PROF. DIEGO MORA (Orcid ID : 0000-0002-6692-9301)

Characterization of airborne viromes in cheese production plants

Stefano Colombo, Stefania Arioli, Giorgio Gargari, Eros Neri, Giulia Della Scala, Diego Mora

Department of Food, Environmental and Nutritional Sciences, University of Milan, Italy

Running title: *Airborne viromes in cheese production plants*

Corresponding author:

Prof. Diego Mora

Department of Food, Environmental and Nutritional Sciences, University of Milan

Via Celoria 2, 20133 Milan, Italy

phone   +39 0250319133

fax   +39 0250319238

e-mail   diego.mora@unimi.it

Abstract

Aims: To characterize airborne virus like particles isolated from two cheese production plants in order to reveal their complexity in terms of viral communities and microbial genes potentially mobilized by viruses.

Methods and results: Airborne virus like particles have been isolated from Grana Padano and Gorgonzola PDO cheese production plants and ripening cellars. A shotgun metagenomics analysis of the isolated viromes highlighted a high complexity of the viral communities both in terms of viral taxonomy and phage-host associations. Bacterial reads in each of the viromes were confirmed to be abundant and their taxonomy appeared to be associated to the environmental parameters and the technological processes that characterize the sampling area. Antibiotic resistance genes have been identified in each virome thus confirming that phages could be involved in the mobilization of antimicrobial resistances among bacterial populations. Interestingly human viruses were also identified even if the contamination source was not revealed.

Conclusions: The environmental conditions, which are imposed by the technology of the dairy process, seams to shape the viral populations as a consequence of the adaptation of microbial taxa to those environments. The identification of sequences belonging to *Legionella pneumophila* and to the human papillomavirus, raised some considerations about the safety of cheese ripening cellars.

Significance and impact of the study: In conclusion, the analysis of the dairy airborne viromes, has revealed a high complexity of the viral communities even if the environments where the samples were collected were confined environments. Metagenomics of airborne

viral population could be a promising monitoring tool for the biological characterization of dairy environments.

Introduction

In many cheese-making activities microorganisms in the air could represent a hidden and dangerous risk factor. The air present in the processing areas can be a source of pathogenic or spoilage microorganisms that could contaminate foods causing problems both on the safety and the quality of the foods themselves. The greatest aerosol sources in dairy plants are personnel, floor drains, ventilation system and water when applied under pressure in the cleaning and sanitizing procedures (Andersen 1958; Vickers *et al*. 1986; Kang *et al*. 1990; Salustiano *et al*. 2003). Air can easily move microorganisms, spores and dust to reach food in preparation. Identification of microorganisms present in the air cannot be limited to bacteria and fungi but should consider also viral populations. Viruses, and specifically phages, have a meaningful impact due to their role in modulating the abundance of microbial populations (Heldal and Bratbak, 1991; Breitbart and Rohwer 2005; Danovaro *et al*. 2017;), and because they represent a tool for horizontal gene transfer mechanisms (Mann and Slauch 1997; Schmieger and Schicklmaier 1999; Muniesa *et al*. 2013).

Beyond the problems that could affect the final products, production steps could be highly damaged by contamination of microorganisms and in particular phages. Despite huge scientific and technological advances over the last 70 years, phages remain the largest single cause of fermentation failure in the dairy industry. In the most severe cases phage infections

can cause a complete loss of starter activity resulting in dead vats, and the associated problems of disposing of large quantities of partly acidified milk (Marcò *et al*. 2012).

In this study two cheese-making factories specialized in the production of two very different cheeses, Gorgonzola and Grana Padano, have been considered for airborne sampling. Grana Padano (GP) cheese is an Italian hard cheese granted a Protected Designation of Origin (PDO) seal. It is made from partially skimmed raw cow's milk and natural whey cultures (NWC) of lactic acid bacteria used as a starter (SLAB), cooked at 53 °C and matured for 9 to 20 months and above. In the end product GP microbial community composition comprehend SLAB and secondary microbiota that develops during a long ripening period, which also includes non-starter lactic acid bacteria (NSLAB) that usually come from milk and dairy environment (Santarelli *et al*. 2013). Gorgonzola is an Italian blue-veined cheese. It is one of the most popular Italian cheeses and is produced under controlled denomination of origin (DOC). It is made of pasteurized cow's milk inoculated with thermophlic Lactic Acid Bacteria (LAB) *Streptococcus thermophilus* and *Lactobacillus delbrueckii* subsp *bulgaricus*. It has been shown that the initial growth and production of lactic acid by thermophilic bacteria is indispensable for producing environmental conditions (low pH and lactic acid as a carbon source) suitable for the rapid growth of yeast (commonly *Saccharomyces cerevisiae*) and *Penicillium roqueforti* (Florèz and Mayo 2006; Cheeseman *et al*. 2014). Aging lasts usually from 60 to 90 days depending on the type of Gorgonzola.

Air was sampled in different areas of a Grana Padano and a Gorgonzola production plants in order to study the viruses that are present during different cheese production steps. Virus like particles have been collected, purified and identified through a metagenomic shotgun sequencing. A flow chart of the methodological approach applied is represented in Fig. 1. Due to the high amounts of bacterial reads associated to viromes as reported in several studies (Fancello *et al*. 2013; Modi *et al*. 2013; Colombo *et al*. 2016; Colombo *et al*. 2017),

we put a special attention to the taxonomy and functional characterization of bacterial genes associated with the air-borne viromes collected in each dairy environment. The study of food contamination used to focus the attention on the possible risks coming from raw materials, food handling and contaminate tools. On the other hand, the potential of the air and aerosol as a risk of microbial contamination is often underestimated (Verreault *et al*. 2011a, 2011b; Joung *et al*. 2017).

Materials and methods

Sampling

Samples were collected in Grana Padano and Gorgonzola production plants. Thirty cubic meters of air surrounding the food samples were collected using Coriolis system (Bertin technologies, Montigny-le-Bretonneux, France). Air was sucked by the system and directed in ten milliliters of sterile water supplemented with 0.01% Tween 20: the cyclonic effect developed by the Coriolis system trapped microorganisms and particles in the liquid.

Virus-like particle isolation

Samples collected with Coriolis system were filtered through 0.2 µm pore size filter with a Sartorious filtering system (Sartorius AG, Goettingen, Germany). Briefly, the permeate containing all the virus-like particles (VLPs) was then precipitated overnight at 4 °C using PEG 8000 at a final concentration of 10% and then centrifuged at 13,000 x g for 30 min at 4 °C. The pellet was suspended in TE buffer (pH 8.0) and then prepared for CsCl gradient centrifugation. The solution was subsequently deposited on top of a 2.5 mL step gradient composed of multiple 0.8 mL CsCl solutions with corresponding densities of 1.7 g ml^{-1}, 1.5 g ml^{-1} and 1.35 g ml^{-1}. Samples were centrifuged for 2 h at 60,000 x g (4 °C) in a SW41

swinging bucket rotor (Beckman Instruments Inc.; Fullerton, CA) (Thurber *et al*. 2009). According to the protocol of Thurber *et al*. (Thurber *et al*. 2009), we recovered approximately 2 ml of the 1.5 g ml^{-1} layer, because material in this density range should be enriched for VLPs.

Extraction and amplification of VLP DNA for metagenomic analysis

Viral DNA was extracted according to a protocol previously described by Thurber *et al*. (2009). The 1.5 g ml^{-1} layer collected from the step gradient was treated with DNase, following the manufacturer's instructions (Sigma Aldrich; final concentration, 2.5 U ml^{-1}), to remove residual host and bacterial DNA. To extract the virions, 0.1 volumes of 2 mole l^{-1} Tris-HCl/0.2 mole l^{-1} EDTA, 1 volume of formamide and 100 µl of a 0.5 M EDTA solution was added to the 10 mL sample, and the resulting mixture was incubated at room temperature for 30 min. The sample was subsequently washed with 2 volumes of 70% ethanol and pelleted by centrifugation at 8,000 x g for 20 min at 4 °C. The pellet was washed twice with 70% ethanol and resuspended in 567 µl of TE buffer, followed by 30 µl of 10% SDS and 3 µl of a 20 mg ml^{-1} solution of proteinase K (Fisher Scientific, Waltham, MA, USA). The mixture was incubated for 1 h at 55 °C and supplemented with 100 µL of 5 mole l^{-1} NaCl and 80 µl of a solution of 10% cetyltrimethylammonium bromide/0.7 mole l^{-1} NaCl. After a 10 min incubation at 65 °C, an equal volume of chloroform was added and the mixture was centrifuged (5 min at 8,000 x g at room temperature). The resulting supernatant was transferred to a new tube and an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added, followed by centrifugation (5 min at 8,000 x g at room temperature). The supernatant was recovered and an equal volume of chloroform was introduced. Following centrifugation, the supernatant was collected and 0.7 volumes of isopropanol were used to precipitate the DNA. After centrifugation (15 min at 13,000 X g and 4 °C), the DNA

pellet was washed with 500 µl of cold 70% ethanol, air-dried and resuspended in 50 µl TE (Thurber *et al*. 2009).

Finally, 2 µl of each DNA solution was quantified using the PowerWave XS Microplate Spectrophotometer at 260 nm (BioTek, Instruments, Inc., CA, USA) and the Take3 Multi-Volume Plate (BioTek Instruments, Inc., CA, USA).

Calculation of bacterial contamination using qPCR

An aliquot of the purified DNA was used as template in qPCR assays to quantify the amount of contaminating, non-viral DNA. A CFX96 thermocycler (BioRad Laboratories S.r.l., Milan, Italy) was used to quantify the bacterial DNA. The analysis was performed using a Fast Eva Green Supermix SYBR Green PCR Master Mix (Bio-Rad Laboratories) in a reaction volume of 15 µl per well. qPCR amplification was carried out with an initial denaturation at 95 °C for 5 min, followed by 40 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s and extension at 72 °C for 30 s. The primers EubF1 (GTGSTGCAYGGYTGTCGTCA) and EubR1 (GAGGAAGGTGKGGAYGACGT) were used to target *16S rRNA* to detect any eubacteria. Data were recorded as threshold cycles (Ct) and expressed as mean values; standard deviations were computed using the BioRad CFX Manager. DNA from pure cultures of *Lactococcus garviae* strain TB25 was subjected to six 10-fold serial dilutions and used as a templates for qPCR. Standard curves were generated by plotting the $\log_{10}$ of the bacterial cell numbers against the corresponding Ct values obtained from the amplification of diluted DNA. Calibration curves showed good correlation between Ct values and the number of cells over the considered range (r^2 regression coefficients were between 0.964 and 0.999). The resulting detection limits were between 1.1 and 1.9 $\log_{10}$ bacterial cells per reaction mix (Guglielmetti *et al*. 2013). On the basis of the Ct values obtained using viral DNA, the actual concentration of bacterial DNA was determined and expressed as a percentage of the total

DNA previously extracted. Bacterial contaminating DNA accounted in less than 0.2% of the total viral DNA extracted.

Metagenomic shotgun sequencing of viral DNA

A genomic library was generated from 3 μg of genomic DNA using the TruSeq DNA PCR-Free Sample Preparation Kit (Illumina) and the MiSeq Reagent, according to the user's guide (Illumina). The quality of the library was estimated prior to loading onto a flow cell, and the library sample was subsequently sequenced using 500 sequencing cycles, according to the MiSeq (Illumina) instructions. The 500 sequencing cycles resulted in an average read length of approximately 250 nucleotides for both paired-end sequences. The MIRA program version 4.0.2 (Chevreux *et al*. 1999) was used for *de novo* assembly of contigs. Mira options were set as follow: "job = genome, denovo, accurate parameters = -nw:cac=no -ge:not=1 illumina_settings -as:mrpc=100" ([minimum_reads_per_contig(mrpc)= 100]). Assembled reads were searched for ORFs predicted with PRODIGAL v2_60 linux(http://prodigal.ornl.gov/) (Hyatt *et al*. 2010)

Bioinformatic analysis

A BLAST analysis was performed against non-redundant viral database using BLAST+ 2.2.29 ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/. A database was created using all entries for the taxonomy ID 10239 (Viruses) on NCBI (Colombo *et al*. 2016).

Both viral and microbial data underwent a BLAST analysis against the both the non-redundant nucleotide and the non-redundant protein databases of NCBI; BLASTn, BLASTp and BLASTx were performed with a threshold of 10^{-5} on the E value. In order to identify reads associated to antibiotic-resistance genes (ARGs), the data were compared with the CARD (Comprehensive Antibiotic Resistance Database); the database is composed of 3228 genes specifically tagged for antibiotic resistance (McArthur *et al*. 2013). BLASTp and BLASTn were performed in order to identify the antibiotic resistance-encoding reads with a

threshold of 10^{-5} on the E value. Only sequences with an aminoacidic identity > 30% and a nucleotide identity > 70% were considered.

Accession numbers. The sequence data have been submitted to the EMBL-EBI database under accession number PRJEB22115

Results

Viromes metagenomics

In this study, we characterized the airborne viromes collected from three areas of two different cheeses production plants. Metagenomic shotgun sequencing highlighted the presence of the same families of virus in the cheese making plants, and in different areas of the same plant with few exceptions; only moderate variations in the relative abundance of each viral family were observed. BLAST analysis resulted in a larger presence of virus related ORFs in the air surrounding the Gorgonzola curd (13%) in comparison to the other two (4% and 5%). As expected, most of ORFs were assigned to bacteria kingdom, from 62% to 91% (Fig. 2). In order to check the quality of the extracted viral DNA, we quantified the relative amount of ribosomal sequences in all three samples by qPCR. The quality control assays revealed that the levels of bacterial contaminating DNA in virome samples accounted for less than 0.2% of the total extracted viral DNA extracted, a threshold considered acceptable as suggested by Mody *et al*. (2013). We therefore exclude that this low level of bacterial DNA contamination in viromes could have affected the subsequent bioinformatic analysis of viromes. The huge amounts of bacterial reads in each virome were after all comparable to those previously described by other authors for viromes isolated from freshwater samples (López-Bueno *et al*. 2009; Rosario *et al*. 2009; Roux *et al*. 2012; Fancello *et al*. 2013; Modi *et al*. 2013), as well as in a previous study carried out in experimental aquaculture plant, and in water river samples (Colombo *et al*. 2016; Colombo *et al*. 2017).

Taxonomy of viral communities

The three samples showed a similar taxonomic distribution of viral families with some exceptions (Table 1). In the three viromes the sequences identified as viruses infecting bacteria and archea ranged between 31% to 46% of the total viral sequences, whereas viruses infecting eukaryotes ranged from 28% to 42% thus indicating a discrete level of non-bacteria contamination in the three environments. Virome collected from air sample in Grana Padano production plant showed the highest amount of ORFs unclassifiable at a family level and generically identified as double stranded DNA viruses (24%). *Myoviridae*, *Mimiviridae*, *Phycodnaviridae* and *Siphoviridae* related ORFs where in a range of 12%-17% of the total. Grana Padano related virome was the only sample with an appreciable abundance of *Podoviridae* related ORFs. The virome associated with Gorgonzola production area showed the highest abundance, with respect to the other two samples, both for *Myoviridae* and *Mimiviridae* as well as *Phycodnaviridae* (19% to 25%); on the other hand, *Siphoviridae* related ORFs were the lowest in abundances in the three samples.

Taxonomy of the bacterial reads associated to viromes

While the taxonomy of viral families showed few differences in the three airborne viromes, a huge diversity was disclosed analyzing the taxonomy of the bacterial reads associated to each virome (Table 2). As observed by several authors, phage metagenomes are enriched with microbial functions (Fancello *et al*. 2013; Modi *et al*. 2013; Colombo *et al*. 2016, 2017) whose ecological role is still unknown. The airborne viromes collected in this study were therefore subjected to the characterization of their bacterial sequences. At a phylum level the main differences between the Grana Padano and the Gorgonzola air samples were related to the relative abundances of *Bacteroidetes* and *Firmicutes*. The first was more abundant in Gorgonzola air samples, whereas the second was the dominating phylum in the Grana Padano

sample (Table 2). At a class level (Table 3), a further differentiation was possible between the two Gorgonzola air samples. In fact, the airborne viromes collected in Gorgonzola production plant (day 1 air sample) showed higher presence of bacterial ORFs from *Sphingobacteria*, *Flavobacteriia* and *Bacilli*, whereas virome from the ripening cellar (day 50 air sample) was enriched in ORFs from *Cytophagia*, *Chlamydiia* and unclassified *Alphaproteobacteria* classes. *Gammaproteobacteria*, equally abundant in both the Gorgonzola samples, were mainly represented by *Moraxellaceae* (37 %), *Altermonadaceae* (12%), and *Legionellaceae* (10%) in day 1 air sample, and by *Legionellaceae* (30%), *Moraxellaceae* (12%) and *Coxiellaceae* (7%) in day 50 air sample. Interestingly, among the 19 species of *Legionella* species identified in day 50 air samples, the most abundant species was the opportunistic pathogen *L. pneumophila* (22% of relative abundance) (Qiu and Luo, 2017), which was present also in day 1 air sample with a relative abundance of 24% being *L. fallonii* the most represented species in this sample (42% of relative abundance).

A completely different taxonomy has been observed in air sample collected from Grana Padano production plant: ORFs from families highly abundant in the two Gorgonzola related samples, were barely present while the other never identified over 3% threshold. However, ORFs related to *Bacilli* (38%), *Clostridia* (19%) and *Bacteroidia* (19%) class were extremely abundant in Grana Padano airborne virome (Table 3). Within *Bacilli* the *Paenibacillaceae*, *Lactobacillaceae* and *Bacillaceae* families showed respectively a relative abundance of 44%, 35% and 15% respectively. *Clostridia* were mainly represented by the family *Heliobacteriaceae* (54%) and by the less abundant *Syntrophomonadaceae* (11%).

Antibiotic resistance genes associated to viromes

Antibiotic resistance genes (ARGs) were identified through comparison with CARD database. Airborne virome collected from Gorgonzola ripening room showed the highest

frequency (2.97%, day 50 air sample) of ARGs calculated as number of ARGs related ORFs over the number of ORFs identified in NCBI. Lower ARGs frequency was observed in the airborne virome associated to the Gorgonzola production area (1.87%, day 1 air sample) and in airborne virome associated to the Grana Padano plant (0.42%). In addition to the differences observed in ARGs frequency, the three airborne viromes also differed in relation to the number of antibiotic molecules identified as targets of the ARGs. Airborne virome associated to the Gorgonzola production plant showed the highest diversity of ARGs.

We also performed ARGs classification based on the antibiotic drug target (Fig. 3). Virome associated to Gorgonzola production plant showed several classes of ARGs being the antibiotic efflux related genes (34 %) and the DNA synthesis inhibitor related genes (17 %) the most and the least represented, respectively. The other two airborne viromes showed a less diverse distribution of ARGs classes, being the antibiotic efflux class the most abundant. Noteworthy, both DNA and protein synthesis inhibitor ARGs classes were represented with values ranging from 13% to 26% in both the viromes. Looking even more in detail (Table 4) we measured a relative abundance of genes associated to lipopetide, aminocoumarin and chlorampenicol resistance (all at 14%) in airborne virome associated to Gorgonzola production area. Whereas the relative abundance of genes associated to lipopetide, aminocoumarin glycopeptide and beta-lactam were 7% and 9% respectively in airborne viromes collected at Gorgonzola ripening room. ARGs associated to elfamycin and tetracycline resistance were identified in airborne virome collected in Grana Padano production plant over the 5% threshold (Table 4).

Phage-host interactions

A prediction of the phage-host interaction has been performed on the basis of the taxonomic information of phages identified in the three airborne viromes. The prediction shows that the

Proteobacteria is the most stricken phylum in all the samples with 12 different genera identified as target of phages. However, other phyla were also represented as phage-targets, among them *Firmicutes* and genera commonly involved in dairy fermentation process, such as *Lactococcus* and *Streptococcus*. In Grana Padano airborne virome the highest abundance of phage-host interaction were targeted against genera *Vibrio, Acinetobacter* and *Pseudomonas* (Fig. 4), whereas in the two airborne viromes collected in Gorgonzola production and ripening plants the highest abundance was associated to the *Firmicutes* phylum, and in particular to *Bacillus* and *Streptococcus* genera. The highest abundance of phages in the airborne-virome collected from Gorgonzola production area were associated to the genus *Streptococcus*, *Cronobacter* and *Cellulophaga*, whereas the genus *Bacillus* and *Leptospira* were those mostly represented for the airborne virome collected from the Gorgonzola ripening room (Fig. 4).

Discussion

The analysis of the dairy airborne viromes, has revealed a high complexity of the viral communities even if the environments where the samples were collected are the confined ones. The complexity of the viral populations is certainly related to the diversity of the microbial communities, which come mainly from milk but also from air, personnel, equipment, working surfaces, tap water. The environmental conditions (temperature, air humidity), which are imposed by the technology of the dairy process, select some microbial taxa and the related viral populations. Beside the information related to the ecology of viral populations, the metagenomics analysis of airborne viromes, the identification of sequences belonging to *Legionella pneumophila* and to the human papillomavirus, have raised some considerations about the safety of the sampled dairy environments. Nevertheless, a final statement on the safety of the environments sampled should be corroborated by the isolation

of pathogens by standard microbiological procedures. Whereas airborne associated virome collected from Gorgonzola ripening cellar showed a similar profile in comparison to the other samples, with the only exception of the human viruses belonging to the *Papillomaviridae* family that showed a high abundance (17%) (Table 1). *Papillomaviridae* reads were organized in 17 contigs and all ORFs showed an amino acidic identity higher than 90% with human papillomavirus type 24. Interestingly, beside the prevalence of viruses belonging to families known to infect bacteria and archea, viruses infecting eukaryotes included human have been identified. Human viruses belonging to the *Papillomaviridae* family were identified in the air sample collected form the ripening cellar of Gorgonzola cheese where the human presence is only limited to workers that move cheese wheels and check their ripening status. Although papillomavirus diseases are transmitted through direct skin-to-skin contact and are considered as sexually transmitted diseases, other type of transmission of the virus are still under debate also in consideration of their high stability outside the host (Ryndock and Meyers, 2014). The presence of papillomavirus in fresh lettuce has been reported by Aw et al. (2016), and in that case the irrigation water was identified as the possible source of contamination. Moreover, members of the family *Papillomaviridae* have been identified in water samples collected from human-managed aquatic systems (Kim et al. 2017). In our sample, the possible source of contamination could be restricted to the workers operating in the ripening cellar handling daily the cheese wheels. In this context, it is worth of mentioning that members of the family *Papillomaviridae* have been identified on the palm of healthy individuals (Hannigan et al. 2015). The stability of papilloma viral particles in the air could be ascribed to the low temperature (4-5 °C) and the high percentage of humidity of that area (99%).

Upon analysis of the taxonomy of bacterial reads associated to viromes was raised the following considerations raised. Concerning the *Coxiellaceae* family, all the reads were

identified as *Coxiella burnetii* another opportunistic pathogen and a causative agent of the zoonosis Q fever (Qiu and Luo, 2017). *Legionella* and *Coxiella* are Gram-negative bacteria genera which comprise species described as opportunistic pathogens, such as *L. pneumophila* and *Coxiella brunetii*. The latest replicates in free-living amoebae, and inhalation of contaminated aerosols by immunocompromised individuals can cause a severe and potentially fatal form of bacterial pneumonia or the mild flu-like Pontiac fever (Newton *et al*. 2010). *Legionella* species is frequently associated to water environments or environments characterized by high air humidity and low temperature such as room humidifiers, cooling towers and air conditioning systems (Brenner *et al*. 1985). These systems often exhibit favorable growth conditions for *Legionella* and transmission to humans occurs via contaminated water aerosolization (Van Heijnsbergen *et al*. 2015). Presence of *Legionella* related ORFs in the air virome was therefore in accordance with the particular environment, which characterize the ripening rooms of Gorgonzola cheese. *Coxiella brunetii* is the causative agent of the zoonosis Q fever and humans get infected mainly through inhalation of contaminated aerosols (Filippitzi *et al*. 2017).

When the taxonomy of bacterial genes associated to the airborne viromes was examined, we were able to clearly differentiate samples based on the type of cheese produced. Interestingly, the two airborne viromes collected from Gorgonzola production plant resulted highly similar even if they belonged to production areas that totally differ in terms of temperature, humidity, light, and presence of workmen. The production area, where milk was inoculated with rennet and the starter culture (*Streptococcus thermophilus*, *Lactobacillus delbrueckii* subsp. *bulgaricus*, *Penicillium roqueforti* and *Saccharomyces cerevisiae*) was an area with filtered air and forced ventilation, characterized by 82% humidity, wet floor, a temperature ranging from 20-22 °C, and the simultaneously presence of 6-10 operators, as well as high frequency of aerosol due to the processes related to curd production, whey loss and cheese wheel

preparation. The ripening area (where cheese wheels were stored to allow the growth of *P. roquefortii* and the development of the related proteolytic and lipolytic activity) was an area with filtered air and forced ventilation, characterized by a high level of humidity (99%), a temperature of 4-5 °C and the absence of workmen which were present only few times a week to move cheese wheels and verify cheese ripening.

The excessive use of antibiotics may select for resistant bacterial populations in the environment, other biotic and abiotic factors may have a role in the spreading of ARGs (Berendonk et al. 2015). Antibiotic resistance may be acquired by spontaneous mutations or acquired by the acquisition of ARGs by horizontal gene transfer events. Mobile genetic elements such as plasmids, transposons, integrons and bacteriophages, mediate the spread of ARGs among bacterial cells (Lupo et al. 2012). Among several bacterial gene categories, the role of viruses in the mobilization of antibiotic resistance genes (ARGs) has been recently experimentally documented, showing that the treatment with antibiotics leads to an increase in the ARGs abundance in the virome (Fancello et al. 2013, Modi *et al*. 2013; Abeles et al. 2015;). As observed in previous studies (Modi *et al*. 2013; Colombo *et al*. 2016; Colombo *et al*. 2017;), viromes can represent a genetic reservoir of bacterial genes. However, the extent to which they are accessible to members of the environmental microbiota remains unclear. In our study, we showed that ARGs are present in airborne viromes of dairy plants and that the distribution of drug classes, based on the inhibition target of the antibiotic, was dependent on the sampling site.

Finally, the phage-host association distribution revealed that phages able to infect the bacterial taxa commonly used for the production of Grana Padano and Gorgonzola cheeses were present in the airborne samples, thus confirming that phage monitoring procedures should take into consideration also the bio-aerosol formation in dairy plants (Neve et al. 1995; Verreault et al. 2011a, 2011b). Interestingly, due to the complexity of the microbial

communities in the three environments, the phage-host association distribution revealed also the presence of phages whose targets were bacterial genera not involved in the food fermentation process, but associated to the environments where the air samples were collected.

The overall data collected from the metagenomics of airborne virus like particles disclosed a huge amount of information that could be a promising monitoring tool for the biological characterization of dairy production plants. The potential usefulness of viromes for the monitoring of ARGs spreading, or for detection and control of pathogens is promising. However, further investigations are necessary. Beside the metagenomics of microbial populations, viromes could represent a further source of information useful for the description of environments associated to productive human activities such as those involved in dairy processes.

Acknowledgements

Conflict of interest

The authors have no conflict of interest to declare.

References

Abeles, S.R., Ly, M., Santiago-Rodriguez, T.M., Pride, D.T. (2015) Effects of Long Term Antibiotic Therapy on Human Oral and Fecal Viromes. *PLOS One* **26** 10(8), e0134941.

Andersen, A.A. (1958) New Sampler for the collection, sizing, and enumeration of viable airborne particles. *J Bacteriol* **76**, 471-484.

Aw, T.G., Wengert, S., Rose, J.B. Metagenomic analysis of viruses associated with field-grown and retail lettuce identifies human and animal viruses.(2016) *I J Food Microbiol* **223**, 50-56.

Berendonk, T.U., Manaia, C.M., Merlin, C., Fatta-Kassinos, D., Cytryn, E., Walsh, F., Bürgmann, H., Sørum, H., Norström, M., Pons, M-N., Kreuzinger, N., Huovinen, P., Stefani, S., Schwartz, T., Kisand, V., Baquero, F., Martinez, J.L. (2015). Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol* **13**, 310-317.

Breitbart, M. and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol* **13**, 278-28;

Brenner, D.J., Steigerwalt, A.G., Gorman, G.W., Wilkinson, H.W., Bibb, W.F., Kackel, M., Tyndal, R.L., Campbel, J., Feeley, J.C., Thacker, W.L., Skaliy, P., Martin, W.T., Brake, B.J., Fields, B.S., McEachern, H.V., Corcoran, L.K. (1985) Ten new species of *Legionella*. *I J Syst Bacteriol* **35**, 50-59.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Gonzalez Peña, Goodrich, J.K, Gordon, J., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widman, J., Yatsunenko, T., Zaneveld, J., Knight, R. (2010) QIIME allows analysis of high-

throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat Methods* **7**, 335–336.

Cheeseman, K., Ropars, J., Renault, P., Dupont, J., Gouzy, J., Branca, A., Abraham, A-L., Ceppi, M., Conseiller, E., Debuchy, R., Malagnac, F., Goarin, A., Silar, P., Lacoste, S., Sallet, E., Bensimon, A., Giraud, T., Brygoo, Y. (2014) Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nat Commun* **5**, 2876.

Chevreux, B., Wetter, T., and Suhai, S. (1999) "Genome sequence assembly using trace signals and additional sequence information". German Conference on Bioinformatics.

Colombo, S., Arioli, S., Guglielmetti, S., Lunelli, F., Mora, D. (2016) Virome-associated antibiotic-resistance genes in an experimental aquaculture facility. *FEMS Microbiol Ecol* **92**, fiw003.

Colombo, S., Arioli, S., Neri, E., Della Scala, G., Gargari, G., Mora, D. (2017) Viromes as genetic reservoir for the microbial communities in aquatic environments: a focus on antimicrobial-resistance genes. *Front Microbiol* **8**, 1095.

Danovaro, R., Dell'Anno, A., Corinaldesi, C., Rastelli, E., Cavicchioli, R., Krupovic, M. (2016) Virus-mediated archaeal hecatomb in the deep seafloor. *Sci Adv* **2**(10) e1600492. DOI: 10.1126/sciadv.1600492

DeSantis, T.Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E.L., Keller, K., Huber, T., Dalevi, D., Hu, P., Andersen, G.L. (2006) Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol* **72**, 5069–5072.

Fancello, L., Trape, S., Robert, C., Boyer, M., Popgeorgiev, N., Raoult, D., Desnues, C. (2013). Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J* **7**, 359-369.

Filippitzi, M.E., Goumperis, T., Robinson, T., Saegerman, C. (2017) Microbiological zoonotic emerging risks, transmitted between livestock animals and humans (2007-2015). *Transbound Emerg Dis* **64**, 1059-1070.

Guglielmetti, S., Fracassetti, D., Taverniti, V., Del Bó, C., Vendrame, S., Klimis-Zacas, D., Arioli, S., Riso, P., Porrini, M. (2013) Differential modulation of human intestinal *Bifidobacterium* populations after consumption of a wild blueberry (*Vaccinium angustifolium*) drink. *J Agric Food Chem* **61**, 8134−8140.

Hannigan, G.D., Meisel, J.S., Tyldsley, A.S., Zheng, Q., Hodkinson, B.P., SanMiguel, A.J., Minot, S., Bushman, F.D., Grice, E.A. (2015) The human skin double-stranded DNA virome: topographical and temporal diversity, genetic enrichment, and dynamic associations with the host microbiome. *mBio* 6:e01578-15. doi: 10.1128/mBio. 01578- 15.

Heldal, M., and Bratbak, G. (1991) Production and decay of viruses in aquatic environments. *Mar Ecol Prog Ser* **72**, 205-212.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**, 119.

Joung, Y.S., Ge, Z., Buie, C.R. (2017) Bioaerosol generation by raindrops on soil. *Nat Commun* **8**, 14668.

Kang, Y.J., and Frank. F.J. (1990) Characteristics of biological aerosols in dairy processing plants. *J Dairy Sci* **73**, 621-626.

Kim, Y., Van Bonn, W., Aw, T.G., Rose, J.B. (2017) Aquarium viromes: viromes of human-managed aquatic systems. *Front Microbiol* **8**, 1231.

Lupo, A., Coyne, S., Berendonk, T.U. (2012) Origin and evolution of antibi- otic resistance: the common mechanisms of emergence and spread in water bodies. *Front Microbiol* **3**, 18. http://dx.doi.org/10.3389/fmicb .2012.00018.

Mann BA, Slauch JM. (1997) Transduction of low-copy number plasmids by bacteriophage P22. *Genetics* **146**, 447–456.

Marcó, M.B., Moineau, S., Quiberoni, A. (2012) Bacteriophages and dairy fermentations. *Bacteriophage* **2**, 149–158.

McArthur, A.G., Waglechner, N., Nizam, F., Yan, A., Azad, M.A., Baylay, A.J., Bhullar, K., Canova, M.J., De Pascale, G., Ejim, L., Kalan, L., King, A.M., Koteva, K., Morar, M., Mulvey, M.R., O'Brien, J.S., Pawlowski, A.C., Piddock, L.J.V., Spanogiannopoulos, P., Sutherland, A.D., Tang, I., Taylor, P.L., Thaker, M., Wang, W., Yan, M., Yu, T., Wright, G.D. (2013) The Comprehensive Antibiotic Resistance Database. *Antimicrob Agents Chemother* **57**, 3348-3357.

McIntyre, K., Heap, H.A., Davey, G.P., and Limsowtin, G.K.Y. (1991) The distribution of lactococcal bacteriophage in the environment of a cheese manufacturing plant. *Int Dairy J* **1**, 183-197.

Modi, S.L., Lee, H.H., Spina, C.S., Collins, J.J. (2013) Antibiotic tretment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* **499**, 219-23.

Muniesa, M., Colomer-Lluch, M., Jofre, J. (2013) Could bacteriophages transfer antibiotic resistance genes from environmental bacteria to human-body associated bacterial populations? *Mob Genet Elements* e25847.

Neve, H., Berger, A., and Heller, K.J. (1995) A method for detecting and enumerating airborne virulent bacteriophage of dairy starter cultures. *Kieler Milchw Forsch* **47**,193-207.

Newton, H.J., Ang, D.K., van Driel, I.R., and Hartland, E.L. (2010). Molecular pathogenesis of infections caused by Legionella pneumophila. *Clin Microbiol Rev* **23**, 274–298.

Qiu, J., Luo, Z-Q. (2017). *Legionella* and *Coxiella* effectors: strength in diversity and activity. *Nat Rev Microbiol* doi: 10.1038/nrmicro.2017.67.

Ryndock, E.J., Meyers, C. (2014) A risk for non-sexual transmission of human papillomavirus? Expert Rev. *Anti-Infect Ther* **12**, 1165–1170.

Salustiano, V.C., Andrade, N.J., Brandão, S.C.C., Azeredo, R.M.C., Lima, S.A.K. (2003) Microbiological air quality of processing areas in a dairy plant as evaluated by the sedimentation technique and a one-stage air sampler. *Braz J Microbiol* **34**, 255-259.

Schmieger H, Schicklmaier P. (1999) Transduction of multiple drug resistance of *Salmonella enterica* serovar *typhimurium* DT104. *FEMS Microbiol Lett* **170**, 251–256.

Thurber, R.V., Haynes, M., Breitbart, M., Wegley, L., Rohwer, F. (2009) Laboratory procedures to generate viral metagenomes. *Nat Protoc* **4**, 470–483.

Van Heijnsbergen, E., Schalk, J.A., Euser, S.M., Brandsema, P.S., den Boer, J.W., de Roda Husman, A.M. (2015) Confirmed and Potential Sources of *Legionella* Reviewed. *Environ Sci Technol* **49**, 4797-4815.

Verreault, D., Moineau, S., Duchaine, C. (2011) Methods for sampling airborne viruses. *Microbiol Mol Biol Rev* **72**, 413-444.

Verreault, D., Gendron, L., Roussueau, G.M., Veillette, M., Massé, D., Lindsley, W.G., Moineau, S., Duchaine, C. (2011) Detection of airborne lactococcal bacteriophages in cheese manufacturing plants. *Appl Environ Microbiol* **77**, 491-497.

Vickers, V.T.(1986) Control of airborne contamination in dairy processing plants. *N Z J Dairy Sci Technol* **21**, 89-98.

Table 1 Taxonomic affiliations* at family levels of the reads from metagenomic
shotgun sequencing of the three viral samples, assigned by BLAST against
our customized viral NCBI taxonomy database

Family	Grana padano air sample	Gorgonzola day 1 air sample	Gorgonzola day 50 air sample	Viral Target
Iridoviridae	1%	0%	3%	Amphibia, Fish, Invertebrates, Lepidoptera, Orthoptera
Marseilleviridae	1%	2%	2%	Amoeba
Microviridae	3%	0%	0%	Bacteria
Mimiviridae	14%	25%	13%	Amoeba and other protists
Myoviridae	17%	25%	20%	Bacteria; Archaea
unclassified dsDNA viruses	24%	18%	18%	
Papillomaviridae	0%	0%	17%	Mammals
Phycodnaviridae	12%	19%	9%	Eukaryotic algae
Podoviridae	14%	0%	2%	Bacteria
Siphoviridae	12%	6%	12%	Bacteria; Archaea

* Taxonomic affiliation was assigned by BLAST with a threshold of 10^{-5} for the E
value against the NCBI taxonomy database.

Table 2 Taxonomic affiliations at phylum level of the reads from metagenomic shotgun sequencing in viromes*

Phylum	Grana padano air sample	Gorgonzola day 1 air sample	Gorgonzola day 50 air sample
Bacteroidetes	26%	61%	51%
Firmicutes	58%	5%	2%
Proteobacteria	9%	24%	29%
Chlamydiae	1%	1%	8%
Other (<3%)	1%	5%	6%
Unassigned	0%	0%	5%

* Taxonomic affiliation was assigned by BLAST with a threshold of 10^{-5} for the E value against the NCBI taxonomy database.

Table 3   Taxonomic affiliations at class levels of the reads from metagenomic shotgun

sequencing in viromes

Class	Grana padano air sample	Gorgonzola day 1 air sample	Gorgonzola day 50 air sample
Alphaproteobacteria	2%	6%	11%
Bacteroidia	19%	2%	1%
Bacilli	38%	3%	1%
Betaproteobacteria	0%	3%	2%
Chlamydiia	1%	1%	8%
Clostridia	19%	2%	1%
Cytophagia	3%	13%	23%
Flavobacteriia	3%	16%	6%
Gammaproteobacteria	5%	14%	15%
Sphingobacteriia	1%	30%	20%
Other (< 2%)	2%	8%	9%
Unclassified	7%	6%	7%

* Taxonomic affiliation was assigned by BLAST with a threshold of 10^{-5} for the E

value against the NCBI taxonomy database.

Table 4   Distribution of reads identified as ARGs expressed as relative abundance with respect to the total ARGs-related reads identified

Gene function	ARG	Grana padano air sample	Gorgonzola air sample	Gorgonzola day 1 air sample	Gorgonzola day 50 air sample
Antibiotic efflux	Efflux pump conferring antibiotic resistance	39%	26%		34%
Antibiotic efflux	Gene modulating antibiotic efflux	17%	9%		14%
Protein synthesis inhibitor	Elfamycin resistance gene	9%	3%		1%
Protein synthesis inhibitor	Tetracycline resistance gene	9%	3%		1%
DNA synthesis inhibitor	Fluoroquinolone resistance gene	4%	0%		1%
Cell wall synthesis inhibitor	Glycopeptide resistance gene	4%	3%		9%
Cell wall disruption	Lipopeptide antibiotic resistance gene	4%	14%		8%
Cell wall disruption	Polymyxin resistance gene	4%	0%		2%
RNA synthesis inhibitor	Rifampin resistance gene	4%	0%		1%
Protein synthesis inhibitor	Streptothricin resistance gene	4%	0%		1%
DNA synthesis inhibitor	Aminocoumarin resistance gene	0%	14%		7%
Protein synthesis inhibitor	Aminoglycoside resistance gene	0%	3%		3%
Cell wall synthesis inhibitor	Beta-lactam resistance gene	0%	6%		8%
Protein synthesis inhibitor	Chloramphenicol resistance gene	0%	14%		3%
Protein synthesis inhibitor	Lincosamide resistance gene	0%	0%		4%
RNA synthesis inhibitor	Peptide antibiotic resistance gene	0%	0%		1%
Protein synthesis inhibitor	Streptogramin resistance gene	0%	3%		1%
DNA synthesis inhibitor	Sulfonamide resistance gene	0%	0%		1%
DNA synthesis inhibitor	Trimethoprim resistance gene	0%	3%		1%

Figure 1. Flow chart of the methodological approach applied for airborne virome isolation and characterization.

Figure 2. Taxonomic affiliations at kingdom levels of the reads from metagenomic shotgun sequencing of the three viral samples, assigned by BLAST against our customized viral NCBI taxonomy database. Virus (white bars); Eukaryota (gray bars); Bacteria (black bars).

Figure 3. Comparison of the distribution of antibiotic resistance-related reads in viromes. Distribution of the drug classes is based on the inhibition target of the antibiotic. (■) Antibiotic efflux; (■) Cell wall distruption and synthesis inhibitor; DNA synthesis inhibitor (■); Protein synthesis inhibitor (■); RNA synthesis inhibitor (■).

Figure 4. Distribution of phage-hosts association. Indicated percentages are in comparison to the total amount of identified phage-related ORFs identified. Bacteria were grouped based on taxonomic level of phyla. Black bars refer to Grana Padano air sample. Gray bars refers to Gorgonzola day 1 air sample. White bars refers to Gorgonzola day 50 air sample.

Sample 30 m³ air surrounding Grana Padano,
Gorgonzola curd (day 1) and Gorgonzola rind (day 50) air samples

Sartorious filtering 0.2 μm (microbes removal)   Virus like particle (VLP) isolation

PEG 8000 precipitation

CsCl Ultracentrifugation

DNAse treatment for exogenous DNA removal   VLP DNA extraction

Ethanol washing

Proteinase K treatment

Phoenol/Chloroform extraction

DNA precipitation with isopropanol

Contaminants control (qPCR) ← VLP DNA → Shotgun Metagenomic sequencing (Miseq Illumina)

Viral custom database   Bioinformatic analysis

NCBI nr database

CARD database