Metatranscriptomics of cheese microbial communities: Efficiency of RNA extraction from various cheese types and of mRNA enrichment

Christophe Monnet, Célia Carbonne, Karine Labadie, Corinne Cruaud, Elodie Brun, Valérie Barbe

This is the submitted version of the manuscript published in International Journal of Food Microbiology (https://doi.org/10.1016/j.ijfoodmicro.2022.109701)

Metatranscriptomics of cheese microbial communities: efficiency of RNA extraction from various cheese types and of mRNA enrichment

Célia Carbonne[a,1], Karine Labadie[b], Corinne Cruaud[b], Elodie Brun[b], Valérie Barbe[c], Christophe Monnet[a,*]

[a] Université Paris-Saclay, INRAE, AgroParisTech, UMR SayFood, 1 avenue Lucien Brétignières, 78850 Thiverval-Grignon, France

[b] Genoscope, Institut François Jacob, CEA, Université Paris-Saclay, 2 Rue Gaston Crémieux, 91057 Evry, France

[c] Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Université Evry, Université Paris-Saclay, 91057 Evry, France

[1] Present address: INRAE, UMR Micalis, Domaine de Vilvert, 78350 Jouy-en-Josas, France

*Corresponding author. E-mail address: christophe.monnet@inrae.fr

Abstract

Microbial communities from cheeses contribute to the development of typical organoleptic properties. Metatranscriptomic analyses can be used to provide a global picture of the functioning of these communities. Our objective was to evaluate the efficiency of RNA extraction from various cheese types and to evaluate mRNA enrichment procedures for metatranscriptomic analyses. For the 32 tested cheese brands, corresponding to five cheese types, the extraction yield varied from 1 µg to 363 µg RNA per gram of cheese and, overall, the yield was lower for fresh cheeses and for the core of pressed cooked cheeses than for the other cheese types. Pressed cooked cheeses also had a lower RNA integrity than the other cheese types. For total RNA extracts from four cheeses, approximately 99% of the sequencing reads corresponded to ribosomal RNA, and mRNA enrichment by RiboPOOL and FastSelect kits decreased this percentage to a range of 75 to 97% and of 53 to 76%, respectively. Comparison of RNA libraries after mRNA enrichment with libraries of undepleted total RNA showed that the FastSelect mRNA enrichment had a lower impact on the gene expression profiles of five target cheese species than the riboPOOL kit and the oligo (dT) selection method. The procedures that we describe in the present study may be useful for metatranscriptomic analysis of various cheese types.

Keywords

Cheese ; RNA extraction ; rRNA depletion ; RNA sequencing ; Metatranscriptome ; Microbial community

Abbreviations: FC: fresh cheese; SUC: soft unpressed cheese; PUC: pressed uncooked cheese; PCC: pressed cooked cheese; BVC: blue veined cheese; PCC: Pearson's correlation coefficient

1. Introduction

During the cheese production process, microorganisms contribute to the development of the typical sensory properties of the final product. They form microbial communities with varying population structures, depending on the cheese type, manufacturing practices, location on the cheese (surface or core), and manufacturing time. These communities include deliberately inoculated microorganisms, strains from milk, brine baths and salt, and strains from the cheese-manufacturing environment itself (Donnelly, 2014; Irlinger et al., 2015; Jonnala et al., 2018; Montel et al., 2014). Gene expression studies help to better understand the growth and activity of these cheese microbial communities. Expression of selected genes can be quantified by reverse transcription real-time PCR, and a more global picture of the functioning of the microbial community can be provided by metatranscriptomic analyses (De Filippis et al., 2016; Dugat-Bony et al., 2019, 2015; Duru et al., 2018; Lessard et al., 2014; Monnet et al., 2016; Ojala et al., 2017; Pangallo et al., 2019; Pham et al., 2019). This last approach is becoming increasingly available to researchers due to continuous technical advances and the reduced cost of high-throughput sequencing technologies.

There are several issues concerning the application of metatranscriptomic analyses to cheeses or model cheeses. One of them is the extraction of RNA from the cheese matrix. Two types of methods can be distinguished: those based on the direct extraction from the cheese and those where cells are separated from the cheese matrix prior to RNA extraction (Kase and Pfefer, 2016; Monnet and Bogovic Matijasic, 2012). The advantage of the latter type is that cells can be concentrated by centrifugation, which results in a higher recovery of RNA. Indeed, metatranscriptomic analyses generally require greater amounts of RNA than reverse transcription real-time PCR analyses. Typically, ~ 5 µg of total RNA is processed when a ribosomal RNA depletion step is performed before RNA sequencing. However, undesired modifications of the cell transcriptomes may occur during cell separation from the cheese matrix. In contrast, with direct extraction procedures, the cellular processes can be stopped at the very beginning of the procedure, stabilizing the mRNA transcripts (Monnet et al., 2008).

Another issue concerning metatranscriptomic analyses of cheeses is the efficiency of mRNA enrichment. Indeed, since most RNA present in cheese RNA extracts is composed of rRNA (>95%), the direct sequencing of these extracts mainly results in sequencing reads that do not match the microbial mRNA transcripts. Fungal mRNA can be sequenced after purification by oligo (dT) selection of the polyadenylated transcripts, but since cheeses also contain bacteria, a complete sequencing of the metatranscriptomes can only be performed after an rRNA depletion step. Such depletion can be performed by hybridization with oligonucleotidic probes linked to magnetic beads and the targeting of conserved regions of rRNA (Chen and Duan, 2012; Petrova et al., 2017). Another type of depletion technology is used in the recently released QIAseq FastSelect™ rRNA removal kit, in which a pool of oligonucleic acids selectively blocks the reverse transcription of rRNA. Due to the presence of a wide diversity of bacterial and fungal species in cheeses, the oligonucleic acids used in these two types of rRNA removal procedures should be able to hybridize to a large variety of rRNA targets, which may be difficult to achieve. In addition, rRNA depletion from cheese may also be complicated by the presence of a large amount of degraded rRNA, part of which lacks the sequences targeted by the depletion probes. Cheeses with long ripening times contain a large proportion of dead cells, leading to a considerable abundance of degraded rRNA in their RNA extracts.

Up until now, the publications describing metatranscriptomic analyses of cheeses performed with direct RNA extraction procedures (i.e., without cell separation from the cheese matrix) concerned samples from the surface of soft cheeses or model cheeses with quite high levels of microbial biomass (Dugat-Bony et al., 2015; Monnet et al., 2016; Pham et al., 2019). The objective of the present study was to evaluate the efficiency of direct RNA extraction from a variety of cheeses, both for surface and core samples. Another objective was to test two commercial rRNA depletion kits containing probes targeting both bacterial and fungal rRNA.

2. Materials and methods

2.1. Cheese sampling and microbiological analyses

The cheeses were purchased on the retail market. They corresponded to fresh cheeses (five brands), soft unpressed cheeses (ten brands), pressed uncooked cheeses (eight brands), pressed cooked cheeses (five brands) and blue-veined cheeses (four brands), which were sampled before their "best-before" date. The cheeses were cut perpendicular to the surface in order to produce three equivalent parts, corresponding to three technical replicates for each cheese brand. The rinds were sampled with a knife, at a thickness of ~ 3-5 mm. The cheese cores were recovered after elimination of all rind remains. The samples were subsequently mixed and homogenized with knives, forks, a mortar and pestle, a cheese grater or a garlic press. For the microbiological analyses, one gram of homogenized sample was mixed with 9 ml of physiological saline solution (9 g/L NaCl) and dispersed with a mechanical blender (Ultra Turrax model T25; Ika Labortechnik, Staufen, Germany) for 1 min at 14,000 rpm. Serial dilutions were then performed in physiological saline and plated onto agar plates. The lactic acid bacteria were enumerated on de Man-Rogosa-Sharpe agar (de Man et al., 1960) (pH 6.5; Biokar Diagnostics, Beauvais, France) supplemented with 50 mg/l amphotericin after 3 days of incubation at 30°C under anaerobic conditions. The aerobic bacteria were enumerated on brain heart infusion agar (Biokar diagnostics) supplemented with 50 mg/l amphotericin after 3 days of incubation at 30°C under aerobic conditions. Lactic acid bacteria are not able to grow on this medium in aerobic conditions, or form very small colonies, which were not considered during colony counting. Fungi were enumerated on yeast extract-glucose-chloramphenicol agar (YGCA) (Biokar Diagnostics) after 3 days of incubation at 25°C (Welthagen and Viljoen, 1997).

2.2. Extraction of RNA from cheese samples and DNase treatment

The RNA extraction procedure is based on a previously described method, which does not involve prior separation of the cells from the cheese matrix (Monnet et al., 2008). Some modifications of this

129 method were done in order to increase the extraction yields and the amount of sample processed. In this
130 procedure, the RNA present in the samples is stabilized after the bead-beating treatment in the presence of
131 guanidine thiocyanate and phenol, and the corresponding tubes can then be stored for several months at -
132 80°C before performing the subsequent steps of the procedure. The extractions were performed on the same
133 samples as those used for the microbiological analyses, taking care to begin the extractions immediately
134 after the end of the homogenization step. Two 7-ml bead-beating tubes (ref. P000944-LYSK0-A; Bertin
135 Technologies, Montigny-le-Bretonneux, France) containing 1.6 g of 0.1-mm-diameter zirconium beads
136 (Biospec Products, Bartlesville, OK, USA) and 1.6 g of 0.5-mm-diameter beads were used for each sample.
137 After the addition of ~ 450-500 mg of homogenized cheese in each tube, 5 ml of ice-cold Uptizol reagent
138 (Interchim, Montluçon, France) was added and mixing was performed immediately on a Precellys Evolution
139 bead beater (Bertin Technologies) using two 20-s mixing sequences at a speed of 10,000 rpm. The tubes
140 were cooled on ice for 5 min after each mixing sequence. They were then stored at -80°C, and the extraction
141 procedure was carried out later. After thawing in a cold water bath (~ 4°C), the content of the two tubes was
142 pooled in a 15-ml RNase-free centrifugation tube (Corning ref. 430766; Sigma-Aldrich, St. Louis, MO,
143 USA) and centrifugation was performed for 10 min at 3,500 x g and 4°C. The supernatant was transferred to
144 a new 15-ml tube, taking care to eliminate the fat layer present at the top of the liquid phase. The tube was
145 then incubated for 5 min at room temperature before the addition of a volume of chloroform equivalent to
146 20% of the volume of the Uptizol-cheese mixture. The tube was shaken vigorously by hand for 15 s and
147 subsequently incubated for 3 min at room temperature, and then for 2 min on ice. After centrifugation for 15
148 min at 3,500 x g and 4°C, most of the upper aqueous phase was recovered in a new 15-ml tube, taking care
149 not to recover any part of the organic phase or of the impurities present between the two phases. A volume
150 of phenol-chloroform-isoamyl alcohol (125/24/1) solution, pH 4.7 (ref. P1944; Sigma-Aldrich), equivalent
151 to 100% of the aqueous phase volume, was added and the tube was subsequently shaken vigorously by hand
152 for 15 s. After centrifugation for 10 min at 10,000 x g and 4°C, the aqueous phase (volume equal to
153 approximately 2 ml) was transferred to a new 15-ml tube, taking care not to recover any part of the organic
154 phase or of the impurities present between the two phases. An additional RNA purification step was then
155 performed on a silica membrane-based column (RNeasy Mini Kit; Qiagen, Courtaboeuf, France) using the
156 following procedure, which included an on-column DNase digestion step. The tube was equilibrated at room

temperature and a volume of absolute ethanol equivalent to 55% of the aqueous phase volume was added. After mixing, 700 µl was loaded on the RNeasy spin column and centrifugation was performed for 15 s at 15,000 x g. After elimination of the flow through, the remainder of the sample was loaded onto the column and treated in the same way until the entire sample was used. A quantity of 350 µl of RW1 buffer (Qiagen) was then loaded onto the column, and the tube was centrifuged for 15 s. A quantity of 80 µl of DNase/RDD buffer 1/7 solution (Qiagen) was subsequently loaded onto the column, and after an incubation of 15 min, 350 µl of RW1 buffer (Qiagen) was added and the tube was centrifuged for 15 s. A quantity of 350 µl of RW1 buffer was added again, and after an incubation of 5 min, the tube was centrifuged for 15 s. Two washing steps were then performed with 500 µl of RPE buffer (Qiagen), and the tube was subsequently centrifuged for 1 min in order to remove all traces of RPE buffer. The RNA was recovered after the addition of 30 µl of RNase-free water, incubation for 2 min and centrifugation for 1 min. Purified RNA was quantified with Qubit RNA assay kits on the Qubit 3.0 fluorimeter (Life Technologies, Carlsbad, CA, USA), and RNA quality was analyzed with an Agilent model 2100 Bioanalyzer (Agilent, Palo Alto, CA, USA) using RNA 6000 Nano chips, according to the manufacturer's instructions. If necessary, greater amounts of RNA can be obtained by pooling samples after the phenol-chloroform-isoamyl alcohol extraction step, provided that the amount of RNA loaded onto the silica membrane-based column is lower than 100 µg. A second DNase treatment was performed on 15 µg of purified RNA using the TURBO DNA-free kit (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions, with 2 units of DNase enzyme in a reaction volume of 50 µl.

2.3. mRNA enrichment and RNA sequencing

Ribosomal RNA depletion with the QIAseq FastSelect™ rRNA removal kit (Qiagen) was performed on 100 ng total RNA using a combination of "FastSelect 5S/16S/23S" and "FastSelect - yeast" oligonucleic acids. Depletion was performed according to the manufacturer's instructions (QIAseq FastSelect 5S/16S/23S Handbook), using 1 µl for each of the two oligonucleic acids solutions and the NEBNext Ultra II directional RNA library preparation kit (New England Biolabs, MA, USA). RNA fragmentation was performed for 90 s

at 89°C, and the incubation step at 42°C during the first strand cDNA synthesis was set to 60 min. NEBNext adaptors were diluted 25 times for adaptor ligation and 16 cycles were applied for PCR enrichment of adaptor ligated DNA.

Ribosomal RNA depletion with the RiboPOOL kit (siTOOLs Biotech GmbH, Planegg, Germany) was performed on 4 µg total RNA using a combination of the probe mixtures "Pan bacteria 026" and "Filamentous fungi 006". Depletion was performed according to the manufacturer's instructions, using 100 pmoles for each probe. After rRNA depletion, the RNAs were purified using the Zymo Research RNA clean and concentrator kit (ref. R1015; Ozyme, Saint-Cyr-l'École, France), with the option to keep only the large RNAs (>200 nt). Quantity and quality of rRNA-depleted RNA were assessed with Qubit RNA assay kits and on Agilent Bioanalyzer RNA 6000 Pico chips. Illumina RNAseq libraries were then prepared from 10 ng rRNA-depleted RNA using the NEBNext Ultra II directional RNA library preparation kit, with the same protocol modifications as mentioned above.

Oligo dT selection was performed using the NEBNext Poly(A) mRNA magnetic isolation module and libraries were prepared with the NEBNext Ultra II directional RNA library preparation kit, following the same modifications mentioned above.

Ready-to-sequence Illumina RNA libraries were then quantified by qPCR and library profiles evaluated with an Agilent 2100 Bioanalyzer. Each library was sequenced using 100 bp single end read chemistry on a NovaSeq 6000 Illumina sequencer. After Illumina sequencing, an in-house quality control process was applied to the reads that passed the Illumina quality filters. The first step discards low-quality nucleotides ($Q < 20$) from both ends of the reads. Next, Illumina sequencing adaptors and primer sequences were removed from the reads. The longest sequence without adaptors and low-quality bases was kept. Sequences between the second unknown nucleotide (N) and the end of the read were also trimmed. Then, reads shorter than 30 nucleotides (after trimming) were discarded. These trimming and removal steps were achieved using in-house-designed software, called fastx_clean, based on the FastX library. The last step identifies and discards read pairs that mapped to the Enterobacteria phage PhiX174 genome, used as a run quality control, with the bowtie2 aligner v2.2.9 (Langmead and Salzberg, 2012) and the phiX reference sequence (GenBank accession number NC_001422.1).

2.4. Extraction of RNA from single strain liquid cultures

The strains and growth conditions are reported in Table 1. After two successive liquid cultures for 48 hours (5% inoculation rate), a volume of 10 ml (bacteria) or 20 ml (fungi) was centrifuged for 10 min at 10,000 x g and 4°C, and the cell pellet was recovered with 10 ml ice-cold Uptizol. The cell suspension was transferred into two 7-ml bead-beating tubes containing zirconium beads, and RNA extraction was then performed as described for the cheese samples. Cell counts at the end of culture were determined by colony counting on agar plates. Cell concentration was also estimated by measurement of the biomass dry weight. For that purpose, part of the culture medium at the end of growth was recovered by filtration (0.2 µm), and the biomass weight was measured after drying for 24 h at 80°C.

2.5. Mapping against reference genomes

Ribosomal RNA sequences (5S, 5.8S, 16S, 18S, 23S, 28S, chloroplast rRNA and mitochondrial rRNA) were filtered using SortMeRNA 2.0 (Kopylova et al., 2012). The non-rRNA reads were then processed using FASTX-Toolkit software (available at http://hannonlab.cshl.edu/fastx_toolkit/) to reduce the size of the sequences to the first 35 bases. The sequences were mapped to the genomes of one representative of five microbial species frequently present in cheeses: *Geotrichum candidum* CLIB918, *Lactococcus lactis* IL1403, *Streptococcus thermophilus* LMD9, *Debaryomyces hansenii* CBS767 and *Penicillium roqueforti* F164 (NCBI bioproject accession numbers PRJEB5752, PRJNA57671, PRJNA13773, PRJNA12410, and PRJEB4023, respectively). Mapping was performed using Bowtie 1.2.3 (Langmead et al., 2009) with the following parameters: -a -m - --best --strata -v2 -t -S. The number of reads that mapped onto CDS of the reference genomes was counted using HTSeq-count version 0.10.1 (Kopylova et al., 2012) with the following parameters: -s reverse -t gene -i locus_tag -m union. The sequences of the reads for all the samples after SortMeRNA analysis were deposited in the European Nucleotide Archive of the European Bioinformatics Institute under the accession numbers ERR6891699, ERR6897440 to ERR6897441, ERR6897486 to ERR6897490, and ERR6897492 to ERR6897495.

3. Results

3.1. RNA extraction yield and RNA integrity from various cheeses

In order to test the efficiency of the RNA extraction method described in the Materials and Methods section, we selected 32 cheeses from five cheese families, corresponding to fresh, soft unpressed, pressed uncooked, pressed cooked and blue-veined cheeses (designated as FC, SUC, PUC, PCC and BVC, respectively). For the fresh and blue-veined cheeses, RNA was extracted from the core of the cheeses. For the soft unpressed, pressed uncooked and pressed cooked cheeses, RNA was extracted from the core and also from the surface, except for the cheeses whose surface was covered by a layer of paper or paraffin wax. The mean standard error of the RNA extraction yields of the three technical replicates performed for each sample was 36% (Fig. 1). The yields varied considerably between the samples, even for those belonging to the same cheese family. The highest yields were obtained for SUC-9, a soft unpressed goat cheese (363 and 238 µg RNA/g cheese for surface and core, respectively), and the lowest for PCC-4, a Parmigiano Reggiano PDO cheese (4 and 1 µg RNA/g cheese for surface and core, respectively). Mean extraction yields were higher for the surface than for the core samples (160 vs. 64 µg RNA/g cheese). Overall, the extraction yield was lower for the core samples of fresh and pressed uncooked cheeses than for the other cheese families (Fig. 2). We expected that the RNA extraction yields would increase with the microbial load of the cheeses. However, for both core and surface samples, there was not a strong correlation between the RNA extraction yields and the fungal and bacterial counts (Fig. 3). Several factors may explain this poor correlation, such as the presence of RNA from dead cells or differences in the RNA content and RNA extraction efficiency from the cells of the various microbial species present in the cheeses.

In order to investigate the differences between microbial species, we determined the RNA extraction yields from pure cultures grown in semi-defined media. The RNA extraction yield, expressed as the amount of RNA recovered per colony forming unit (Fig. 4A), varied considerably between the microbial species. Overall, the yields were higher for the fungi than for the bacteria. The highest efficiencies were observed for

Penicillium camemberti and *Geotrichum candidum*, which is probably due to the formation of filaments by these fungi, resulting in a smaller amount of colony forming units per amount of biomass. The large variability of the extraction yields between the microbial species may thus contribute to the poor correlation between the microbial counts and the amounts of RNA recovered from the cheeses. Differences in extraction yields were much lower when the yield was expressed as the amount of RNA recovered per mg of biomass dry weight (Fig. 4B).

Integrity of the RNA preparations from cheeses was evaluated by capillary electrophoresis. With the RNA extraction method used in the present study, the cellular processes and the activity of the RNases present in the samples are stopped at the very beginning of the procedure by the addition of guanidinium thiocyanate-phenol solution and bead-beating. The presence of peaks resulting from the degradation of rRNAs, which are the main constituents of the cellular RNAs, is thus the result of RNA degradations that occur in cheese prior to RNA extraction. The surfaces of the bacterial 16S and 23S, and fungal 18S and 26S rRNAs were determined from the electrophoregrams, and the proportion of the corresponding peaks was used as an indicator of RNA integrity (Additional file 1: Table S1). This proportion varied considerably between the samples, even for those belonging to the same cheese family. However, overall, the RNA integrity was lower for the pressed cooked cheeses than for the other cheese families. Examples of RNA electrophoregrams corresponding to different levels of RNA integrity are presented in Fig. S1 (Additional file 1). A large fraction of RNA corresponded to intact rRNA in samples SUC-3, SUC-8 and FC-2. This fraction was lower for samples PUC-6 and BVC-2, and nearly no intact rRNA was found for the pressed cooked cheese sample, PCC-2.

3.2. mRNA enrichment of cheese RNA preparations

The aim of depleting ribosomal RNA from the cheese RNA preparations is to improve the detection and quantification of mRNA by RNA sequencing. Commercialization of the Ribo-Zero rRNA depletion kits (Illumina) was stopped in 2018 and, to our knowledge, only the riboPOOL rRNA depletion kit and the

QIAseq FastSelect rRNA removal kit are currently commercially available for depletion of both bacterial and fungal rRNA. We tested these two kits for RNA preparations of cheeses BVC-2 (core), SUC-8 (surface), and PUC-6 (surface), and for a mixture of ten different RNA preparations (surfaces of SUC-1, SUC-3, SUC-4, SUC-6, SUC-9, PUC-4, PUC-6 and PCC-1, and the cores of BVC-2 and BVC-4). For the two kits, a combination of fungal and bacterial depletion probes was used, as described in Materials and Methods. We also purified eukaryotic mRNA from these RNA samples by oligo (dT) selection of the polyadenylated transcripts.

RNA sequencing was then performed on the initial total RNA preparations (approx. 300 million sequencing reads per sample) and on the samples after rRNA depletion or oligo (dT) selection of polyadenylated transcripts (approx. 80 million sequencing reads per sample) (Additional file 1: Table S2). When no mRNA enrichment was performed, ribosomal RNA represented nearly all the sequences, ranging from 98.7 to 99.2% of the sequencing reads (Fig. 5). Ribosomal RNAs assigned to eukaryotes (18S and 28S) were dominant for cheeses SUC-8 and PUC-6, whereas those assigned to prokaryotes (16S and 23S) were dominant for BVC-2. Prokaryotic and eukaryotic ribosomal sequences were nearly equivalent for the mixture of ten cheese RNA preparations. After oligo (dT) selection, most sequencing reads (67 to 91%) corresponded to non-rRNA sequences. Treatment of total RNA by the FastSelect or the riboPOOL rRNA removal kit increased the proportion of non-rRNA sequences, but the latter kit was less efficient for the sample BVC-2 and the mixture of ten cheese RNAs. This was related to a lower depletion efficiency of the riboPOOL kit for the prokaryotic rRNA (16S and 23S), which were more abundant in these two samples. For all the samples, the FastSelect kit was more effective for removing 18S rRNA than 28S rRNA.

In order to compare the metatranscriptomes generated from the initial total RNA extracts and from the three mRNA enrichment procedures, we mapped the sequencing reads to the genomes of five microbial species frequently present in cheeses (Table 2). As expected, when no mRNA enrichment was performed, only a very low proportion of the reads mapped to microbial CDSs. The highest CDS abundances were obtained for *P. roqueforti* in cheese sample BVC-2, and for *G. candidum* in cheese sample SUC-8 (0.42% and 0.38% of the sequencing reads, respectively). After oligo (dT) selection of SUC-8, the *G. candidum*

CDS abundance represented 53% of the reads, indicating that this species was the dominant fungi of this cheese. As expected, for the four cheese RNA samples, the transcriptomes of the bacterial species (*L. lactis* and *St. thermophilus*) represented only a low proportion of the sequencing reads after oligo (dT) selection (the maximum was 0.15% for *St. thermophilus* CDSs in cheese SUC-8). For samples SUC-8 and PUC-6, the abundance of fungal CDSs was higher when rRNA was depleted with FastSelect than with riboPOOL, and the opposite was observed for the bacterial CDSs. For the two other cheese samples, the abundance of both bacterial and fungal CDSs was higher with FastSelect than with riboPOOL, which is in accordance with the fact that for these samples, which contained more bacterial 16S and 23S rRNA than the two others, rRNA depletion was less efficient with the riboPOOL kit.

In order to evaluate the impact of the mRNA enrichment procedure on the microbial transcriptomes, we determined the correlations between the CDS hit numbers obtained after mRNA enrichment and the CDS hit numbers obtained by sequencing the undepleted total RNA. Correlations were considered only when the cumulated number of CDS hits for the corresponding genome was higher than 0.01% of the sequencing reads. An example of a graphical representation of correlation for *P. roqueforti* in cheese sample BVC-2 is shown in Fig. S2 (Additional file 1). The highest correlation with the transcriptomes generated by total RNA sequencing was obtained after mRNA enrichment with the FastSelect kit. The corresponding Pearson's correlation coefficient (PCC) was between 0.82 and 0.96, with a mean value of 0.90. PCC values with FastSelect were higher than with oligo (dT) selection, even for the transcriptomes of *G. candidum* in sample SUC-8 or in the mixture of cheese samples, or of *P. roqueforti* in sample BVC-2, for which the number of hits after oligo (dT) selection was high (53, 34 and 25% of the sequencing reads, respectively). For the riboPOOL rRNA depletion kit, we observed some differences in the PCC values obtained for the same cheese samples. For example, in cheese PUC-6, the PCC value for the transcriptome of *P. roqueforti* (0.46) was lower than for *D. hansenii* and *St. thermophilus* (0.80 and 0.84).

In the standard protocol of the FastSelect kit that we used in the present study, a volume of 1 µl is used for each of the two oligonucleotidic solutions (5S/16S/23S and yeast), and we wanted to determine if an increase of this volume improves rRNA depletion efficiency. For the four cheese RNA samples,

depletion was more efficient when the oligonucleotidic solution volumes were increased to 4 µl. After sequencing, the proportion of non-rRNA reads reached 65, 74, 90 and 77% for the samples SUC-8, PUC-6, BVC-2 and the mixture of ten cheese RNAs, respectively. However, this also resulted in a strong decrease in the correlations with the transcriptomes generated from undepleted total RNA. For example, for the cheese sample SUC-8, PCC values were 0.68 and 0.71 for *G. candidum* and *L. lactis* reference genomes, respectively (data not shown), which is why we will not select this modified protocol for future experiments.

4. Discussion

Efficient methods for RNA extraction from cheese and for mRNA enrichment will promote the study of cheese microbial communities by metatranscriptomic analyses. In the present study, we tested, for various cheese samples, an RNA extraction method in which the cellular processes are stopped at the very beginning of the procedure, thereby limiting the modifications of the cell transcriptomes during the extraction. The absence of a cell separation from cheese and cell concentration step prior to RNA extraction in this procedure makes it more difficult to recover large amounts of RNA. However, when we processed one gram of cheese, we recovered at least 5 µg of RNA for most cheese samples (45 out of 52), which is sufficient for common downstream rRNA depletion procedures. There were large differences in RNA extraction yields, even for the same cheese family. However, on average, the extraction yields were lower for fresh cheeses and for the core of pressed cooked cheeses than for the other cheese types. Since cheese RNA is exclusively of microbial origin, it can be assumed that the RNA extraction yield increases with the amount of microbial biomass. However, in our samples, there was no strong correlation between the RNA extraction yields and the bacterial and fungal counts. This can be explained by factors such as the presence of RNA from dead cells, differences in the cell RNA content of the microbial species, or differences in cell resistance to the bead-beating treatment and extraction methods. In addition, for filamentous fungi such as *P. camemberti*, cell counting on agar plates is not a valid method for estimating the amount of biomass since the colonies mainly correspond to the spores (Leclercq-Perlat et al., 2004). In the present study, we analyzed cheeses only after the manufacturing process. It can be assumed that, due to the low microbial growth rate , it would

be more difficult to recover large amounts of RNA for cheeses sampled at the beginning of the manufacturing process.

We observed large differences in the levels of rRNA integrity between the samples, even for the same cheese family. On average, the RNAs extracted from pressed cooked cheeses had lower levels of integrity than for the other types of cheeses. Since the cellular processes and the activity of the RNases are stopped at the very beginning of the extraction procedure, low RNA integrity is not due to the extraction method but, instead, to cell lysis during the ripening or storage of the cheeses (Broome et al., 2011). Even if the presence of high amounts of degraded rRNA does not, *per se*, preclude RNA sequencing, the corresponding samples have a lower microbial fraction with an active transcription activity, which probably has a detrimental impact on the detection threshold for metatranscriptomic analyses.

An efficient mRNA enrichment procedure from cheese RNA is necessary in order to reduce the proportion of sequencing reads that do not correspond to bacterial and fungal mRNA transcripts. The Ribo-Zero™ kit for bacteria was shown to be effective to enrich non-rRNA transcripts in human stool samples (Giannoukos et al., 2012) and bacterial biofilms (Petrova et al., 2017). In recent years, this kit, or a combination of this kit with the Ribo-Zero™ kit for yeasts, was used for several cheese transcriptomic studies (Dugat-Bony et al., 2019; Duru et al., 2018; Monnet et al., 2016; Ojala et al., 2017; Pham et al., 2019), but the company stopped the sale of these kits at the end of 2018. In the present study, we tested the riboPOOL bacterial and fungal depletion kits that use the same principle of probes to selectively capture rRNA. We also tested the recently released FastSelect kits that use oligonucleic acids that selectively block the reverse transcription of rRNA. The bacterial version (5S/16S/23S) of this kit was used in a recent study for the metatranscriptomic analysis of a fecal sample (Galata et al., 2021). When we performed RNA sequencing on samples with no mRNA enrichment, approximately 99% of the reads corresponded to rRNA. One possible explanation for the very high proportion of rRNA in our total RNA samples is the fact that the cheeses were sampled after their manufacturing, at a stage where there is no active cell growth. Another reason could be the presence in the cheeses of dead cells, which contain more rRNAs in comparison to mRNAs, because the latter are less stable.

In the present study, the riboPOOL and FastSelect rRNA depletion kits considerably increased the proportion of non-rRNA sequencing reads, with an equivalent performance for two of the cheese samples, and a higher performance for FastSelect on the two other samples. For mouse hippocampal tissue, treatment with the FastSelect HMR (human, mouse and rat) kit resulted in less than 1.5% rRNA in the final library (Potemkin et al., 2021). The fact that large amounts of rRNA were still present in our mRNA-enriched samples (from 53 to 76% of the sequencing reads for FastSelect and 75 to 97% for riboPOOL) may be due to the low integrity of the cheese rRNA and/or insufficient hybridization efficiency on rRNA from some cheese species. One interesting feature of the FastSelect kit is that it requires only a small amount of input RNA (100 ng for the libraries produced in the present study), which makes it possible to sequence the metatranscriptomes from cheese samples with lower RNA content than what is required when capture probes are used. After oligo (dT) selection, the proportion of rRNA in the sequencing libraries (9 to 33% of the sequencing reads) was much lower than after rRNA depletion with riboPOOL or FastSelect. However, one major drawback of oligo (dT) selection is that the bacterial part of the metatranscriptomes cannot be obtained by this technique.

Metatranscriptomic studies may consist of functional or taxonomical profiling of sequencing reads or of CDS contigs generated by the assembly of the sequencing reads (Shakya et al., 2019). In some cases, it is also possible to perform differential gene expression analysis for the dominant species in order to determine their under- or over-expressed genes. For that purpose, the sequencing reads corresponding to the target species are extracted from the RNA sequencing data, and differential gene expression analyses in different biological conditions are performed by statistical analysis with software such as DESeq2 (Love et al., 2014). For this second type of analysis, in particular, the impact of the mRNA enrichment procedure on the gene expression profiles should be as limited as possible. Comparison of the CDS hit numbers for selected cheese microbial species showed that the FastSelect treatment had a lower impact than the riboPOOL treatment or the oligo (dT) selection, and the Pearson's correlation coefficient of the FastSelect libraries vs. the libraries generated from the deep sequencing of undepleted total RNA was between 0.82 and 0.96, depending on the species and cheese samples.

In summary, the present study shows that RNA can be extracted from various cheese types using a method that does not involve prior cell separation from the cheese matrix. After RNA extraction, the rRNA integrity levels are highly variable, depending on the samples, and rRNA accounts for about 99% of the total RNA. This percentage can be decreased using the riboPOOL or the FastSelect rRNA removal kits. For the latter kit, there was, on average, a lower impact of the rRNA depletion procedure on the gene expression profiles of five selected target species in four cheese RNA samples.

Acknowledgements

This work was carried out within the ADAMOS research project, part of the RMT Fromages de Terroirs, a French technological network. It was supported by the Ministry of Agriculture (CASDAR) and by Genoscope, the CEA and France Génomique (ANR-10-INBS-09-08). We are grateful to the INRAE MIGALE Bioinformatics Facility (MIGALE, INRAE, 2020. Migale Bioinformatics Facility, doi: 10.15454/1.5572390655343293E12) for providing computational resources.

References

Broome, M.C., Powell, I.B., Limsowtin, G.K.Y., 2011. Cheese | Starter Cultures: Specific Properties, in: Fuquay, J.W. (Ed.), Encyclopedia of Dairy Sciences (Second Edition). Academic Press, San Diego, pp. 559–566. https://doi.org/10.1016/B978-0-12-374407-4.00067-4

Chen, Z., Duan, X., 2012. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. Methods Mol Biol 733, 93–103. https://doi.org/10.1007/978-1-61779-089-8_7

De Filippis, F., Genovese, A., Ferranti, P., Gilbert, J.A., Ercolini, D., 2016. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci. Rep. 6, 21871. https://doi.org/10.1038/srep21871

de Man, J.C., Rogosa, M., Sharpe, M.E., 1960. A medium for cultivation of lactobacilli. J. Appl. Bacteriol. 23, 130–135. https://doi.org/10.1111/j.1365-2672.1960.tb00188.x

Donnelly, C.W. (Ed.), 2014. Cheese and Microbes. ASM Press, Washington, DC. https://doi.org/10.1128/9781555818593

Dugat-Bony, E., Bonnarme, P., Fraud, S., Catellote, J., Sarthou, A.-S., Loux, V., Rué, O., Bel, N., Chuzeville, S., Helinck, S., 2019. Effect of sodium chloride reduction or partial substitution with potassium chloride on the microbiological, biochemical and sensory characteristics of semi-hard and soft cheeses. Food Res. Int. 125, 108643. https://doi.org/10/gf85g9

Dugat-Bony, E., Straub, C., Teissandier, A., Onésime, D., Loux, V., Monnet, C., Irlinger, F., Landaud, S., Leclercq-Perlat, M.-N., Bento, P., Fraud, S., Gibrat, J.-F., Aubert, J., Fer, F., Guédon, E., Pons, N., Kennedy, S., Beckerich, J.-M., Swennen, D., Bonnarme, P., 2015. Overview of a surface-ripened cheese community functioning by meta-omics analyses. PLoS ONE 10. https://doi.org/10.1371/journal.pone.0124360

Duru, I.C., Laine, P., Andreevskaya, M., Paulin, L., Kananen, S., Tynkkynen, S., Auvinen, P., Smolander, O.-P., 2018. Metagenomic and metatranscriptomic analysis of the microbial community in Swiss-type Maasdam cheese during ripening. Int. J. Food Microbiol. 281, 10–22. https://doi.org/10.1016/j.ijfoodmicro.2018.05.017

Galata, V., Busi, S.B., Kunath, B.J., de Nies, L., Calusinska, M., Halder, R., May, P., Wilmes, P., Laczny, C.C., 2021. Synergistic effect of short- and long-read sequencing on functional meta-omics. bioRxiv 2021.04.22.440869. https://doi.org/10.1101/2021.04.22.440869

Giannoukos, G., Ciulla, D., Huang, K., Haas, B., Izard, J., Levin, J., Livny, J., Earl, A., Gevers, D., Ward, D., Nusbaum, C., Birren, B., Gnirke, A., 2012. Efficient and robust RNA-seq process for cultured bacteria and complex community transcriptomes. Genome Biol. 13, r23. https://doi.org/10.1186/gb-2012-13-3-r23

Hettinga, D.H., Vedamuthu, E.R., Reinbold, G.W., 1968. Pouch Method for Isolating and Enumerating Propionibacteria1. J. Dairy Sci. 51, 1707–1709. https://doi.org/10.3168/jds.S0022-0302(68)87259-5

Irlinger, F., Layec, S., Helinck, S., Dugat-Bony, E., 2015. Cheese rind microbial communities: diversity, composition and origin. FEMS Microbiol. Lett. 362, 1–11. https://doi.org/10.1093/femsle/fnu015

Jonnala, Y., R, B., McSweeney, P.L.H., Sheehan, J.J., Cotter, P.D., 2018. Sequencing of the Cheese Microbiome and Its Relevance to Industry. Front. Microbiol. 9. https://doi.org/10.3389/fmicb.2018.01020

Kase, J.A., Pfefer, T.L., 2016. Nucleic Acid Sample Preparation from Dairy Products and Milk, in: Sample Preparation Techniques for Soil, Plant, and Animal Samples, Springer Protocols Handbooks. Humana Press, New York, NY, pp. 231–244. https://doi.org/10.1007/978-1-4939-3185-9_16

Kopylova, E., Noé, L., Touzet, H., 2012. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics 28, 3211–3217. https://doi.org/10.1093/bioinformatics/bts611

Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9, 357–359. https://doi.org/10.1038/nmeth.1923

Langmead, B., Trapnell, C., Pop, M., Salzberg, S.L., 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10, R25. https://doi.org/10.1186/gb-2009-10-3-r25

Leclercq-Perlat, M.-N., Buono, F., Lambert, D., Latrille, E., Spinnler, H.-E., Corrieu, G., 2004. Controlled production of Camembert-type cheeses. Part I: Microbiological and physicochemical evolutions. J. Dairy Res. 71, 346–354. https://doi.org/10.1017/s0022029904000196

Lessard, M.-H., Viel, C., Boyle, B., St-Gelais, D., Labrie, S., 2014. Metatranscriptome analysis of fungal strains *Penicillium camemberti* and *Geotrichum candidum* reveal cheese matrix breakdown and potential development of sensory properties of ripened Camembert-type cheese. BMC Genomics 15, 235. https://doi.org/10.1186/1471-2164-15-235

Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15, 550. https://doi.org/10.1186/s13059-014-0550-8

Monnet, C., Bogovic Matijasic, B., 2012. Application of PCR-based methods to dairy products and to non-dairy probiotic products, in: Hernandez-Rodriguez, P., Ramirez Gomez, A.P. (Eds.), Polymerase Chain Reaction. InTech, Rijeka, pp. 11–51.

Monnet, C., Dugat-Bony, E., Swennen, D., Beckerich, J.-M., Irlinger, F., Fraud, S., Bonnarme, P., 2016. Investigation of the Activity of the Microorganisms in a Reblochon-Style Cheese by Metatranscriptomic Analysis. Front. Microbiol. 7. https://doi.org/10.3389/fmicb.2016.00536

Monnet, C., Ulvé, V., Sarthou, A.-S., Irlinger, F., 2008. Extraction of RNA from cheese without prior separation of microbial cells. Appl. Environ. Microbiol. 74, 5724–5730. https://doi.org/10.1128/aem.00716-08

Montel, M.-C., Buchin, S., Mallet, A., Delbes-Paus, C., Vuitton, D.A., Desmasures, N., Berthier, F., 2014. Traditional cheeses: Rich and diverse microbiota with associated benefits. Int. J. Food Microbiol. 177, 136–154. https://doi.org/10.1016/j.ijfoodmicro.2014.02.019

Ojala, T., Laine, P.K.S., Ahlroos, T., Tanskanen, J., Pitkänen, S., Salusjärvi, T., Kankainen, M., Tynkkynen, S., Paulin, L., Auvinen, P., 2017. Functional genomics provides insights into the role of Propionibacterium freudenreichii ssp. shermanii JS in cheese ripening. Int. J. Food Microbiol. 241, 39–48. https://doi.org/10.1016/j.ijfoodmicro.2016.09.022

Pangallo, D., Kraková, L., Puškárová, A., Šoltys, K., Bučková, M., Koreňová, J., Budiš, J., Kuchta, T., 2019. Transcription activity of lactic acid bacterial proteolysis-related genes during cheese maturation. Food Microbiol. 82, 416–425. https://doi.org/10/gf2h88

Petrova, O.E., Garcia-Alcalde, F., Zampaloni, C., Sauer, K., 2017. Comparative evaluation of rRNA depletion procedures for the improved analysis of bacterial biofilm and mixed pathogen culture transcriptomes. Sci. Rep. 7. https://doi.org/10.1038/srep41114

Pham, N.-P., Landaud, S., Lieben, P., Bonnarme, P., Monnet, C., 2019. Transcription Profiling Reveals Cooperative Metabolic Interactions in a Microbial Cheese-Ripening Community Composed of Debaryomyces hansenii, Brevibacterium aurantiacum, and Hafnia alvei. Front. Microbiol. 10. https://doi.org/10/gf6nbh

Potemkin, N., Cawood, S.M.F., Treece, J., Guévremont, D., Rand, C.J., McLean, C., Stanton, J.-A.L., Williams, J.M., 2021. A workflow for simultaneous detection of coding and non-coding transcripts by ribosomal RNA-depleted RNA-Seq. bioRxiv 2021.01.04.425201. https://doi.org/10.1101/2021.01.04.425201

539 Shakya, M., Lo, C.-C., Chain, P.S.G., 2019. Advances and Challenges in Metatranscriptomic Analysis.
540 Front. Genet. 10. https://doi.org/10/ggdc5m
541 Welthagen, J.J., Viljoen, B.C., 1997. Comparison of ten media for the enumeration of yeasts in dairy
542 products. Food Res. Int. 30, 207–211. https://doi.org/10.1016/S0963-9969(97)00044-6
543
544
545

Figures / Tables

Fig. 1

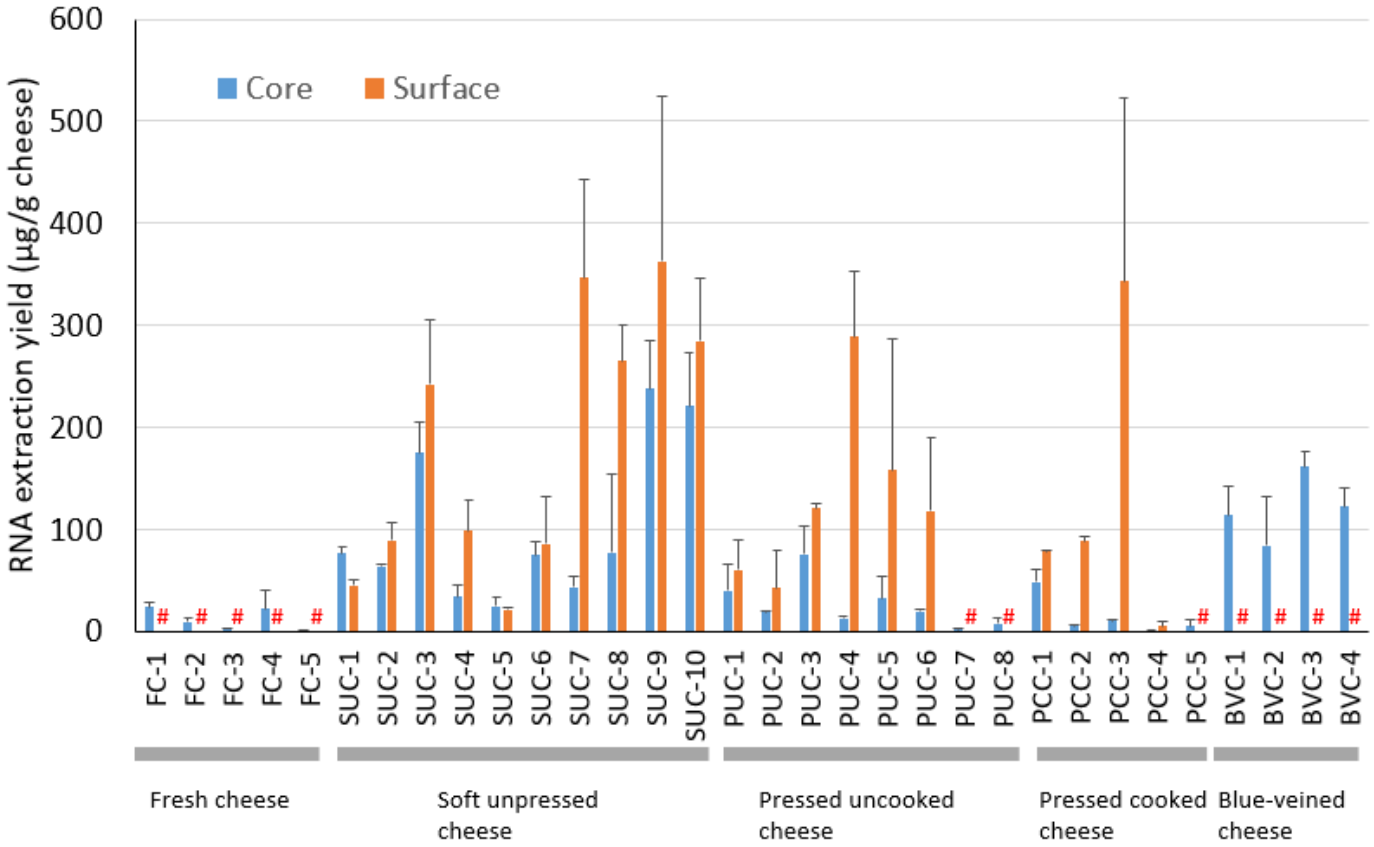
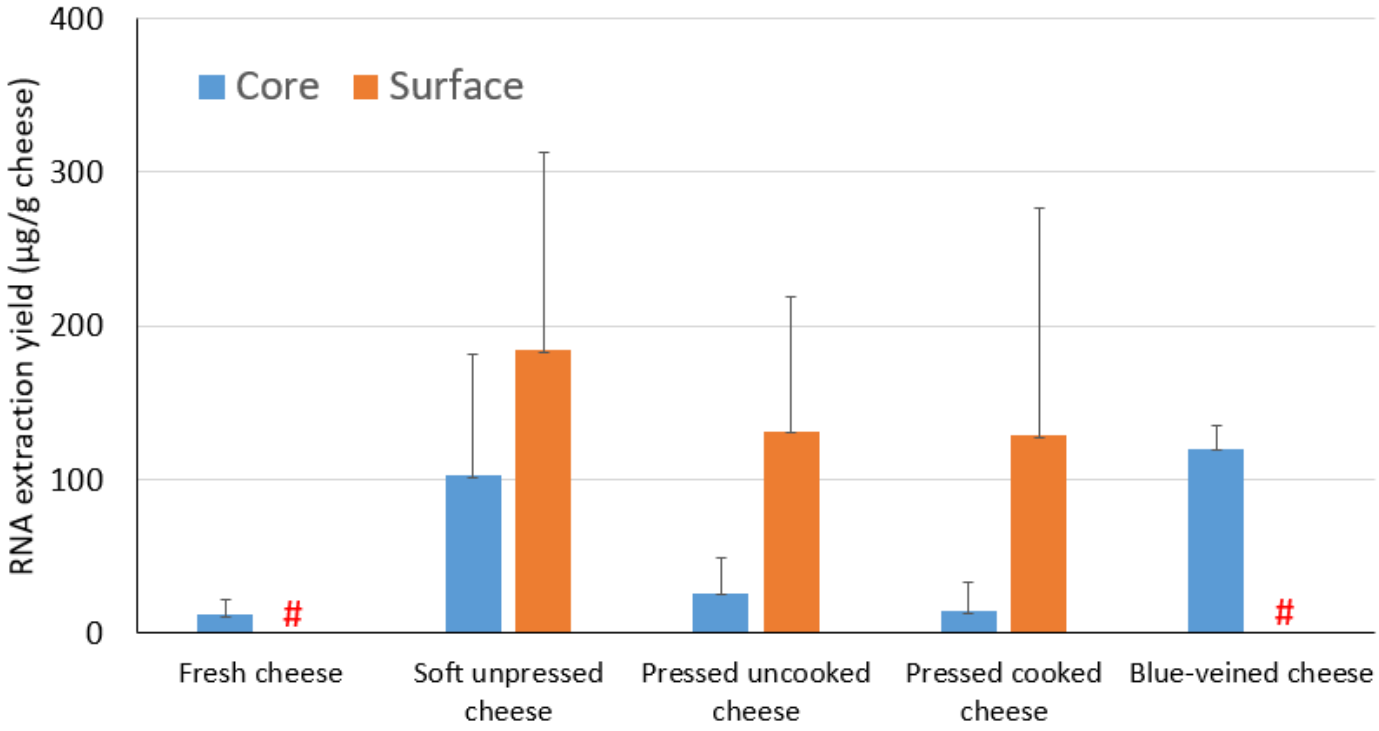


Fig 1. RNA extraction yields from the core or the surface of 32 cheese brands. The error bars represent the standard deviations for three technical replicates. Samples for which the surface was not analyzed are indicated by a # symbol.

555 **Fig. 2**



556

557 **Fig 2.** Mean RNA extraction yields from the five cheese families. The error bars represent the standard

558 deviation within each cheese family. No surface samples were analyzed for two of the cheese families

559 (indicated by a # symbol).

560

Fig. 3

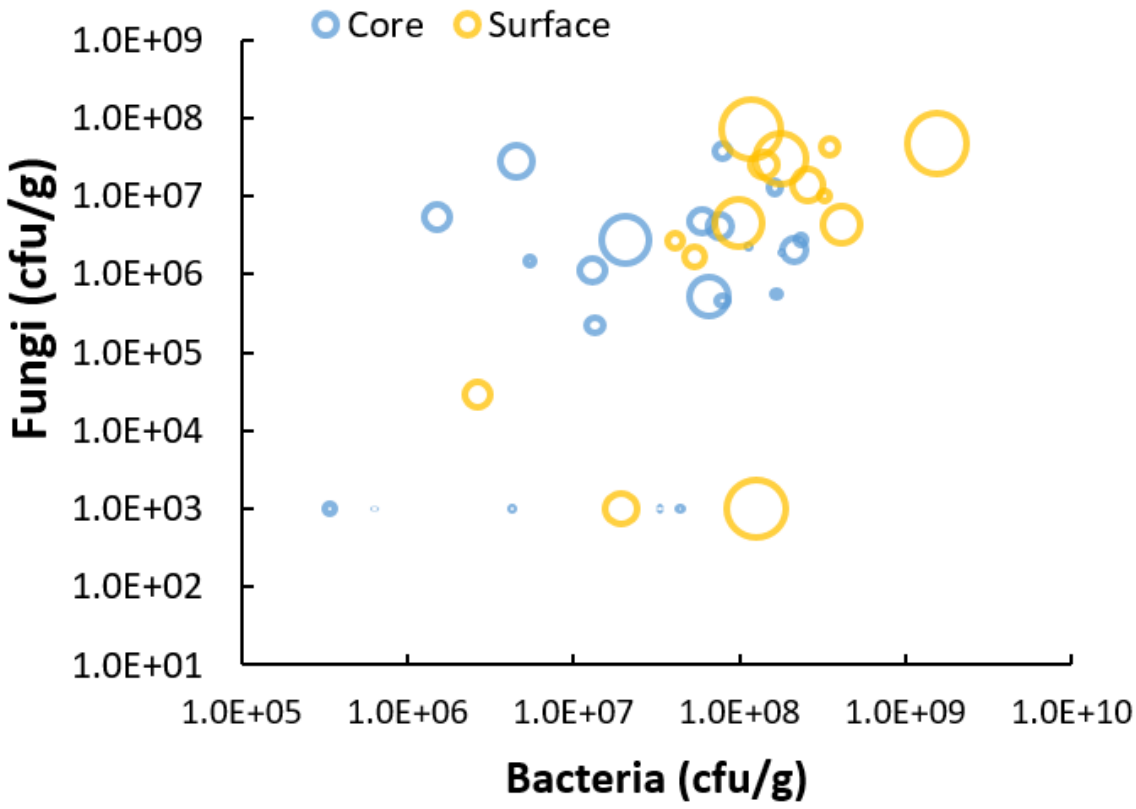


Fig. 3. Graphic representation of fungal counts, bacterial counts (sum of lactic acid bacteria and aerobic bacteria) and RNA extraction yields of the cheese samples. The surface area of the marks is proportional to the RNA extraction yield. The samples for which fungal colony abundance was below the quantification threshold were assigned to a fungal level of $10^{E}3$ cfu/g.

Fig. 4

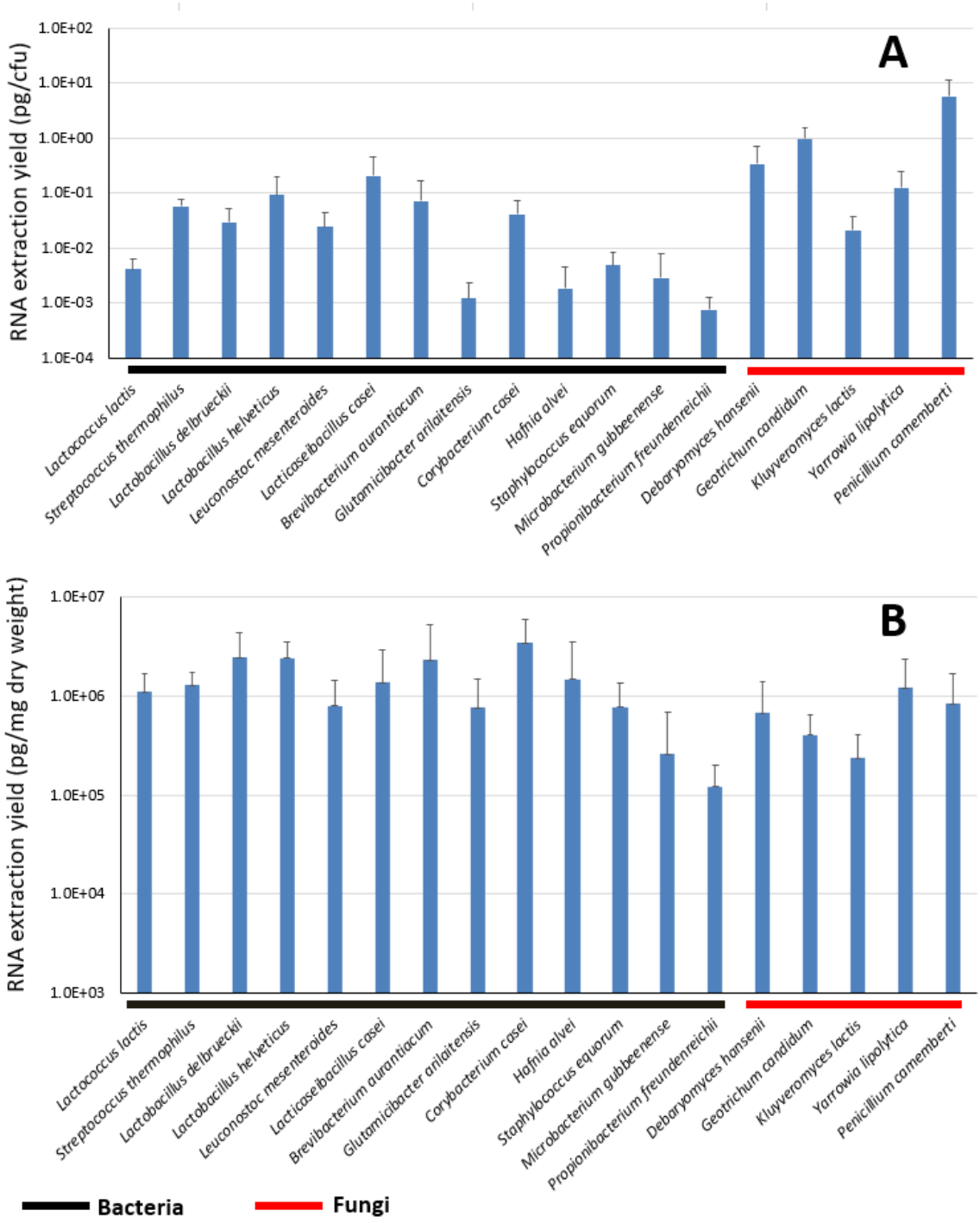


Fig. 4. RNA extraction yields, expressed as the amount of RNA per cfu (A) or per mg biomass dry weight (B) for various cheese-associated species. For each species, one strain was grown as pure culture in a semi-defined medium, and RNA was extracted from the biomass using the same extraction method as the one used for cheese. The error bars represent the standard deviations for three separate cultures.

Fig. 5

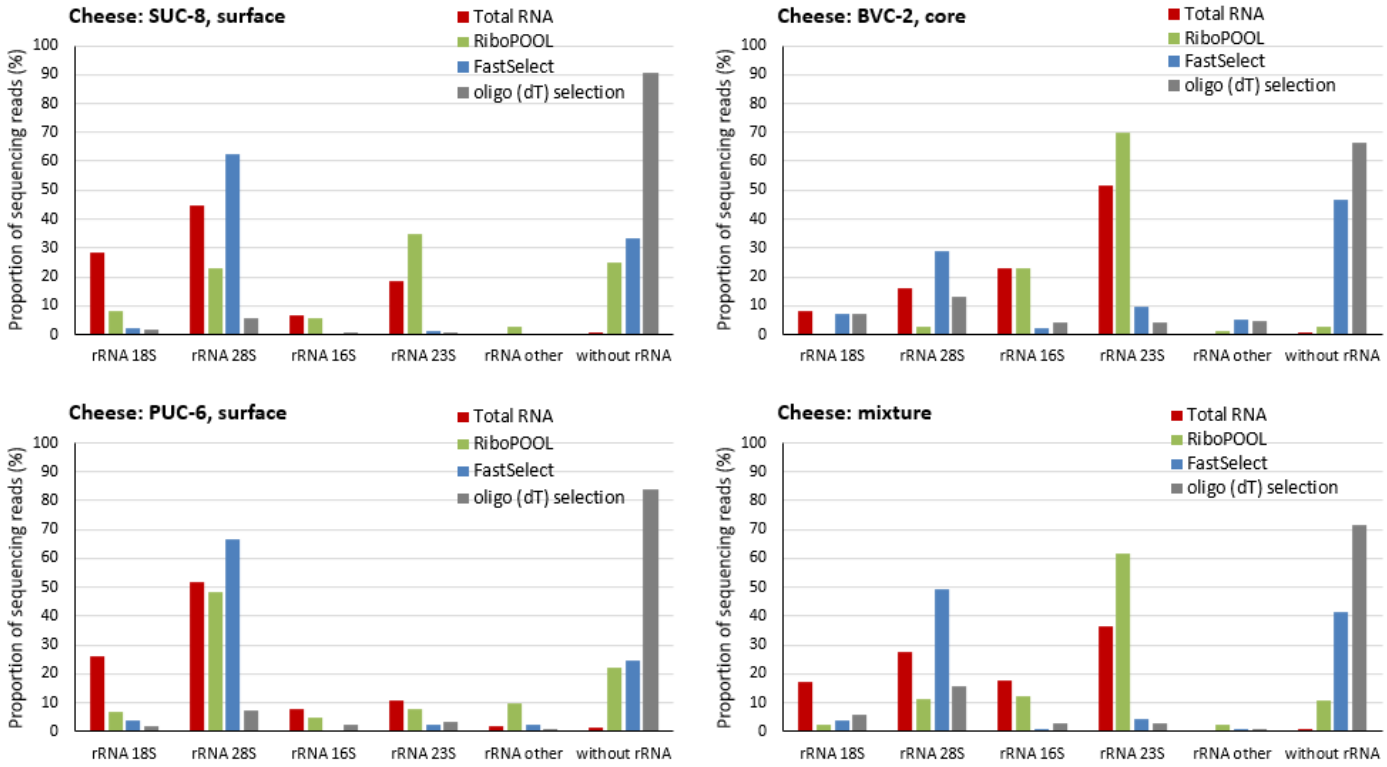


Fig. 5. Presence of ribosomal RNA sequences in the cheese metatranscriptomes before (total RNA) and after ribosomal mRNA enrichment (RiboPOOL and FastSelect rRNA depletion or oligo (dT) selection of polyA RNA). The "Mixture" sample corresponded to a mixture of RNA extracts produced from ten different cheeses. The rRNA sequences were detected using SortMeRNA software.

Table 1

Strains and culture conditions.

Species	Stain[a]	Growth medium[b]	Growth conditions[c]
Lactococcus lactis	S3+	M17-lactose	30°C, partial anaerobiosis
Streptococcus thermophilus	LMD-9	M17-lactose	37°C, partial anaerobiosis
Lactobacillus delbrueckii ssp. *bulgaricus*	ATCC 11842	MRS	37°C, partial anaerobiosis
Lactobacillus helveticus	ATCC 15009	MRS	37°C, partial anaerobiosis
Leuconostoc mesenteroides	DSM 20346	MRS	30°C, partial anaerobiosis
Lacticaseibacillus casei	X4	MRS	30°C, partial anaerobiosis
Brevibacterium aurantiacum	ATCC 9175	Brain Heart Infusion broth	25°C, aerobic conditions
Glutamicibacter arilaitensis	CIP 108037	Brain Heart Infusion broth	25°C, aerobic conditions
Corybacterium casei	DSM 44701	Brain Heart Infusion broth	25°C, aerobic conditions
Hafnia alvei	GB001	Brain Heart Infusion broth	25°C, aerobic conditions
Staphylococcus equorum	Mu2	Brain Heart Infusion broth	25°C, aerobic conditions
Microbacterium gubbeenense	DSM 15944	Brain Heart Infusion broth	25°C, aerobic conditions
Propionibacterium freundenreichii	IC 951	Yeast Extract Lactate broth	30°C, anaerobic conditions
Debaryomyces hansenii	304	Potato Dextrose broth	25°C, aerobic conditions
Geotrichum candidum	ATCC 204307	Potato Dextrose broth	25°C, aerobic conditions
Kluyveromyces lactis	CBS 683	Potato Dextrose broth	25°C, aerobic conditions
Yarrowia lipolytica	CBS 6124	Potato Dextrose broth	25°C, aerobic conditions
Penicillium camemberti	NEIGE LYO 2 D	Potato Dextrose broth	25°C, aerobic conditions

[a]Strains are from the following collections: American Type Culture Collection (ATCC), Rockville, MD, USA; Collection de l'Institut Pasteur (CIP), Paris, France; Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Braunschweig, Germany; Centraalbureau voor Schimmelcultures (CBS), Utrecht, The Netherlands. Strain NEIGE LYO 2 D is from Danisco France SAS (Dangé-Saint-Romain, France), strain IC 951 from CIGC (Comité Interprofessionnel de Gestion du Comté; Poligny, France), and the other strains are from the SayFood culture collection (Thiverval-Grignon, France).

[b]Except for Yeast Extract Lactate broth (Hettinga et al., 1968), growth media are from Biokar Diagnostics (Beauvais, France)

[c]Partial anaerobiosis: static cultures in tubes with headspace volume <10% of the growth medium volume; anaerobic conditions: static cultures with Merck Millipore Anaerocult A (Fischer Scientific, Illkirch, France); aerobic conditions: agitated cultures (150 rpm) in 100-ml conical flasks containing 20 ml of growth medium.

Table 2

Mapping of sequencing reads from cheese RNA samples on CDSs of selected reference genomes.

		Sequencing reads that mapped to CDSs of selected reference genomes (as percentage of total sequencing reads)					
Cheese RNA sample	mRNA enrichment procedure	*Geotrichum candidum* CLIB918	*Lactococcus lactis* IL1403	*Streptococcus thermophilus* LMD9	*Debaryomyces hansenii* CBS767	*Penicillium roqueforti* F164	Sum for the five genomes
SUC-8, surface	no enrichment	0.38	0.02	<0.01	<0.01	<0.01	0.41
	RiboPOOL rRNA depletion	4.32	1.20	0.07	0.06	0.03	5.68
	FastSelect rRNA depletion	17.76	0.77	0.05	0.12	0.16	18.85
	oligo (dT) selection	53.28	0.01	<0.01	0.35	0.45	54.08
PUC-6, surface	no enrichment	<0.01	<0.01	0.04	0.01	0.02	0.08
	RiboPOOL rRNA depletion	0.02	0.21	1.38	0.28	0.17	2.06
	FastSelect rRNA depletion	0.09	0.14	0.95	0.35	0.78	2.31
	oligo (dT) selection	0.32	0.02	0.15	3.70	3.61	7.79
BVC-2, core	no enrichment	<0.01	0.03	0.02	<0.01	0.42	0.47
	RiboPOOL rRNA depletion	0.01	0.16	0.17	0.01	0.91	1.26
	FastSelect rRNA depletion	0.13	2.23	3.73	0.23	18.15	24.47
	oligo (dT) selection	0.30	0.02	0.07	0.69	33.63	34.70
Mixture	no enrichment	0.10	0.02	0.01	0.01	0.13	0.27
	RiboPOOL rRNA depletion	0.48	0.25	0.17	0.11	0.46	1.47
	FastSelect rRNA depletion	8.74	0.84	0.56	0.69	3.95	14.79
	oligo (dT) selection	24.52	0.01	0.02	3.11	5.45	33.11

ND: Not Determined

Table 3

Impact of the mRNA enrichment procedure on the mapping of sequencing reads to the CDSs of selected reference genomes. Pearson's correlation coefficients were measured between the data representing the numbers of hits for each CDS.

Cheese RNA sample	Reference genome	Pearson's correlation coefficient		
		RiboPOOL rRNA depletion vs. total RNA	FastSelect rRNA depletion vs. total RNA	oligo (dT) selection vs. total RNA
SUC-8, surface	*Geotrichum candidum* CLIB918	0.43	0.94	0.86
	Lactococcus lactis IL1403	0.88	0.90	ND
	Streptococcus thermophilus LMD9	ND	ND	ND
	Debaryomyces hansenii CBS767	ND	ND	ND
	Penicillium roqueforti F164	ND	ND	ND
PUC-6, surface	*Geotrichum candidum* CLIB918	ND	ND	ND
	Lactococcus lactis IL1403	ND	ND	ND
	Streptococcus thermophilus LMD9	0.84	0.84	ND
	Debaryomyces hansenii CBS767	0.80	0.88	0.81
	Penicillium roqueforti F164	0.46	0.82	0.77
BVC-2, core	*Geotrichum candidum* CLIB918	ND	ND	ND
	Lactococcus lactis IL1403	0.93	0.93	ND
	Streptococcus thermophilus LMD9	0.89	0.88	ND
	Debaryomyces hansenii CBS767	ND	ND	ND
	Penicillium roqueforti F164	0.87	0.96	0.71
Mixture	*Geotrichum candidum* CLIB918	0.71	0.94	0.89
	Lactococcus lactis IL1403	0.92	0.95	ND
	Streptococcus thermophilus LMD9	0.89	0.84	ND
	Debaryomyces hansenii CBS767	0.75	0.84	0.84
	Penicillium roqueforti F164	0.80	0.95	0.77
	Mean :	*0.78*	*0.90*	*0.81*

ND: Not Determined (correlations were not considered for oligo (dT) selection with prokaryotic reference genomes and when sequencing reads mapping to the reference genome represented less than 0.01% of the reads of the sample with no mRNA enrichment).

Additional File 1

Table S1

Proportion of the sum of the areas of intact 16S, 18S, 23S and 26S rRNA in the electrophoregrams (Agilent Bioanalyzer).

	Proportion of (16S+18S+23S+26S) peak area (%)		Proportion of (16S+18S+23S+26S) peak area (%): Means per cheese type	
Cheese	**Surface**	**Core**	**Surface**	**Core**
Fresh cheese				
FC-1	NA	23.35		
FC-2	NA	30.8		
FC-3	NA	3	NA	14.9
FC-4	NA	17.3		
FC-5	NA	0.1		
Soft unpressed cheese				
SUC-1	19.7	25.1		
SUC-2	14.3	13.4		
SUC-3	16.6	38.3		
SUC-4	27.2	19.2		
SUC-5	30	10.2	29.5	15.2
SUC-6	11.35	14.75		
SUC-7	32.5	16.1		
SUC-8	41.9	2.4		
SUC-9	52.4	32.6		
SUC-10	48.5	29.9		
Pressed uncooked cheese				
PUC-1	43	7.4		
PUC-2	8.45	0.2		
PUC-3	2	8.7		
PUC-4	18.7	46.8	14.3	15.2
PUC-5	3.6	33.5		
PUC-6	10.25	24.7		
PUC-7	NA	0.2		
PUC-8	NA	0.1		
Pressed cooked cheese				
PCC-1	14.1	0.4		
PCC-2	6.4	0.1		
PCC-3	2.95	7.3	6.0	1.6
PCC-4	0.7	0.1		
PCC-5	NA	0.1		
Blue-veined cheese				
BVC-1	NA	10.2		
BVC-2	NA	4.75	NA	15.0
BVC-3	NA	26.1		
BVC-4	NA	18.9		

NA: Not Analyzed

Table S2

Number of sequencing reads generated from cheese RNA samples.

	Sequencing reads from cheese RNA samples			
mRNA enrichment procedure	SUC-8, surface	PUC-6, surface	BVC-2, core	Mixture
no enrichment	342 122 711	333 253 580	262 107 142	307 732 435
RiboPOOL rRNA depletion	101 577 768	101 020 111	86 344 428	96 036 772
FastSelect rRNA depletion	68 312 640	91 788 587	91 377 513	82 987 913
oligo (dT) selection of polyA RNA	79 690 016	70 294 764	69 376 527	81 105 447

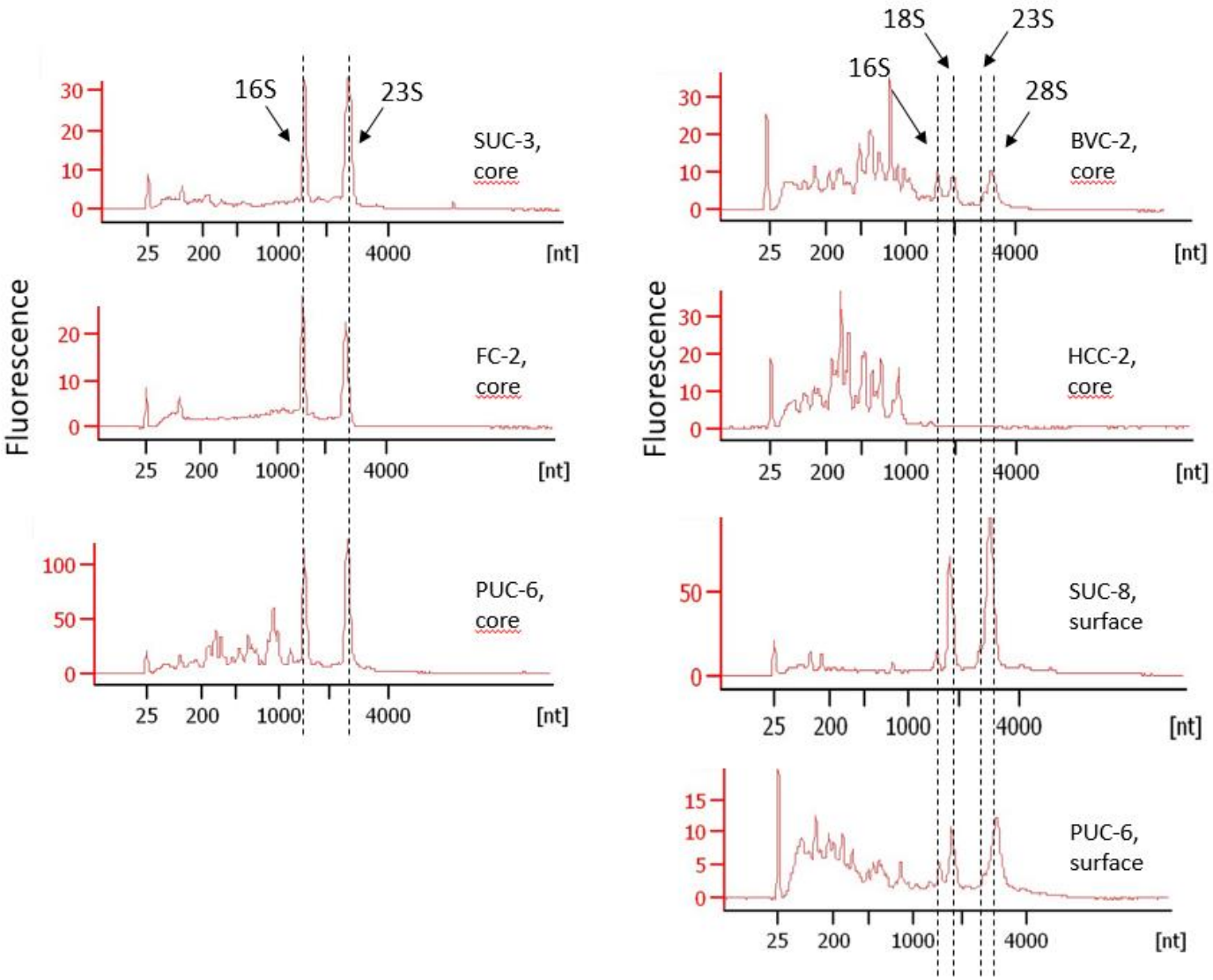


Fig. S1. Examples of electrophoregrams (Agilent Bioanalyzer) of RNA preparations from cheeses.

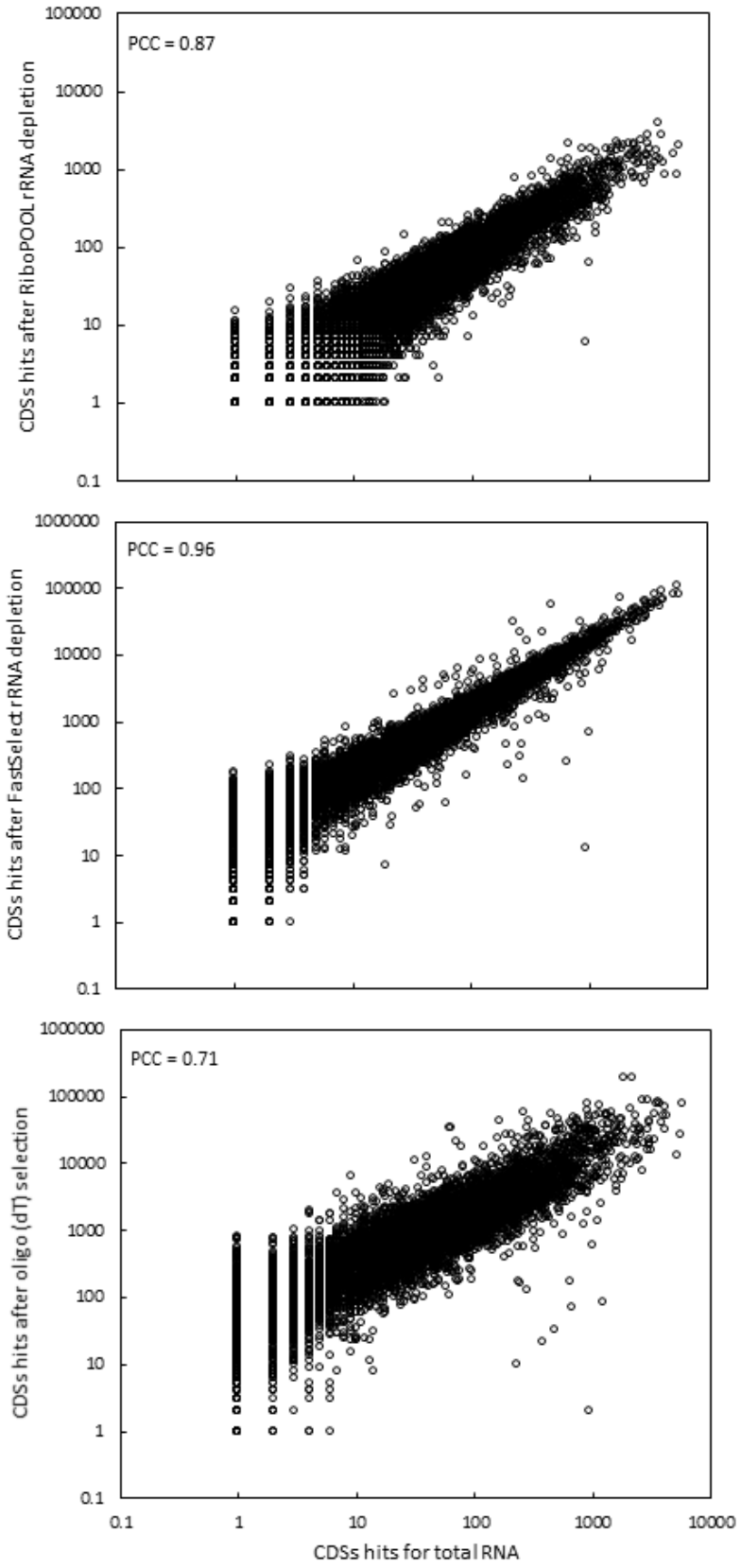


Fig. S2. Examples of correlations between mapping of sequencing reads of total RNA and mapping after mRNA enrichment. The figure presents double-log scatter plots linking the number of sequencing reads mapping to CDSs of the *Penicillium roqueforti* F164 reference genome for RNA samples from cheese BVC-2 (PPC: Pearson correlation coefficient).