Dadrasnia, 2015) and different species of coagulase-negative staphylococci in salami samples (Rebecchi, Pisacane, Callegari, Puglisi, & Morelli, 2015). The presence of antibiotic resistance genes at high levels is often associated with the use of milk from animals treated with antibiotics (Tóth et al., 2020).

5 Conclusion

The data presented here show that the artisanal cheeses produced in the South and Southeast regions of Brazil display an original microbiota, only marginally contaminated by industrial starters. Interestingly, it contains original lineages of bacterial strains belonging to

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

known species of technological interest, such as *L. lactis* and *C. variabile*, as well as yeasts such as *D. hansenii*. It also contains less ordinary food species such as the bacteria *B. seileri, B. alimentarium, S. saprophyticus*, and the yeast *D. catenulata*. Moreover, two new food streptococci, *S. infantarius* and *S. salivarius-like* were found to be the main LAB in several production schemes, confirming the high originality of the Brazilian artisanal cheese microbiota and the value of this small-scale production scheme, which collectively represent a reservoir of biodiversity. Additional studies should be performed in order to improve our understanding of these microbiota and their impact on the sensory aspects of Brazilian cheeses. The further characterization of representative strains will also open the possibility to develop inoculants that maintain the cultural and historical identities of these cheeses and to establish standards of quality for their production. Finally, our study identifies several areas of concern, such as the presence of somatic cell DNA and high levels of antibiotic resistance genes in several microbiota, inferring that the milk used was from diseased herds. Overall, the data from this study highlight the potential value of the traditional and artisanal cheese production network in Brazil, and provide a metagenomic-based scheme to help manage this resource safely.

Authors' contributions

CIK and PR conceived the study and its experimental design. CIK performed microbiological, genomic, metagenomic and functional analyses. NM and PR contributed to genomic and functional analysis. CIK provided data visualization. CIK and PR wrote the manuscript. PR supervised the project.

Acknowledgements

The authors are grateful to Mathieu Almeida and Stéphane Chaillou for sharing their knowledge with us about metagenomic data analysis.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CIK's grant was supported by the "Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)" – Brazil [grant number 202444/2017-1].

Table and Figures Legends

Fig. 1. Map of Brazil showing regions and types of cheeses collected.

Fig. 2. Boxplots of alpha-diversity indices, PCoAs and clustering among bacterial communities identified in the core **(A)** or rind **(B)** of Brazilian cheeses, and fungal communities identified in the rind **(C)**. Samples are colored according to the different regions where they are produced, i.e., blue for South and red for Southeast.

Fig. 3. Bacterial (16S, **A-B**) and fungal (ITS, **C**) plot depicting the relative abundance of the 25 main species found in Brazilian cheese cores and rinds. One asterisk (*) indicates low depth of the sequences and two asterisks (**) indicate samples that are not amplified.

Fig. 4. The relative abundance of microbial composition of 15 Brazilian cheeses (six cores and nine rinds). Values were calculated from the coverage of the *ychF* marker gene assembled in the metagenomes. The asterisk (*) indicates a low depth of the sequences.

Fig. 5. Heatmap showing distribution of the 30 antibiotic resistance genes detected in reads/kb within the 15 metagenomic samples. *MLS: Macrolide, lincosamide and streptogramin.

Supplementary Material

Supplementary Data 1. The production process and characteristics of Araxá, Canastra, Serro, Colonial and Serrano.

Fig. S1. Rarefaction curves depicting the depth of 16S (core and rind) and ITS (rind) sequencing, as well as species richness for the data obtained from Brazilian cheeses. The x-axis represents the sequencing depth (reads) and the y-axis represents the estimated OTU richness detected at species level.

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

587 **Fig S2**. Phylogeny of species detected by *ychF*, highlighting the reconstituted MAGs.

588 **Fig. S3.** Potential new species and genera found by autoMLST analyses.

589 **Fig. S4.** Phylogenomic tree for *Lactococcus lactis* highlighting the different groups of *L.*
590 *lactis*.

591 **Fig. S5**. Overview of approaches used in this study.

592 **Table S1.** Metadata describing the 23 cheese samples analyzed.

593 **Table S2.** Raw reads detected of bacterial and fungal operational taxonomic units (OTUs) in
594 the samples.

595 **Table S3.** Percentage (%) of *Bos Taurus* reads aligned in each cheese metagenome.

596 **Table S4.** The relative abundance of microbial composition of 15 cheese metagenomes from
597 Brazil using MetaPhlAn taxonomic assignment and *ychF* marker gene methods.

598 **Table S5.** Quality of prokaryotic and eukaryotic MAGs and their ANIs with the closest
599 reference genome found in the NCBI database.

600 **Table S6.** Coverage (reads/kb) of antibiotic resistance and genes of technological interest.

601 **Table S7**. Detection of 12 pathogens commonly found in dairy environments by read
602 mapping expressed in percentage of total read mapped on a set of reference genomes.

603

604 **6 References**

605 Alanjary, M., Steinke, K., & Ziemert, N. (2019). AutoMLST: an automated web server for generating
606 multi-locus species trees highlighting natural product potential. *Nucleic Acids Research,*
607 *47*(W1), W276-W282. doi:10.1093/nar/gkz282
608 Almeida, M., Hébert, A., Abraham, A.-L., Rasmussen, S., Monnet, C., Pons, N., . . . Renault, P. (2014).
609 Construction of a dairy microbial genome catalog opens new perspectives for the
610 metagenomic analysis of dairy fermented products. *BMC Genomics, 15*(1), 1101.
611 doi:10.1186/1471-2164-15-1101
612 Antonopoulos, D. A., Assaf, R., Aziz, R. K., Brettin, T., Bun, C., Conrad, N., . . . Yoo, H. (2019). PATRIC as
613 a unique resource for studying antimicrobial resistance. *Briefings in Bioinformatics, 20*(4),
614 1094-1102. doi:10.1093/bib/bbx083
615 Argudín, M. Á., Mendoza, M. C., & Rodicio, M. R. (2010). Food Poisoning and Staphylococcus aureus
616 Enterotoxins. *Toxins, 2*(7). doi:10.3390/toxins2071751
617 Balaban, N., & Rasooly, A. (2000). Staphylococcal enterotoxins. *International Journal of Food*
618 *Microbiology, 61*(1), 1-10. doi:https://doi.org/10.1016/S0168-1605(00)00377-9

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

619 Banjara, N., Suhr, M. J., & Hallen-Adams, H. E. (2015). Diversity of Yeast and Mold Species from a
620 Variety of Cheese Types. *Current Microbiology, 70*(6), 792-800. doi:10.1007/s00284-015-
621 0790-1
622 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Pevzner, P. A.
623 (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell
624 sequencing. *Journal of computational biology : a journal of computational molecular cell*
625 *biology, 19*(5), 455-477. doi:10.1089/cmb.2012.0021
626 Bertuzzi, A. S., Walsh, A. M., Sheehan, J. J., Cotter, P. D., Crispie, F., McSweeney, P. L. H., . . . Rea, M.
627 C. (2018). Omics-Based Insights into Flavor Development and Microbial Succession within
628 Surface-Ripened Cheese. *mSystems, 3*(1), e00211-00217. doi:10.1128/mSystems.00211-17
629 Bintsis, T. (2018). Lactic acid bacteria as starter cultures: An update in their metabolism and genetics.
630 *AIMS Microbiology, 4*(4), 665-684. doi:10.3934/microbiol.2018.4.665
631 Borelli, B. M., Lacerda, I. C. A., Penido, F. C. L., & Rosa, C. A. (2016). Traditional cheeses produced in
632 Brazio: Characterisation, production technologies and health implications. In E. Perkins (Ed.),
633 *Food Microbiology* (pp. 161-190). New York: Nova Publishers.
634 Brito, L. P. d., Silva, E. C. d., Calaça, P. R. d. A., Medeiros, R. S. d., Soares, M. T. C. V., & Porto, A. L. F.
635 (2020). Lactic acid bacteria isolated from Coalho cheese from northeast Brazil in dairy
636 production: A screening for technological application. *Research, Society and Development,*
637 *9*(10), e5249108457. doi:10.33448/rsd-v9i10.8457
638 Cabral, M. L. B., Lima, M. d. S. F. d., Fernandes, G., Costa, E. G. d., Porto, A. L. F., & Cavalcanti, M. T. H.
639 (2016). Artisan cheese: a potential source of wild lactic acid bacteria to obtain new starter
640 cultures. *Journal of Bioenergy and Food Science, 3*, 207-215.
641 Carlsson, J., Grahnén, H., Jonsson, G., & Wikner, S. (1970). Early Establishment of Streptococcus
642 salivarius in the Mouths of Infants. *Journal of Dental Research, 49*(2), 415-418.
643 doi:10.1177/00220345700490023601
644 Casaes Nunes, R. S., Pires de Souza, C., Pereira, K. S., Del Aguila, E. M., & Flosi Paschoalin, V. M.
645 (2016). Identification and molecular phylogeny of coagulase-negative staphylococci isolates
646 from Minas Frescal cheese in southeastern Brazil: Superantigenic toxin production and
647 antibiotic resistance. *Journal of Dairy Science, 99*(4), 2641-2653.
648 doi:https://doi.org/10.3168/jds.2015-9693
649 Ciccarelli, F. D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B., & Bork, P. (2006). Toward
650 Automatic Reconstruction of a Highly Resolved Tree of Life. *Science, 311*(5765), 1283.
651 doi:10.1126/science.1123061
652 Costa, A., Neves Pereira, A., Silva, A., Silva, F., Ribeiro, K., Torres, I., . . . Alves, V. (2018). Antilisterial
653 and antistaphylococcal activity of a Lactococcus lactis strain isolated from Brazilian fresh
654 Minas cheese. *Journal of Food Safety, 39*. doi:10.1111/jfs.12593
655 Coton, E., Desmonts, M.-H., Leroy, S., Coton, M., Jamet, E., Christieans, S., . . . Talon, R. (2010).
656 Biodiversity of Coagulase-Negative Staphylococci in French cheeses, dry fermented sausages,
657 processing environments and clinical samples. *International Journal of Food Microbiology,*
658 *137*(2), 221-229. doi:https://doi.org/10.1016/j.ijfoodmicro.2009.11.023
659 Cretenet, M., Even, S., & Le Loir, Y. (2011). Unveiling Staphylococcus aureus enterotoxin production
660 in dairy products: a review of recent advances to face new challenges. *Dairy Science &*
661 *Technology, 91*(2), 127-150. doi:10.1007/s13594-011-0014-9
662 de Medeiros Carvalho, M., de Fariña, L. O., Strongin, D., Ferreira, C. L. L. F., & Lindner, J. D. D. (2019).
663 Traditional Colonial-type cheese from the south of Brazil: A case to support the new Brazilian
664 laws for artisanal cheese production from raw milk. *Journal of Dairy Science, 102*(11), 9711-
665 9720. doi:https://doi.org/10.3168/jds.2019-16373
666 Delbes, C., Monnet, C., & Irlinger, F. (2015). Des communautés microbiennes au service de la qualité
667 des fromages : diversité et dynamique adaptative et fonctionnelle des populations
668 endogènes et ensemencées.
669 Dittmann, K. K., Chaul, L. T., Lee, S. H. I., Corassin, C. H., Fernandes de Oliveira, C. A., Pereira De
670 Martinis, E. C., . . . Oxaran, V. (2017). Staphylococcus aureus in Some Brazilian Dairy

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Industries: Changes of Contamination and Diversity. *Frontiers in Microbiology, 8*(2049). doi:10.3389/fmicb.2017.02049

Dugat-Bony, E., Garnier, L., Denonfoux, J., Ferreira, S., Sarthou, A. S., Bonnarme, P., & Irlinger, F. (2016). Highlighting the microbial diversity of 12 French cheese varieties. *International Journal of Food Microbiology, 238*, 265-273. doi:10.1016/j.ijfoodmicro.2016.09.026

Dugat-Bony, E., Sarthou, A. S., Loux, V., Vidal, M., Bonnarme, P., Irlinger, F., & Layec, S. (2016). Draft Genome Sequence of *Corynebacterium variabile* Mu292, Isolated from Munster, a French Smear-Ripened Cheese. *Genome announcements, 4*(4), e00669-00616. doi:doi:10.1128/genomeA.00669-16

Escudié, F., Auer, L., Bernard, M., Mariadassou, M., Cauquil, L., Vidal, K., . . . Pascal, G. (2018). FROGS: Find, Rapidly, OTUs with Galaxy Solution. *Bioinformatics, 34*(8), 1287-1294. doi:10.1093/bioinformatics/btx791

Felsenstein, J. (1985). Confidence-Limits on Phylogenies - an Approach Using the Bootstrap. *Evolution, 39*(4), 783-791. doi:DOI 10.1111/j.1558-5646.1985.tb00420.x

Flórez, A. B., Delgado, S., & Mayo, B. (2005). Antimicrobial susceptibility of lactic acid bacteria isolated from a cheese environment. *Canadian journal of microbiology, 51*(1), 51-58. doi:10.1139/w04-114

Flórez, A. B., Vázquez, L., & Mayo, B. (2017). A Functional Metagenomic Analysis of Tetracycline Resistance in Cheese Bacteria. *Frontiers in Microbiology, 8*(907). doi:10.3389/fmicb.2017.00907

Gaglio, R., Todaro, M., & Settanni, L. (2021). Improvement of Raw Milk Cheese Hygiene through the Selection of Starter and Non-Starter Lactic Acid Bacteria: The Successful Case of PDO Pecorino Siciliano Cheese. *International Journal of Environmental Research and Public Health, 18*, 1834. doi:10.3390/ijerph18041834

García-Díez, J., & Saraiva, C. (2021). Use of Starter Cultures in Foods from Animal Origin to Improve Their Safety. *International Journal of Environmental Research and Public Health, 18*(5). doi:10.3390/ijerph18052544

Gardner, S. N., Slezak, T., & Hall, B. G. (2015). kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics, 31*(17), 2877-2878. doi:10.1093/bioinformatics/btv271

Ge, B., Jiang, P., Han, F., Saleh, N. K., Dhiman, N., Fedorko, D. P., . . . Meng, J. (2007). Identification and Antimicrobial Susceptibility of Lactic Acid Bacteria from Retail Fermented Foods. *Journal of Food Protection, 70*(11), 2606-2612. doi:10.4315/0362-028X-70.11.2606

Hammer, P., Jordan, J., Jacobs, C., & Klempt, M. (2019). Characterization of coagulase-negative staphylococci from brining baths in Germany. *Journal of Dairy Science, 102*(10), 8734-8744. doi:https://doi.org/10.3168/jds.2018-15610

Irlinger, F., Layec, S., Hélinck, S., & Dugat-Bony, E. (2015). Cheese rind microbial communities: diversity, composition and origin. *Fems Microbiology Letters, 362*(2), 1-11. doi:10.1093/femsle/fnu015

Irlinger, F., & Monnet, C. (2021). Temporal differences in microbial composition of Epoisses cheese rinds during ripening and storage. *Journal of Dairy Science, 104*(7), 7500-7508. doi:10.3168/jds.2021-20123

Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., & Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nature Communications, 9*(1), 5114. doi:10.1038/s41467-018-07641-9

Jamali, H., Paydar, M., Radmehr, B., Ismail, S., & Dadrasnia, A. (2015). Prevalence and antimicrobial resistance of Staphylococcus aureus isolated from raw milk and dairy products. *Food Control, 54*, 383-388. doi:https://doi.org/10.1016/j.foodcont.2015.02.013

Jans, C., Follador, R., Hochstrasser, M., Lacroix, C., Meile, L., & Stevens, M. J. (2013). Comparative genome analysis of Streptococcus infantarius subsp. infantarius CJ18, an African fermented camel milk isolate with adaptations to dairy environment. *BMC Genomics, 14*, 200. doi:10.1186/1471-2164-14-200

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Jans, C., Kaindi, D. W., Bock, D., Njage, P. M., Kouame-Sina, S. M., Bonfoh, B., . . . Meile, L. (2013). Prevalence and comparison of Streptococcus infantarius subsp. infantarius and Streptococcus gallolyticus subsp. macedonicus in raw and fermented dairy products from East and West Africa. *Int J Food Microbiol, 167*(2), 186-195. doi:10.1016/j.ijfoodmicro.2013.09.008

Jans, C., Meile, L., Kaindi, D. W. M., Kogi-Makau, W., Lamuka, P., Renault, P., . . . Bonfoh, B. (2017). African fermented dairy products – Overview of predominant technologically important microorganisms focusing on African Streptococcus infantarius variants and potential future applications for enhanced food safety and security. *International Journal of Food Microbiology, 250*, 27-36. doi:https://doi.org/10.1016/j.ijfoodmicro.2017.03.012

Jans, C., Meile, L., Lacroix, C., & Stevens, M. J. A. (2015). Genomics, evolution, and molecular epidemiology of the Streptococcus bovis/Streptococcus equinus complex (SBSEC). *Infection, Genetics and Evolution, 33*, 419-436. doi:https://doi.org/10.1016/j.meegid.2014.09.017

Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., . . . McArthur, A. G. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research, 45*(D1), D566-D573. doi:10.1093/nar/gkw1004

Kadariya, J., Smith, T. C., & Thapaliya, D. (2014). *Staphylococcus aureus* and Staphylococcal Food-Borne Disease: An Ongoing Challenge in Public Health. *BioMed Research International, 2014*, 827965. doi:10.1155/2014/827965

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. doi:https://doi.org/10.1016/j.fm.2018.12.014

Kamimura, B. A., Magnani, M., Luciano, W. A., Campagnollo, F. B., Pimentel, T. C., Alvarenga, V. O., . . . Sant'Ana, A. S. (2019). Brazilian Artisanal Cheeses: An Overview of their Characteristics, Main Types and Regulatory Aspects. *Comprehensive Reviews in Food Science and Food Safety, 18*(5), 1636-1657. doi:https://doi.org/10.1111/1541-4337.12486

Kang, D. D., Froula, J., Egan, R., & Wang, Z. (2015). MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *Peerj, 3*, e1165. doi:10.7717/peerj.1165

Kothe, C. I., Bolotin, A., Kraïem, B.-F., Dridi, B., team, F. M., & Renault, P. (2021). Unraveling the world of halophilic and halotolerant bacteria in cheese by combining cultural, genomic and metagenomic approaches. *International Journal of Food Microbiology*, 109312. doi:https://doi.org/10.1016/j.ijfoodmicro.2021.109312

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution, 33*(7), 1870-1874. doi:10.1093/molbev/msw054

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods, 9*(4), 357-359. doi:10.1038/nmeth.1923

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Subgroup., G. P. D. P. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics, 25*(16), 2078-2079. doi:10.1093/bioinformatics/btp352

Li, X., Li, Y., Alvarez, V., Harper, W. J., & Wang, H. H. (2011). Effective antibiotic resistance mitigation during cheese fermentation. *Applied and Environmental Microbiology, 77*(20), 7171-7175. doi:10.1128/AEM.05069-11

Li, Y., Cao, W., Liang, S., Yamasaki, S., Chen, X., Shi, L., & Ye, L. (2020). Metagenomic characterization of bacterial community and antibiotic resistance genes in representative ready-to-eat food in southern China. *Scientific Reports, 10*(1), 15175. doi:10.1038/s41598-020-72620-4

Lima, C. D. L. C., Lima, L. A., Cerqueira, M. M. O. P., Ferreira, E. G., & Rosa, C. A. (2009). Bactérias do acido láctico e leveduras associadas com o queijo-de-minas artesanal produzido na região da Serra do Salitre, Minas Gerais. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 61*, 266-272. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-09352009000100037&nrm=iso

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

775 Luiz, L., Castro, R., Silva, J., Oliveira, L., Sales, G., Nunes, A. C., . . . Sandes, S. (2016). Isolation and
776 identification of lactic acid bacteria from Brazilian Minas artisanal cheese. *CyTA - Journal of*
777 *Food, 15*, 125. doi:10.1080/19476337.2016.1219392
778 Mahe, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2014). Swarm: robust and fast
779 clustering method for amplicon-based studies. *Peerj, 2*. doi:10.7717/peerj.593
780 MAPA (Ministério da Agricultura, Pecuária e Abastecimento). Dispõe sobre a identidade e os
781 requisitos de qualidade que deve apresentar o produto denominado queijo minas meia cura,
782 INSTRUÇÃO NORMATIVA Nº 74, DE 24 DE JULHO DE 2020 C.F.R. (2020).
783 Margalho, L. P., Feliciano, M. D. E., Silva, C. E., Abreu, J. S., Piran, M. V. F., & Sant'Ana, A. S. (2020).
784 Brazilian artisanal cheeses are rich and diverse sources of nonstarter lactic acid bacteria
785 regarding technological, biopreservative, and safety properties—Insights through
786 multivariate analysis. *Journal of Dairy Science, 103*(9), 7908-7926.
787 doi:https://doi.org/10.3168/jds.2020-18194
788 McArthur, A. G., Waglechner, N., Nizam, F., Yan, A., Azad, M. A., Baylay, A. J., . . . Wright, G. D. (2013).
789 The comprehensive antibiotic resistance database. *Antimicrobial agents and chemotherapy,*
790 *57*(7), 3348-3357. doi:10.1128/AAC.00419-13
791 McCarthy, C., Snyder, M. L., & Parker, R. B. (1965). The Indigenous Oral Flora of Man. I. The Newborn
792 to the 1-Year-Old Infant. *Arch Oral Biol, 10*, 61-70. doi:10.1016/0003-9969(65)90058-0
793 McMurdie, P. J., & Holmes, S. (2013). phyloseq: An R Package for Reproducible Interactive Analysis
794 and Graphics of Microbiome Census Data. *PLOS ONE, 8*(4), e61217.
795 doi:10.1371/journal.pone.0061217
796 Medeiros, R. S., Araújo, L. M., Queiroga Neto, V., Andrade, P. P., Melo, M. A., & Gonçalves, M. M. B.
797 P. (2016). Identification of lactic acid bacteria isolated from artisanal Coalho cheese produced
798 in the Brazilian Northeast. *CyTA - Journal of Food, 14*(4), 613-620.
799 doi:10.1080/19476337.2016.1185468
800 Montel, M.-C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D. A., Desmasures, N., & Berthier, F.
801 (2014). Traditional cheeses: Rich and diverse microbiota with associated benefits.
802 *International Journal of Food Microbiology, 177*, 136-154.
803 doi:https://doi.org/10.1016/j.ijfoodmicro.2014.02.019
804 Moradi, M., Omer, A. K., Razavi, R., Valipour, S., & Guimarães, J. T. (2021). The relationship between
805 milk somatic cell count and cheese production, quality and safety: A review. *International*
806 *Dairy Journal, 113*, 104884. doi:https://doi.org/10.1016/j.idairyj.2020.104884
807 Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., & Tyson, G. W. (2015). CheckM: assessing
808 the quality of microbial genomes recovered from isolates, single cells, and metagenomes.
809 *Genome Research, 25*(7), 1043-1055. doi:10.1101/gr.186072.114
810 Penna, A. L. B., Gigante, M. L., & Todorov, S. D. (2021). Artisanal Brazilian Cheeses—History,
811 Marketing, Technological and Microbiological Aspects. *Foods, 10*(7), 1562. Retrieved from
812 https://www.mdpi.com/2304-8158/10/7/1562
813 Perin, L. M., Sardaro, M. L. S., Nero, L. A., Neviani, E., & Gatti, M. (2017). Bacterial ecology of artisanal
814 Minas cheeses assessed by culture-dependent and -independent methods. *Food*
815 *Microbiology, 65*, 160-169. doi:10.1016/j.fm.2017.02.005
816 Perkins, V., Vignola, S., Lessard, M. H., Plante, P. L., Corbeil, J., Dugat-Bony, E., . . . Labrie, S. (2020).
817 Phenotypic and Genetic Characterization of the Cheese Ripening Yeast Geotrichum
818 candidum. *Front Microbiol, 11*, 737. doi:10.3389/fmicb.2020.00737
819 Petzer, I.-M., Karzis, J., Donkin, E. F., Webb, E. C., & Etter, E. M. C. (2017). Somatic cell count
820 thresholds in composite and quarter milk samples as indicator of bovine intramammary
821 infection status. *The Onderstepoort journal of veterinary research, 84*(1), e1-e10.
822 doi:10.4102/ojvr.v84i1.1269
823 Pham, N. P., Landaud, S., Lieben, P., Bonnarme, P., & Monnet, C. (2019). Transcription Profiling
824 Reveals Cooperative Metabolic Interactions in a Microbial Cheese-Ripening Community
825 Composed of Debaryomyces hansenii, Brevibacterium aurantiacum, and Hafnia alvei. *Front*
826 *Microbiol, 10*, 1901. doi:10.3389/fmicb.2019.01901

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

827 Poirier, S., Rué, O., Peguilhan, R., Coeuret, G., Zagorec, M., Champomier-Vergès, M.-C., . . . Chaillou,
828 S. (2018). Deciphering intra-species bacterial diversity of meat and seafood spoilage
829 microbiota using gyrB amplicon sequencing: A comparative analysis with 16S rDNA V3-V4
830 amplicon sequencing. *PLOS ONE, 13*(9), e0204629. doi:10.1371/journal.pone.0204629
831 Pontarolo, G. H., Melo, F. D., Martini, C. L., Wildemann, P., Alessio, D. R. M., Sfaciotte, R. A. P., . . .
832 Ferraz, S. M. (2017). Quality and safety of artisan cheese produced in the serrana region of
833 Santa Catarina. *Semina-Ciencias Agrarias, 38*(2), 739-748. doi:10.5433/1679-
834 0359.2017v38n2p739
835 Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., . . . Glöckner, F. O. (2013). The
836 SILVA ribosomal RNA gene database project: improved data processing and web-based tools.
837 *Nucleic Acids Research, 41*(D1), D590-D596. doi:10.1093/nar/gks1219
838 Rebecchi, A., Pisacane, V., Callegari, M. L., Puglisi, E., & Morelli, L. (2015). Ecology of antibiotic
839 resistant coagulase-negative staphylococci isolated from the production chain of a typical
840 Italian salami. *Food Control, 53*, 14-22. doi:https://doi.org/10.1016/j.foodcont.2015.01.001
841 Resende, M. F. S., Costa, H. H. S., Andrade, E. H. P., Acurcio, L. B., Drummond, A. F., Cunha, A. F., . . .
842 Souza, M. R. (2011). Influence of altitute on lactic acid bacteria population of minas artisanal
843 cheese from Serra da Canastra. *Arquivo Brasileiro de Medicina Veterinária e Zootecnia, 63*(6),
844 1567-1573. doi:Doi 10.1590/S0102-09352011000600039
845 Rizzotti, L., La Gioia, F., Dellaglio, F., & Torriani, S. (2009). Characterization of tetracycline-resistant
846 Streptococcus thermophilus isolates from Italian soft cheeses. *Applied and Environmental*
847 *Microbiology, 75*(12), 4224-4229. doi:10.1128/AEM.01380-08
848 Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahe, F. (2016). VSEARCH: a versatile open source
849 tool for metagenomics. *Peerj, 4*. doi:10.7717/peerj.2584
850 RS (Rio Grande do Sul). Dispõe sobre a produção e a comercialização do queijo artesanal serrano no
851 Estado do Rio Grande do Sul, (2016).
852 Saitou, N., & Nei, M. (1987). The Neighbor-Joining Method - a New Method for Reconstructing
853 Phylogenetic Trees. *Molecular Biology and Evolution, 4*(4), 406-425. doi:DOI
854 10.1093/oxfordjournals.molbev.a040454
855 Schröder, J., Maus, I., Trost, E., & Tauch, A. (2011). Complete genome sequence of Corynebacterium
856 variabile DSM 44702 isolated from the surface of smear-ripened cheeses and insights into
857 cheese ripening and flavor generation. *BMC Genomics, 12*(1), 545. doi:10.1186/1471-2164-
858 12-545
859 Sunagawa, S., Mende, D., Zeller, G., Izquierdo-Carrasco, F., Berger, S., Kultima, J., . . . Bork, P. (2013).
860 Metagenomic species profiling using universal phylogenetic marker genes. *Nature Methods,*
861 *10*. doi:10.1038/nmeth.2693
862 Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997). The CLUSTAL_X
863 windows interface: flexible strategies for multiple sequence alignment aided by quality
864 analysis tools. *Nucleic Acids Research, 25*(24), 4876-4882. doi:DOI 10.1093/nar/25.24.4876
865 Tóth, A. G., Csabai, I., Krikó, E., Tőzsér, D., Maróti, G., Patai, Á. V., . . . Solymosi, N. (2020).
866 Antimicrobial resistance genes in raw milk for human consumption. *Scientific Reports, 10*(1),
867 7464. doi:10.1038/s41598-020-63675-4
868 Truong, D. T., Franzosa, E., Tickle, T., Scholz, M., Weingart, G., Pasolli, E., . . . Segata, N. (2015).
869 MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nature Methods, 12*, 902-903.
870 doi:10.1038/nmeth0116-101b
871 Vandera, E., Kakouri, A., Koukkou, A.-I., & Samelis, J. (2019). Major ecological shifts within the
872 dominant nonstarter lactic acid bacteria in mature Greek Graviera cheese as affected by the
873 starter culture type. *International Journal of Food Microbiology, 290*, 15-26.
874 doi:https://doi.org/10.1016/j.ijfoodmicro.2018.09.014
875 Walsh, A., Macori, G., Kilcawley, K., & Cotter, P. (2020). Meta-analysis of cheese microbiomes
876 highlights contributions to multiple aspects of quality. *Nature Food, 1*, 500-510.
877 doi:10.1038/s43016-020-0129-3

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

878 Wescombe, P. A., Hale, J. D. F., Heng, N. C. K., & Tagg, J. R. (2012). Developing oral probiotics from
879 Streptococcus salivarius. *Future Microbiology, 7*(12), 1355-1371. doi:10.2217/fmb.12.113
880 Wingett, S., & Andrews, S. (2018). FastQ Screen: A tool for multi-genome mapping and quality
881 control. *F1000Research, 7*, 1338. doi:10.12688/f1000research.15931.1
882 Wolfe, B. E., Button, J. E., Santarelli, M., & Dutton, R. J. (2014). Cheese Rind Communities Provide
883 Tractable Systems for In Situ and In Vitro Studies of Microbial Diversity. *Cell, 158*(2), 422-433.
884 doi:10.1016/j.cell.2014.05.041
885 Yadav, S., & Kapley, A. (2021). Antibiotic resistance: Global health crisis and metagenomics.
886 *Biotechnology Reports, 29*, e00604. doi:https://doi.org/10.1016/j.btre.2021.e00604

bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

SOUTHEAST

SOUTH

Canastra

Araxa

Serro

Colonial

Serrano

(A) CORE 16S

(B) RIND 16S

(C) RIND ITS

Relative abundance

1.00

0.75

0.50

0.25

0.00

C-01 C-02 C-03 C-04 C-05 C-06 C-07 C-08 C-09 C-10 C-11 C-12 C-13 C-14 C-15 C-16 C-17 C-18 C-19 C-20 C-21 C-22 C-23

C-R01 C-R02 C-R03 C-R04 C-R05 C-R06 C-R07 C-R08 C-R09 C-R10 C-R11 C-R12 C-R13 C-R14 C-R15 C-R16 C-R17 CR-18 C-R19 C-R20 C-R21 C-R22 CR-23

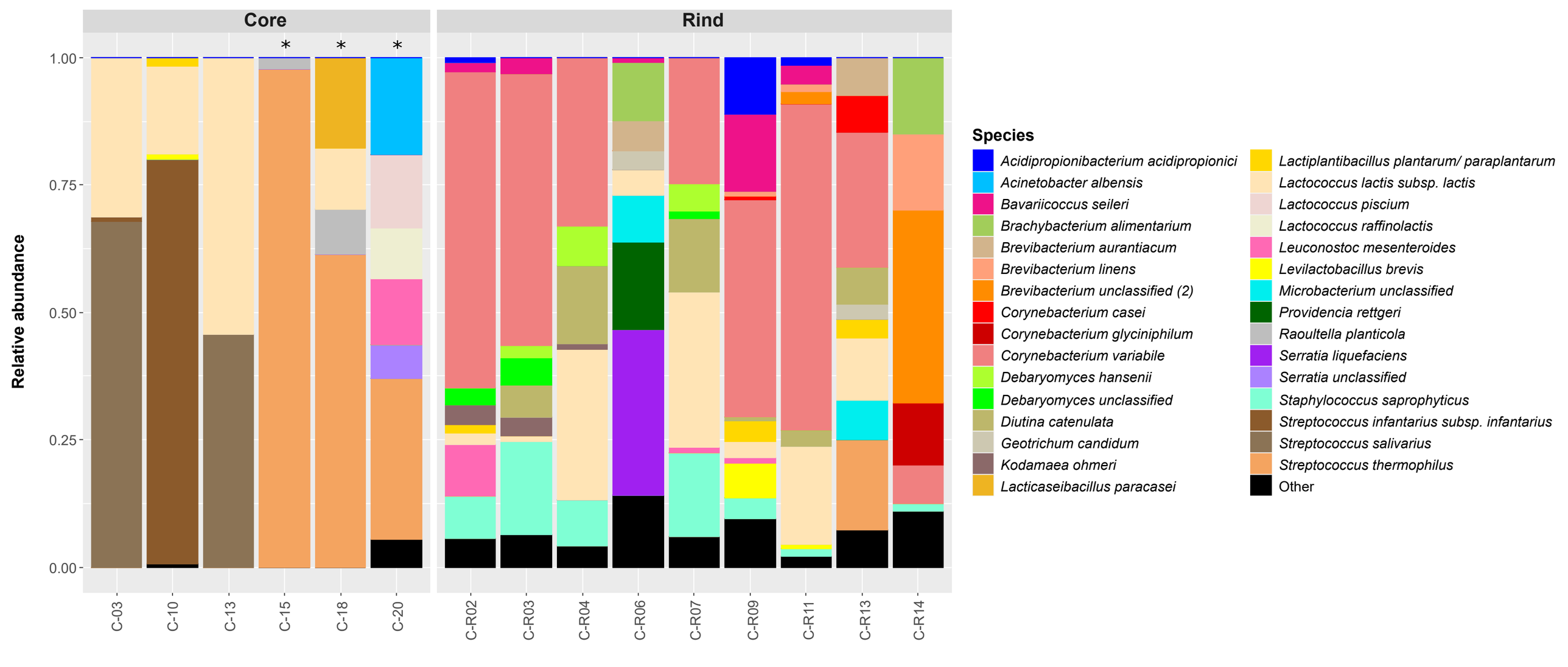
Samples

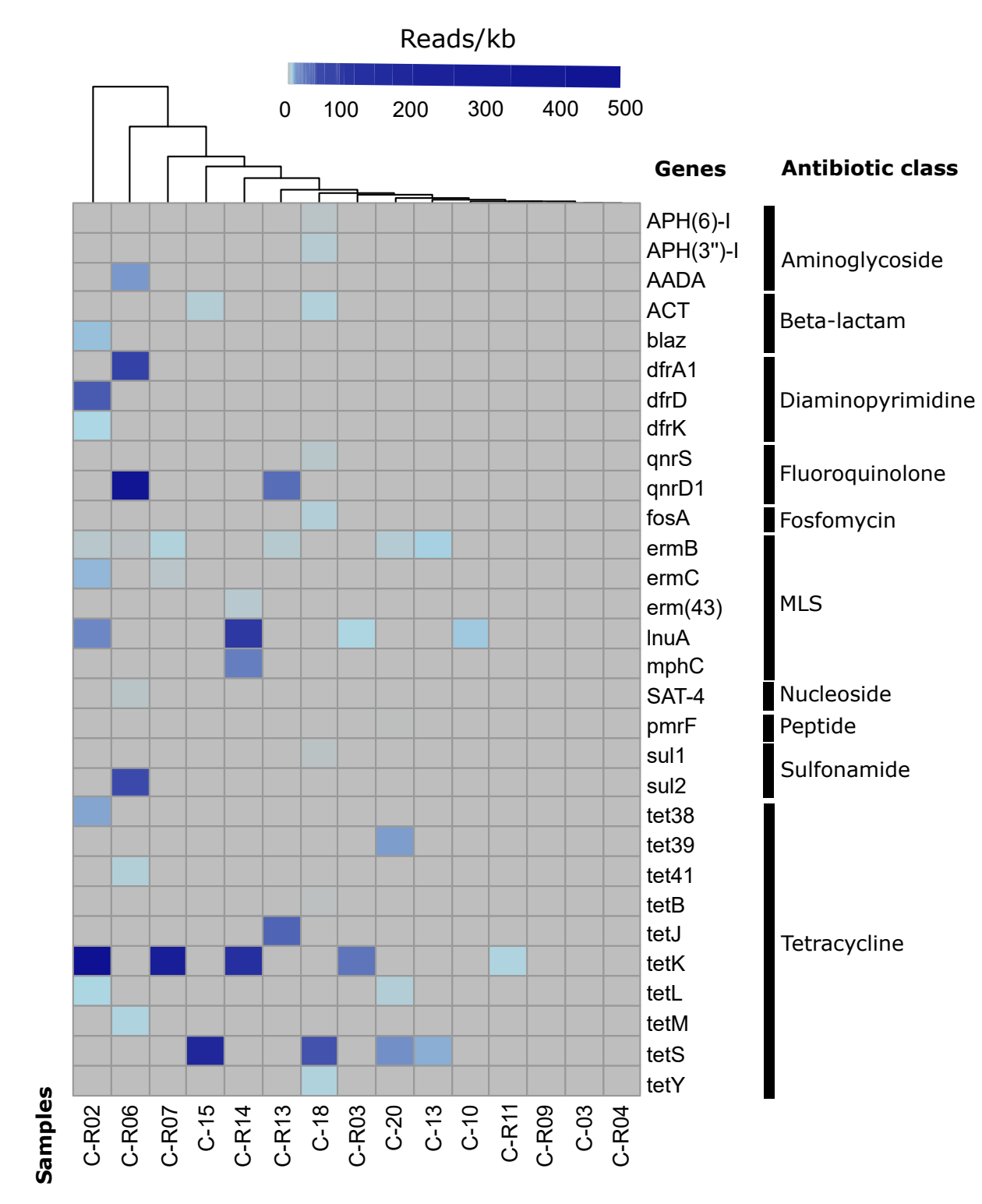
Species

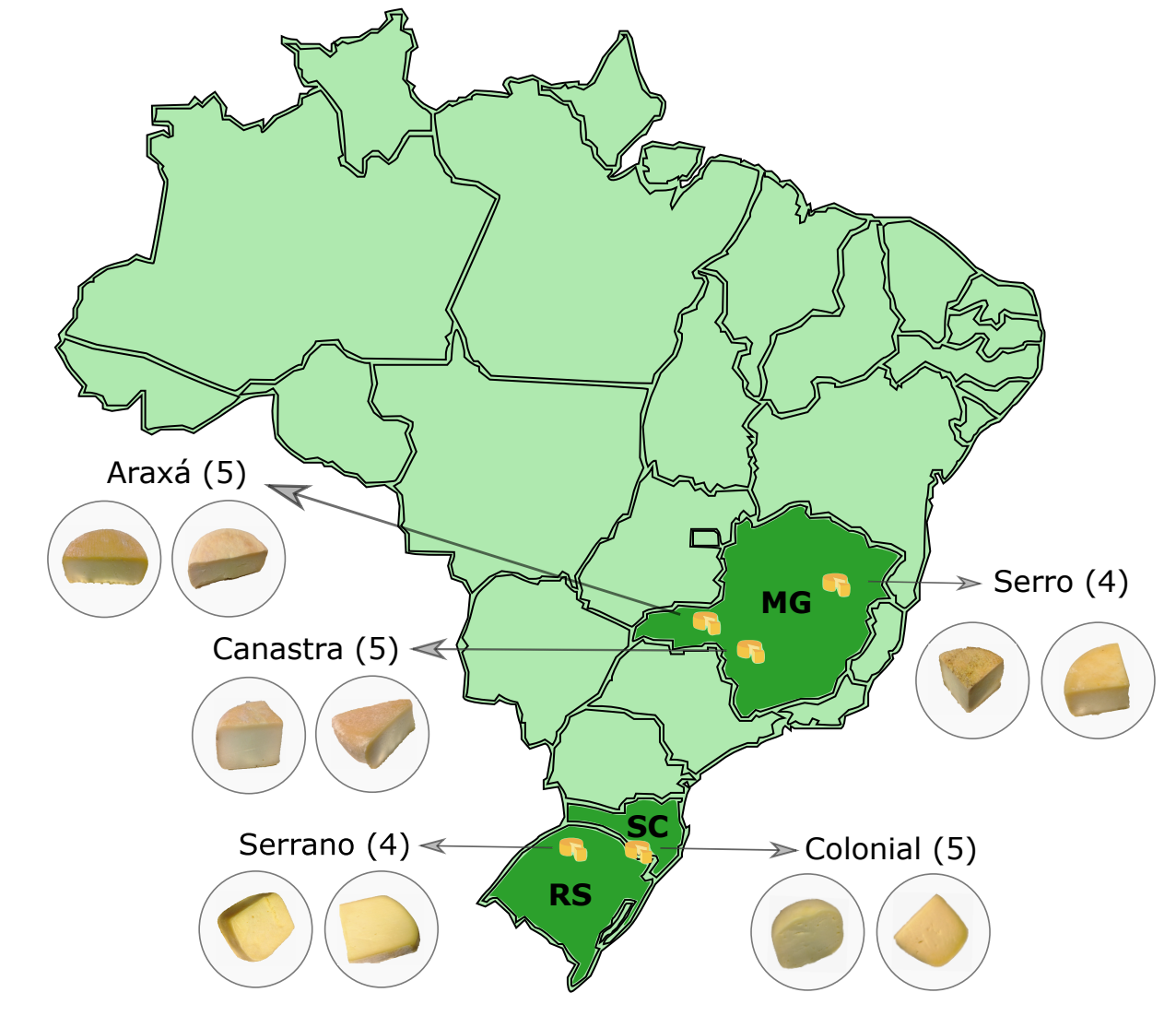
- *Cobetia marina*
- *Cobetia sp.*
- *Enterobacter sp.*
- *Escherichia coli*
- *Kluyvera cryocrescens*
- *Kurthia gibsonii*
- *Levilactobacillus brevis*
- *Lacticaseibacillus casei-paracasei*
- *Lactobacillus delbrueckii*
- *Lactiplantibacillus plantarum*
- *Lactococcus lactis*
- *Lactococcus piscium*
- *Lactococcus raffinolactis group*
- *Lactococcus sp. 1*
- *Leuconostoc mesenteroides*
- *Raoultella ornithinolytica*
- *Staphylococcus aureus*
- *Streptococcus equinus-lutetiensis-infantarius*
- *Streptococcus parauberis*
- *Streptococcus porcorum/sanguinis*
- *Streptococcus sp. 1*
- *Streptococcus sp. 2*
- *Streptococcus sp. 3*
- *Streptococcus sp. 4*
- *Streptococcus thermophilus-salivarius*
- Other

- *Bavariicoccus seileri*
- *Brevibacterium sp. 1*
- *Celerinatantimonas sp.*
- *Chromohalobacter canadensis*
- *Cobetia marina*
- *Corynebacterium glyciniphilum*
- *Corynebacterium sp. 1*
- *Corynebacterium sp. 2*
- *Corynebacterium variabile*
- *Lactococcus lactis*
- *Lactococcus sp. 1*
- *Leuconostoc mesenteroides*
- *Providencia rettgeri*
- *Psychrobacter sp. 1*
- *Rouxiella sp. 2*
- *Serratia liquefaciens*
- *Serratia sp.*
- *Staphylococcus equorum*
- *Staphylococcus saprophyticus*
- *Streptococcus equinus-lutetiensis-infantarius*
- *Streptococcus porcorum-sanguinis*
- *Streptococcus sp. 1*
- *Streptococcus sp. 2*
- *Streptococcus thermophilus-salivarius*
- *Weissella jogaejeotgali*
- Other

- *Acremonium citrinum*
- *Candida palmioleophila*
- *Candida parapsilosis*
- *Candida zeylanoides*
- *Clavispora lusitaniae*
- *Clavispora sp.*
- *Debaryomyces hansenii*
- *Diutina catenulata*
- *Fusarium solani*
- *Fusarium sp.*
- *Geotrichum candidum*
- *Gibberella intricans*
- *Kluyveromyces lactis*
- *Kodamaea ohmeri*
- *Microascus brevicaulis*
- *Moniliella sp.*
- *Penicillium commune*
- *Penicillium sp.*
- *Trichosporon sp.*
- *Yarrowia lipolytica*
- Other


Core
Rind
Relative abundance
1.00
0.75
0.50
0.25
0.00
*
*
*
C-03
C-10
C-13
C-15
C-18
C-20
C-R02
C-R03
C-R04
C-R06
C-R07
C-R09
C-R11
C-R13
C-R14
Species
Acidipropionibacterium acidipropionici
Acinetobacter albensis
Bavariicoccus seileri
Brachybacterium alimentarium
Brevibacterium aurantiacum
Brevibacterium linens
Brevibacterium unclassified (2)
Corynebacterium casei
Corynebacterium glyciniphilum
Corynebacterium variabile
Debaryomyces hansenii
Debaryomyces unclassified
Diutina catenulata
Geotrichum candidum
Kodamaea ohmeri
Lacticaseibacillus paracasei
Lactiplantibacillus plantarum/ paraplantarum
Lactococcus lactis subsp. lactis
Lactococcus piscium
Lactococcus raffinolactis
Leuconostoc mesenteroides
Levilactobacillus brevis
Microbacterium unclassified
Providencia rettgeri
Raoultella planticola
Serratia liquefaciens
Serratia unclassified
Staphylococcus saprophyticus
Streptococcus infantarius subsp. infantarius
Streptococcus salivarius
Streptococcus thermophilus
Other

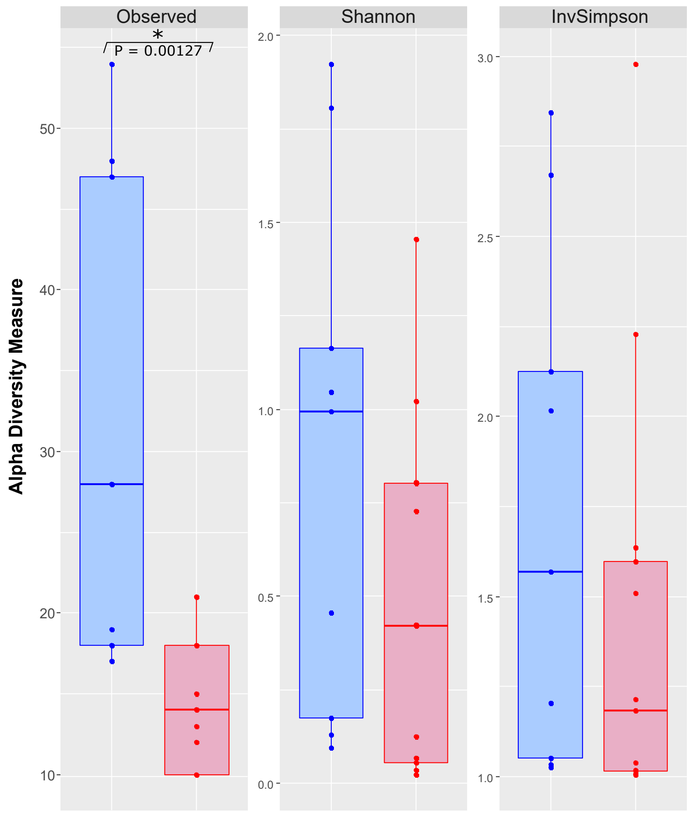

Reads/kb
0
100
200
300
400
500
Genes
Antibiotic class
APH(6)-I
APH(3")-I
AADA
ACT
blaz
dfrA1
dfrD
dfrK
qnrS
qnrD1
fosA
ermB
ermC
erm(43)
lnuA
mphC
SAT-4
pmrF
sul1
sul2
tet38
tet39
tet41
tetB
tetJ
tetK
tetL
tetM
tetS
tetY
Aminoglycoside
Beta-lactam
Diaminopyrimidine
Fluoroquinolone
Fosfomycin
MLS
Nucleoside
Peptide
Sulfonamide
Tetracycline
Samples
C-R02
C-R06
C-R07
C-15
C-R14
C-R13
C-18
C-R03
C-20
C-13
C-10
C-R11
C-R09
C-03
C-R04

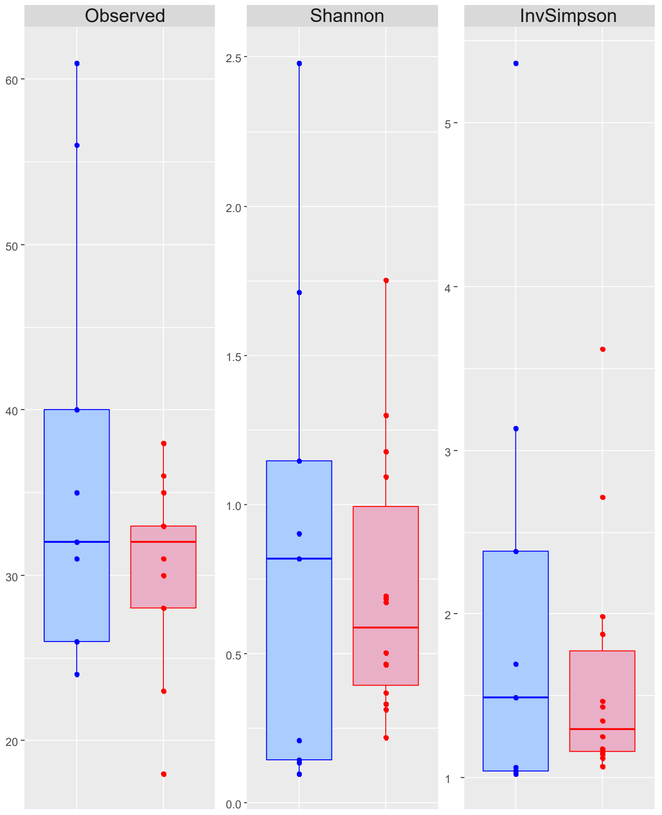

Araxá (5)
Canastra (5)
Serrano (4)
Serro (4)
Colonial (5)
MG
SC
RS

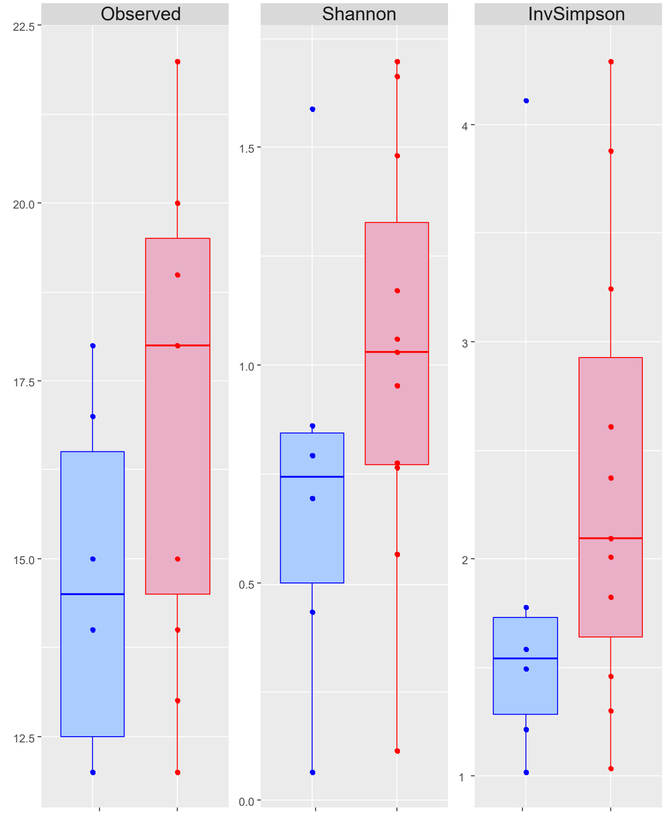
(A) CORE 16S

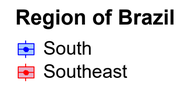
(B) RIND 16S

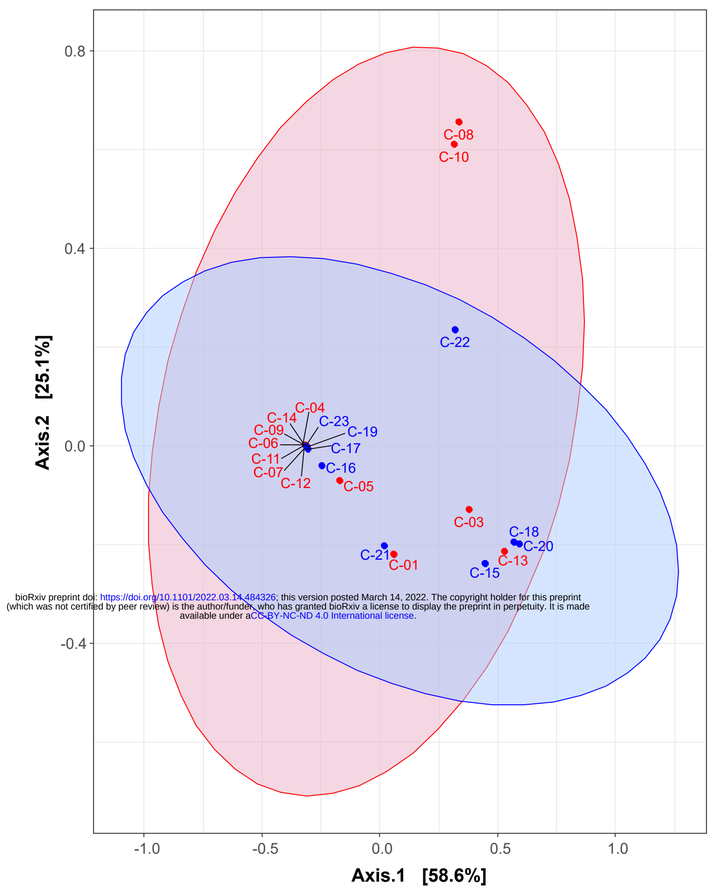
(C) RIND ITS

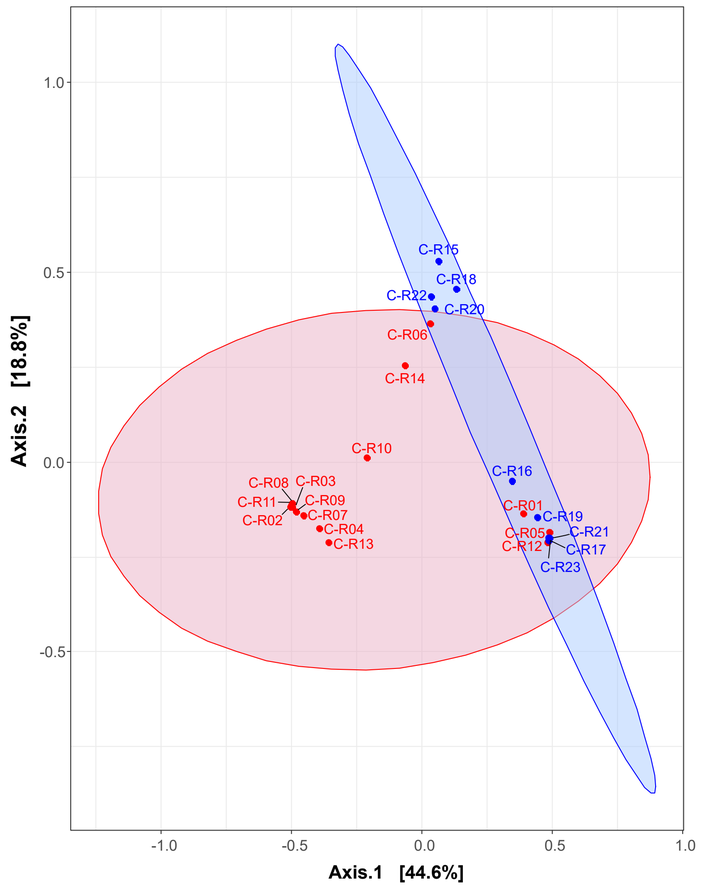


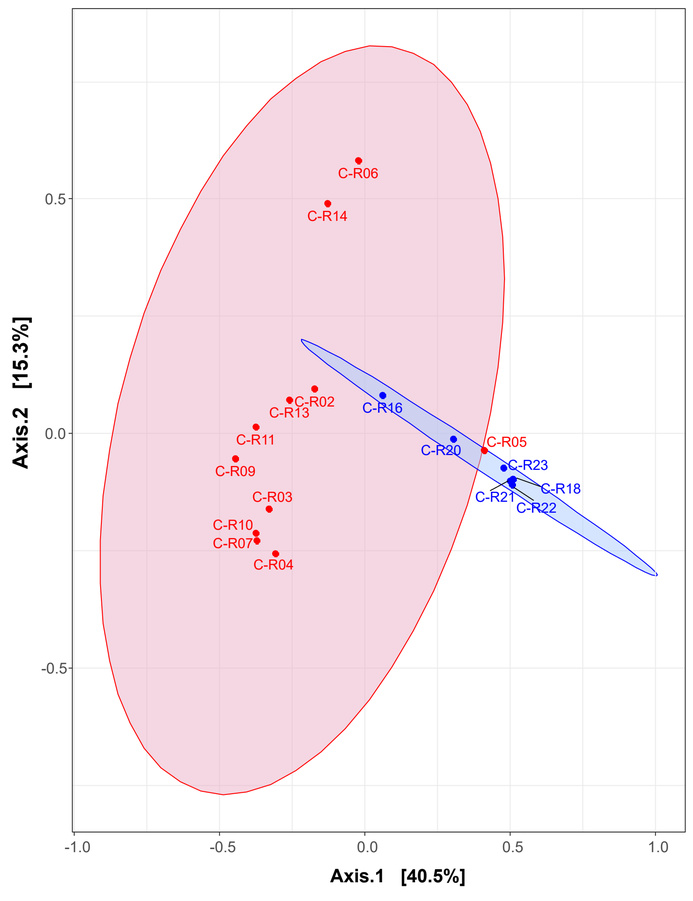












bioRxiv preprint doi: https://doi.org/10.1101/2022.03.14.484326; this version posted March 14, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.