UNIVERSITÀ DEGLI STUDI DI MILANO

Ph.D. SCHOOL IN FOOD SYSTEMS

Department of Food, Environmental and Nutritional Sciences

XXIX Cycle

Disclosing the phage-mediated antibiotic resistances in the food chain (diPHARE-FOOD)

[AGR 16]

STEFANO COLOMBO
(R10539 –R34)

Tutor: Prof. Diego Mora

Ph.D. Dean: Prof. Francesco Bonomi

2016/2017

Indice

1. State Of The Art 4
  1.1 The importance of antibiotics 4
  1.2 The antibiotic resistome 5
  1.3 Mobilization of antibiotic resistance determinants 6
    1.3.1 HGT: Transformation 7
    1.3.2 HGT: Conjugation 8
    1.3.3 HGT: Transduction 8
  1.4 Ubiquity and abundance of bacteriophage 9
  1.5 Food safety: Virus hazards from food, water 10
2. Aims and Rationale 12
3. Virome-associated antibiotic-resistance genes in an experimental aquaculture plant 13
  3.1 Introduction 13
  3.2 Results 13
    3.2.1 Viral taxonomic characterization 13
    3.2.2 Distribution and taxonomy of ARGs in the virome and microbiome 15
    3.2.3 The virome and microbiome share ARGs 18
    3.2.4 ARGs present in publically available viromes 19
  3.3 Discussion 19
4. Exploring microbiomes and viromes for antibiotic-resistance genes in aquatic environments 22
  4.1 Introduction 22
  4.2 Results 23
    4.2.1 16s RNA gene profiling analysis 23
    4.2.2 Taxonomic description of microbial communities using the shotgun metagenomic approach 25
    4.2.3 Taxonomic description of viral communities using the shotgun metagenomic approach 26
    4.2.4 Functional characterization of microbiome data 28
    4.2.5 Functional characterization of viromes data 29
    4.2.6 Identification of antibiotic resistance genes in the microbiome and virome 30
    4.2.7 Identification of heavy metal and biocide resistance genes 32
    4.2.8 Identification of gene coding bacteriocin in microbiomes and viromes 32

4.2.9 Identification of antibiotic resistance genes co-resident with phage-genes in the same sequence contig 33
4.2.10 Frequency of antibiotic resistant colonies in water samples 33
4.3 Discussion 35
5. Evaluation of airborne viruses from two different dairy plants with metagenomic approach 39
5.1 Introduction 39
5.2 Results 41
5.2.1 16S rRNA gene profiling analysis 41
5.2.2 Viral shotgun metagenomic – Taxonomy 43
5.2.3 Viral shotgun metagenomic – Functional classification 45
5.3.4 Identification of the antibiotic resistance genes in air viromes 47
5.3.5 Identification of Legionella related ORFs 48
5.3.6 Phage – host interaction 49
5.3 Discussion 50
6. Products 52
7. Conclusion 60
8. Materials And Methods 61
8.1 Sampling 61
8.1.1 Sampling of wastewater from an aquaculture plant (Chapter 3) 61
8.1.2 Sampling and evaluation of antibiotic-resistant colony frequencies in water samples (Chapter 4) 61
8.1.3 Sampling of air an cheese related matrices in two dairy plant (Chapter 5) 61
8.2 Virus-like particle isolation 62
8.3 Extraction and amplification of VLP DNA for metagenomic analysis 62
8.4 Calculation of bacterial contamination using qPCR 63
8.5 Extraction of microbial DNA for metagenomic analysis 64
8.6 16S rRNA gene amplification 64
8.7 Ion Torrent PGM sequencing of *16S rRNA* gene-based amplicons 65
8.8 16S rRNA-based microbiota analysis 65
8.9 Metagenomic shotgun sequencing of microbial and viral DNA 66
8.10 Quality control of metagenomic reads 66
8.11 Bioinformatic analysis 67
9. Supplementary Tables 68
Table S1. Taxonomic abundance of ORFs (BLASTp with a threshold of 10^5 for the E value) identified by MetaVir in the aquaculture virome 68

Table S2. ARGs genes shared between viral and microbial metagenomes 69
Table S3. Relative abundance of ARGs, with respect to the total amount of ORFs identified, in several freshwater viromes. (BLASTx with a threshold of 10^5 for the E value against CARD database). 70
Table S4. Relative abundance of reads associated to MRGs and biocide resistant genes.......................... 71
Table S5. Relative abundance of reads associated to bacteriocin genes... 73
Table S6. Virome-contigs harboring phage genes and ARGs 74
Table S7. Relative abundance of ARGs, with respect to the total amount of ORFs identified, in several freshwater viromes. (BLASTx with a threshold of 10^5 for the E value against CARD database). 75
10. References.. 76

1. State Of The Art

1.1 The importance of antibiotics

Antibiotics are life-saving medications, and are considered one of the most important medical discoveries of all time. Antibiotics, also called antibacterials, are a type of molecules used to treat and prevent bacterial infections. They may either kill (bactericidal effect) or inhibit the growth of bacteria (bacteriostatic effect). Antibiotics are divided in more than 15 classes whose targets are involved in essential physiological or metabolic functions of the bacterial cell (Table 1) (Levy and Marshall, 2004).

Mechanism of action	Antibiotic families
Inhibition of cell wall synthesis	Penicillins; cephalosporins; carbapenems; daptomycin; monobactams; glycopeptides
Inhibition of protein synthesis	Tetracyclines; aminoglycosides; oxazolidonones; streptogramins; ketolides; macrolides; lincosamides
Inhibition of DNA synthesis	Fluoroquinolones
Competitive inhibition of folic acid synthesis	Sulfonamides; trimethoprim
Inhibition of RNA synthesis	Rifampicin
Other	Metronidazole

Antibiotics have existed in nature for billions of years (Wright and Poinar, 2012), since they were used by microorganism to inhibit or kill other bacteria that would be competitors for environmental resources (Martinez, 2008; Aminov, 2009; von Wintersdorff *et al.,* 2016). In addition to this main function, antibiotics could function as signaling molecules which are involved in developmental processes as biofilm formation (van Schaik, 2015).

In addition to naturally present antibiotics, development and large scale production of synthetic antibiotics led to a massive utilization in many human activities as medicine, agriculture and farming. This massive utilization had a large impact on the microbial biosphere. In the work of Van Boeckel *et al*. has been reported that the global human antibiotic consumption amounted to 54,083,964,813 standard units (pills, capsules, or ampoules) in 2000 and had increased by 36% to

73,620,748,816 standard units by 2010 (Van Boeckel *et al.,* 2014). Even worse was the antibiotic consumption in food animals, that exceeded 63 million kg in 2010 and will also drastically increase in the coming years.
The excessive use of antibiotics in medicine, agriculture, and farming has been linked to the emergence of resistant bacteria (Penders and Stobberingh, 2008; Economou and Gousia, 2015). The majority of consumed antibiotics are then released in the environment either directly or through wastewaters; their persistence in the environment lead to the development of a selective pressure that favors the increase and the mobilization of antibiotic resistance genes both in non-pathogenic and pathogenic bacteria (von Wintersdorff *et al.,* 2016).

1.2 The antibiotic resistome

As antibiotic appeared and developed since ancient time, antibiotic resistance genes (ARGs) have been developed since the beginning of life. Evidence can be found studying microbiomes of particular soil (i.e. permafrost) or areas (isolated cave) where no human activities has ever be found and interaction with civilized areas was not possible (D'Costa *et al.,* 2011; Perron *et al.,* 2015).
Since spreading of antibiotic resistance genes is gaining more and more importance in the last decades, study of resistomes, the collection of all the antibiotic resistance genes in a particular environment is skyrocketing. Analysis of resistomes has also underlined the importance of microorganism that acts as reservoir of genes that favors the mobilization of the genes itself (viruses) (von Wintersdorff *et al.,* 2016).
For most antibiotics, the appearance of resistance occurs through the development of extremely efficient enzymes, efflux proteins and other transport systems that are the results of thousands years of evolution (Bhullar *et al.,* 2012).
Genes that confer resistance to antibiotic may have an entirely different function in its original bacterial host: a classic example is the 20-N-acetyltransferase encoding gene in *Providencia stuartii*. Even if the original function was the modification of peptidoglycan, since aminoglycoside antibiotics are structurally similar to the natural substrate of 20-N-acetyltransferase, the enzyme can also inactivate aminoglycosides, providing intrinsic resistance to this class of antibiotics to *P. stuartii* (Macinga DR and Rather, 1999). Mobilization of genes that confer resistance to antibiotics despite a totally different main function gives a meaningful boost to the proliferation of antibiotic resistant bacteria (van Schaik, 2015).

Sequencing of metagenomes from various environments has provided a large amount of data that are usually freely consultable. Such data can be mined for the presence of resistance genes, even if it was not the initial purpose of the studies (von Wintersdorff *et al.,* 2016).
Studies on different environment has let to define that ARGs predominantly cluster by ecology, implying that the resistome in soils, and wastewater treatment plants differ significantly from that of human pathogens (Gibson *et al.,* 2015; Munck *et al.,* 2015). On the other hand different ARGs identified in soil ecosystems showed a perfect nucleotidic match with ARGs identified in human pathogens probably due to gene exchange, so the two resistomes even if very different, are characterized by shared genes (Forsberg *et al.,* 2012; von Wintersdorff *et al.,* 2016).

The resistome includes all the genes present in a specific environment that can confer resistance both to pathogens in the clinic, but also to other nonpathogenic species.
The 'resistome' concept should also include proto-resistance genes, which are precursor genes that could one day turn in actual resistance genes. In addition should be considered the so-called 'silent' resistance genes, which, like proto-resistance genes, do not cause phenotypic resistance until their expression changes through mobilization or mutation to associated regulatory elements (Perry *et al.,* 2014a). Nucleotide polymerase could be considered a proto-resistance genes since has been demonstrated that the structure of this enzyme is extremely similar to lincosamide and aminoglycoside nucleotidyltranferases underlining a close evolutionary relationship. In the light of this study, the ancestral nucleotide polymerases were proto-resistance genes that evolved into these antibiotic-modifying genes (Morar *et al.,* 2009 ; Perry *et al.,* 2014a).

1.3 Mobilization of antibiotic resistance determinants

Resistance to antibiotics can occur either by mutations or by acquisition genes that confers resistance via horizontal gene transfer (HGT). Acquisition of resistance by mutation is a process that could potentially create unlimited and more and more effective mechanisms that confer resistance, but on the other hand, it requires high amount of time in order to effectively select the most efficient ones. Probably the most effective system to transfer ARGs in an environment is HGT since the efficient gene are already present. Even if it is an ancient mechanism, due to the increasing presence of a selective pressure generated by man released antibiotics, HGT has given an impressive boost to the spreading of ARGs and consequently to the number of resistant strains. (von Wintersdorff *et al.,* 2016). To underline the importance that HGT events have on the human life, it has recently been reported that human-associated bacteria are 25-fold more probable to undergo

HGT than ecologically diverse non-human isolates (Smillie *et al.,* 2011; Perry *et al.,* 2014b). This can be explained by the fact that bacteria that inhabit the same niche (like in the human body) are often closely related and therefore highly susceptible to gene exchange, compared to those present in complex environment that are usually more taxonomically different. In the light of these consideration must be kept in mind that even if individual HGT events may be rare, the density of bacteria in the soil suggests that the sum total of HGT events is high, even with genetic barriers to gene exchange. Once the gene has been transferred is then necessary the presence of a selective pressure in order to maintained the acquired gene. Without selective pressure in the environment, the acquired gene can be lost by simple genetic drift.

If the acquired gene confers a selective advantage, it will fix and cause a rapid expansion in the population carrying the advantageous gene. In this scenario, (1) all genes conferring resistance will be clonal, and (2) other genes conferring similar selective advantages will be less likely to fix in the population because the genetic environment is already saturated by the "colonizing" HGT event (Perry *et al.,* 2014b).

When HGT occurs through a plasmid (conjugation) or bacteriophage vector (transduction), host and recipient must be compatible for DNA expression and replication purposes, otherwise DNA will undergo destruction by restriction modification system. Once in the cell, transferred DNA must either replicate autonomously (such as with a plasmid) or integrate into the host chromosome (Wiedenbeck and Cohan 2011; Perry *et al.,* 2014b).

HGT events occur in three different ways: transformation, conjugation and transduction.

1.3.1 HGT: Transformation

Some bacteria are capable of uptake, integrate and express naked fragments of extracellular DNA, in a process called transformation. A classic example is the transfer of ARGs in order to evade antibiotics. Transformation can occur only if several condition are met: 1) there must be DNA present in the extracellular environment; 2) the recipient bacteria must be in a state of competence; 3) the translocated DNA must be stabilized, either by integration into the recipient genome, or by recircularization (Thomas and Nielsen, 2005). Has been also reported that exposure to antibiotics can induce competence in many species of bacteria underlining that antibiotic can stimulate the transfer of the respective ARGs (von Wintersdorff *et al.,* 2016).

1.3.2 HGT: Conjugation

Conjugation is the transfer of DNA through a multi-step process requiring cell to cell contact via cell surface pili or adhesins. It is facilitated by the conjugative machinery which is encoded either by genes on autonomously replicating plasmids or by integrative conjugative elements in the chromosome. Compared to the transformation, conjugation provides better protection from the surrounding environment and a more efficient means of entering the host cell, while with respect to transduction it often has a broader host range (von Wintersdorff *et al.,* 2016).

1.3.3 HGT: Transduction

In the transduction, transfer of genetic material is mediated bacteriophages and the transferred genes are advantageous to their microbial hosts, in turn promoting their own survival and dissemination (Modi *et al.,* 2013). Recent studies applying metagenomic approaches to samples from various environments have suggested that bacteriophages may play a bigger part in the spread of ARGs than previously recognized as shown in studies regarding subtropical freshwater (Parsley *et al*., 2010; Tseng *et al*., 2013), activated sludge or human gut (Minot *et al.,* 2011).
This system let genetic material to be introduced into a bacterium by a phage that has previously replicated in another bacterium, in which it packaged random DNA fragments (generalized transduction) or the DNA adjacent to the prophage attachment site (specialized transduction). The size of the DNA fragments that can be packaged into a bacteriophage particle is limited by the size of the phage capsid, but can reach upwards of 100 kb. All kind of DNA both chromosomal fragments and mobile elements (plasmids, transposons and insertion elements) could be transferred with transduction (Schmieger and Schicklmaier, 1999; Mann and Slauch, 1997). Once inside the new cell, the transferred DNA must escape degradation by the bacterial restriction systems of the cell, and then be incorporated into the recipient's genome. Incorporation can be achieved either by homologous recombination or integration (Muniesa *et al.,* 2013b).
With respect to transformation and conjugation, transduction can have a larger impact since it does not require 'donor' and 'recipient' cells to be present at the same place or even at the same time; beyond that bacteriophages are resistant to very harsh condition (temperature, salinity, radiation), much more than free DNA, making them particularly adapt to transfer genetic material among different biomes (Sano *et al.,* 2004).
In addition, transduction has so far been considered as a rare event, occurring approximately once every 10^7–10^9 phage infections (Bushman 2002). But recent works showed that transduction evens

might occur at frequencies several orders of magnitude greater than previously thought (Chiura 1997; Kenzaka 2010; Muniesa *et al.,* 2013b).

In a recent work, Modi *et al*. demonstrated that mices treated with antibiotics increased the number of ARGs in the intestinal phageome compared to the control showing that the selective pressure generated by the antibiotic increase the number of ARGs transferred to the phages that could mobilized again (Modi *et al.,* 2013). Furthermore, other studies based on qPCR analysis have shown to detect ARGs in bacteriophages underlining the significant role of reservoirs of ARGs(Colomer-Lluch *et al.,* 2014a,b; Colomer-Lluch *et al.,* 2011b; Marti *et al.,* 2014).
Different studies have reported that certain bacteriophages showed wide host range of infection that crosses between different species or even different taxonomic classes, underlining the large impact that this mechanism has in the spread of ARGs (Mazaheri Nezhad Fard *et al.,* 2011; von Wintersdorff *et al.,* 2016).

1.4 Ubiquity and abundance of bacteriophage

Bacteriophages, or phages, are viruses that infect bacteria. Bacteriophages are probably the most abundant life form on Earth in fact they use to outnumber bacteria by a range from 1 to 10(Muniesa *et al.,* 2013a).They have a very complex role in the ecosystems: on the one hand, by infecting and lysing infected bacteria, they contribute remarkably to bacterial mortality; for example, up to 15% in the case of bacterioplankton (Suttle 1994). Bacterial lysis they are responsible of, release in the environment large amount of organic compounds that has an important impact on the cycling of organic matter in the biosphere at a global level, so on all the other organisms that have not been directly in contact with the phage itself. On the other hand, they control microbial diversity by selecting for some types of bacteria that are resistant to their attacker (Scanlan 2012), thus changing the proportions of bacterial species or strains in a community, and consequently influencing the evolution of bacterial genomes through horizontal gene transfer by transduction (Muniesa *et al.,* 2013b).
Bacteriophages can only replicate in a susceptible host cell and they present two different life cycles, the lytic and the lysogenic. In the lytic cycle, following infection the bacteriophage redirects the host metabolism towards the production of new phages that are released by lysis of the host cell. Bacteriophages that can only follow the lytic cycle are known as virulent bacteriophages.
Other bacteriophages, known as temperate bacteriophages, can follow the lysogenic cycle, in which the genome of the temperate phage remains in the host, replicating along with the host, either

integrated in the cell chromosome or as an independent replicon. At this stage, the bacteriophage is known as a prophage, which can be induced to follow the lytic cycle. Induction occurs either spontaneously or when stimulated by inductors. Lysogenic inductors can be natural, such as host starvation (Muniesa *et al.,* 2013b).
Evidence that confirm that the majority of viruses in the viral fraction of most environments are bacteriophages, came from metagenomic studies that have shown that a large proportion of viral particles contain bacterial DNA sequences (Fancello *et al.,* 2013; Roux *et al.,* 2012; Rosario *et al.,* 2009; López-Bueno *et al.,* 2009; Modi *et al.,* 2013). While phages do not need these genes for their replication, they probably give phages or their host a selective advantage (Muniesa *et al.,* 2013b).

1.5 Food safety: Virus hazards from food, water

In addition to bacteriophages that has seen before can have an impact on the human life in an indirect since they can increase the risk of pathogenic bacteria, it must be considered also viruses that have a direct impact on human life.
Food and environmental virology focused the attention mainly on that can be transmitted through water, sewage, soil, air, (matrices where microbial pathogens can be easily found) or food (due to the direct impact on the human life)(Bidawid *et al.,* 2009). Most of the viruses considered are enteric viruses transmitted via the fecal–oral route; it must be considered that infected humans can excrete large amounts of human pathogenic viruses. Usually these viruses are nonenveloped and thus very stable in the environment (Rzez˙utka & Cook, 2004) and include major etiological agents, some of which are thought to be emerging zoonotic pathogens. Due to the high difficulty to remove these viruses with sewage treatment, they can be cause of contamination of the environment from treated as well as untreated wastewater. In addition it should be consider the direct fecal contamination of the environment from humans and animals, for example by bathers or by defecation of free-range or wild animals onto soil or surface waters (Rodríguez-Lázaro *et al.,* 2012). Contaminated water can come in contact after few steps with vegetables and fruit (leafy vegetables, cucumber, tomatoes) as irrigation water and so virus can be reintroduced in the human or animal diet as contaminated foods bringing viral pathogens into human and animal populations (Herman *et al.,* 2015).
Human exposure to even low levels of these pathogenic viruses in the environment, such as norovirus (NoV), can cause infection and disease.
Environmentally transmitted viruses include major etiological agents of mild diseases such as gastroenteritis as well as agents of more severe diseases such as meningitis and hepatitis. Most of

these viruses belong to the families *Adenoviridae, Caliciviridae, Hepeviridae Picornaviridae* and *Reoviridae* (Dubois *et al.,* 1997; Muscillo *et al.,* 2001; Rodríguez-Lázaro *et al.,* 2012)

2. Aims and Rationale

Despite the clear importance that phages, and more in general viruses, have in many biological mechanisms and despite their role in all the ecosystems very little is known. This decade will probably be remembered by the great boost that NGS techniques have given to the study of viruses overcoming all the limitation related to the in vivo study. With this premise it could be interesting to focus our attention to understand the actual impact that viruses have in the mobilization and spreading of microbial genes and in detail, antibiotic resistance genes.

In a context of a Food Systems PhD school, I decided to investigate possible effects and characteristics of viruses, limiting my research to food related environments mainly because they can have a direct and indirect impact on the human health. Water from aquaculture was the first sample analyzed (i) being a matrix that has a strong direct contact with the respective food; secondly, it has been analyzed water from Lambro river (ii) since it is used in the field's irrigation so it directly came in contact with cereals and vegetables. Last project involved the analysis of air surrounding different moment of cheesemaking production (iii). Once again the viruses present in the air could deposit on the surface of the food and they can be ingested by humans.

The three projects had the purpose to fully characterize the respective microbial communities. To do so, I directly took care of the old school microbiology aspects, of the molecular biology parts and the bioinformatical analysis. Through the bioinformatic analysis I focused the attention on the presence of antibiotic resistance genes in the viromes obtaining a level of description that cannot be reached with the wet lab techniques.

Even if the increasing amount of literature involving phages and their interactions with the environment has given a meaningful boost to a better comprehension, due to the complexity of the problem addressed, this project wants to improve our knowledge about phages and their impact on the spreading of antibiotic resistance genes, in order to move forward on the comprehension of this unknown world.

3. Virome-associated antibiotic-resistance genes in an experimental aquaculture plant

3.1 Introduction

According to recently released data (FAO - Fisheries and Aquaculture Department), world aquacultural production of fish for consumption has exceeded 60 million tons in 2011, with an increase of 6.2% over the previous year (FAO 2011). This large growth has been accompanied by an increased usage of a wide range of antibiotics (Armstrong *et al.,* 2005). In aquaculture, the main use for antibiotics is the prevention and treatment of bacterial infections in fish; the prophylactic use of antibiotics is a common practice as well. The necessity of antibiotic use in aquaculture is a consequence of lowered host defenses associated with high-density breeding under suboptimal hygienic conditions (Grave *et al.,* 1999; Defoirdt *et al.,* 2007; Cabello *et al.,* 2013). Despite the potential importance of bacteriophages in transferring resistance genes from the environment to human and animal microbiomes, studies on this topic are limited (Modi *et al.,* 2013). In this work, we report a comprehensive characterization of viral and microbial communities in an experimental aquaculture sample using a metagenomics approach. The contemporary study of both communities resulted in the identification of different genes that are mobilized in the virome, with particular attention paid to antibiotic resistance genes (ARGs).

3.2 Results

3.2.1 Viral taxonomic characterization

Metagenomic shotgun data of the viral DNA extracted from 20 liters of the wastewater of a fish breeding tank were performed with BLASTp against a custom viral database. According to several other studies (Fancello *et al.,* 2012; Roux *et al.,* 2012*;* Zablocki *et al.,* 2014), the majority of the identified ORFs (accession number SAMEA3334740) originated from bacteria (86%) whereas those from viruses were the second most represented (13%). Thirteen families were identified (Table S1, Supporting Information), and prokaryotic viruses were the most abundant in the sample. *Caudovirales* accounted for more than 50% of the total taxonomic abundance. *Myoviridae* and

Siphoviridae were the most represented families with 28% and 18% of the total ORFs, respectively. Analysis revealed the presence of viruses whose hosts include *amoebae* and *algae*; *Mimiviridae* and *Phycodnaviridae* were the main representatives of this group, at 13-15% of the total. Furthermore, the largest variety of phages identified interacts with *Proteobacteria*, which is consistent with the abundance of this phylum in the microbiome (Figure 1).

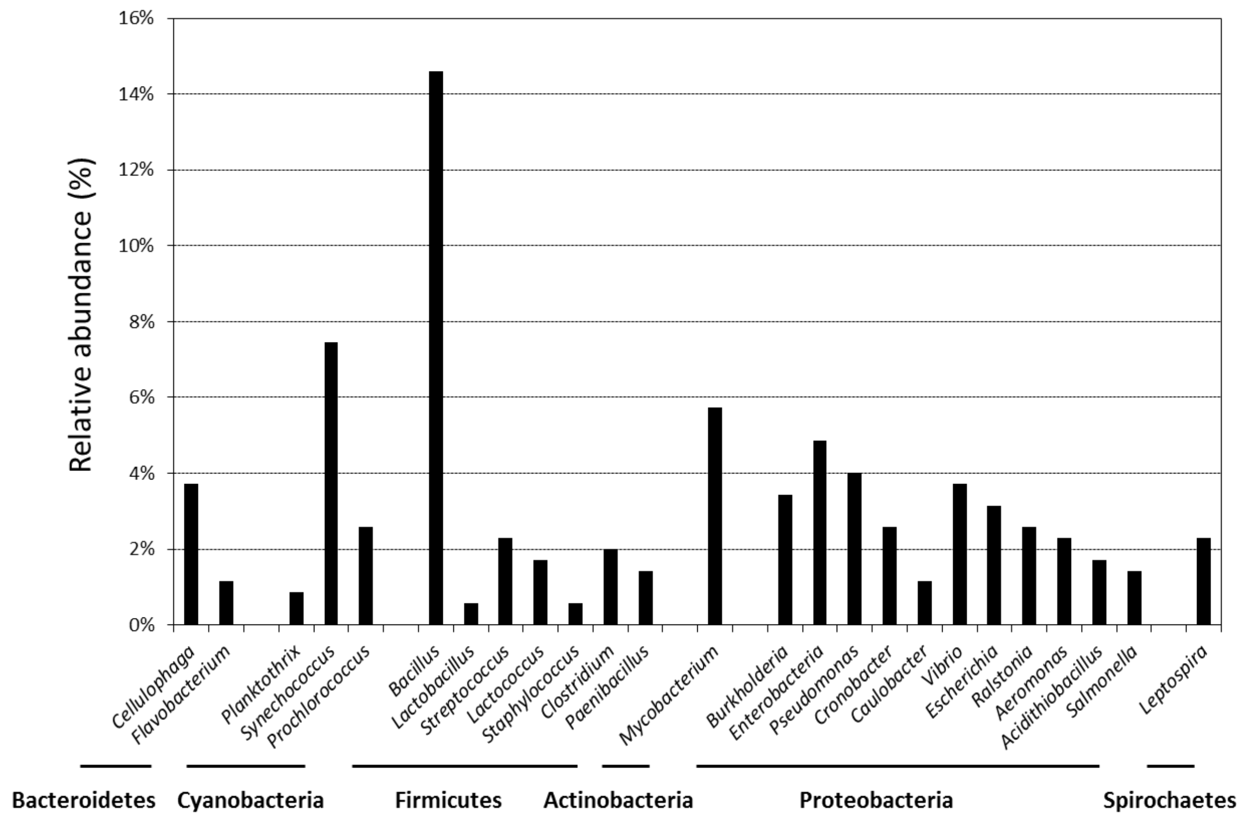


Figure 1: *Distribution of phage-hosts association. Percentages were respect to the total amount of phage-related ORFs identified. Bacteria were grouped based on taxonomic level of phyla.*

The two main families of *Proteobacteria* represented in the water sample, *Sphingomonadaceae* and *Comamonadaceae,* and their related genera are typical constituents of freshwater environments, because bacterioplankton are opportunistic pathogens (Kallman *et al.,* 2006; Kilic *et al.,* 2007; Lin *et al.,* 2010; Vaz-Moreira *et al.,* 2011; Chen *et al.,* 2013). However, the most abundant ORFs identified in the virome were best matched with phages that infect *Bacillus*, *Synechococcus* and *Mycobacterium*, bacterial genera that have not been identified in the microbiome, thus suggesting that these genera were probably underrepresented in the water sample because of a phage infection that resulted in a lytic cycle. In this context, the sharp reduction or local extinction of microbial taxa by viruses is a phenomenon that would be expected according to the "Kill-the-Winner" dynamics

hypothesis (Thingstad 2000). This dynamic model postulates a repetitive cycle in which an increase in prey population leads to an increase in the predator population that in turn decreases the prey population, thus causing its own subsequent decline.

3.2.2 Distribution and taxonomy of ARGs in the virome and microbiome

Microbial cells were isolated to extract their DNA. *16S rRNA* gene profiling (accession number SAMEA3333506) and metagenomics shotgun analysis (accession number SAMEA3334741) were performed on this sample. The ORFs obtained from shotgun metagenomic sequencing were analyzed using the CARD database. The total number of ARGs identified was 950 (4.13% of the total ORFs identified with NCBI) in the microbiome and 214 (2.64%) in the virome. The distribution of the ARG classes identified in the two metagenomes (Figure 2) showed that the most abundant genes were those encoding proteins involved in the modulation of antibiotic efflux pumps (CARD nomenclature), with values ranging from 16% to 19% of the total identified ARGs. Among the different ARG classes, antibiotic efflux pumps had the most general function, acting on different target molecules and having a cell detoxifying activity (Pao *et al.,* 1998). Macrolide-resistance genes were the second most abundant group and included erythromycin, telithromycin and clarithromycin resistance genes. With few exceptions, the distribution of ARGs was similar between microbial and viral metagenomes. Glycopeptide efflux pump and lincosamide ARGs had higher values in the virome than in the microbiome (Figure 2), whereas gene modulating antibiotic efflux pumps and fluoroquinolone resistance genes were the most abundant in the microbiome.

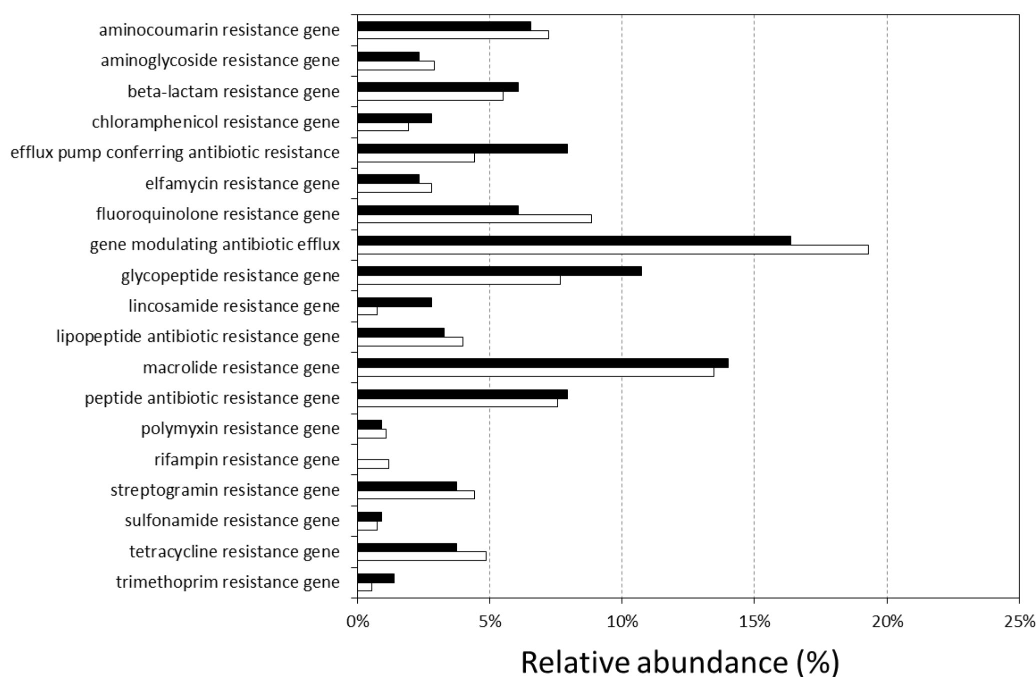


Figure 2: *Distribution of the ARGs identified in the microbial metagenome (white bars) and viral metagenome (black bars) as relative abundance (%).*

For each ARG identified, the taxonomy of the respective ORF, based on the results obtained from the BLAST comparison with the NCBI refseq database, was assigned. The taxonomy of the ARGs in the microbiome (Figure 3) showed the presence of three main families (*Comamonadaceae* 26-37%, *Sphingomonadaceae* 18-24% and *Sinobacteraceae* 14-21%) equally distributed among the different drug classes, and a large number (15-26%) of the ARGs were shared among different microbial families, accounting for an amount lower than 4%. The virome analysis revealed that ARGs within this metagenome were mainly distributed between *Flavobacteriaceae* (38% - 53%) and *Methylophilaceae* (19% - 33%) families and that other bacterial families (7-13%) had a relative abundance below the 5% threshold (Figure 3). None of the families present in both the virome and the microbiome, accounted for a relative abundance higher than 5%.

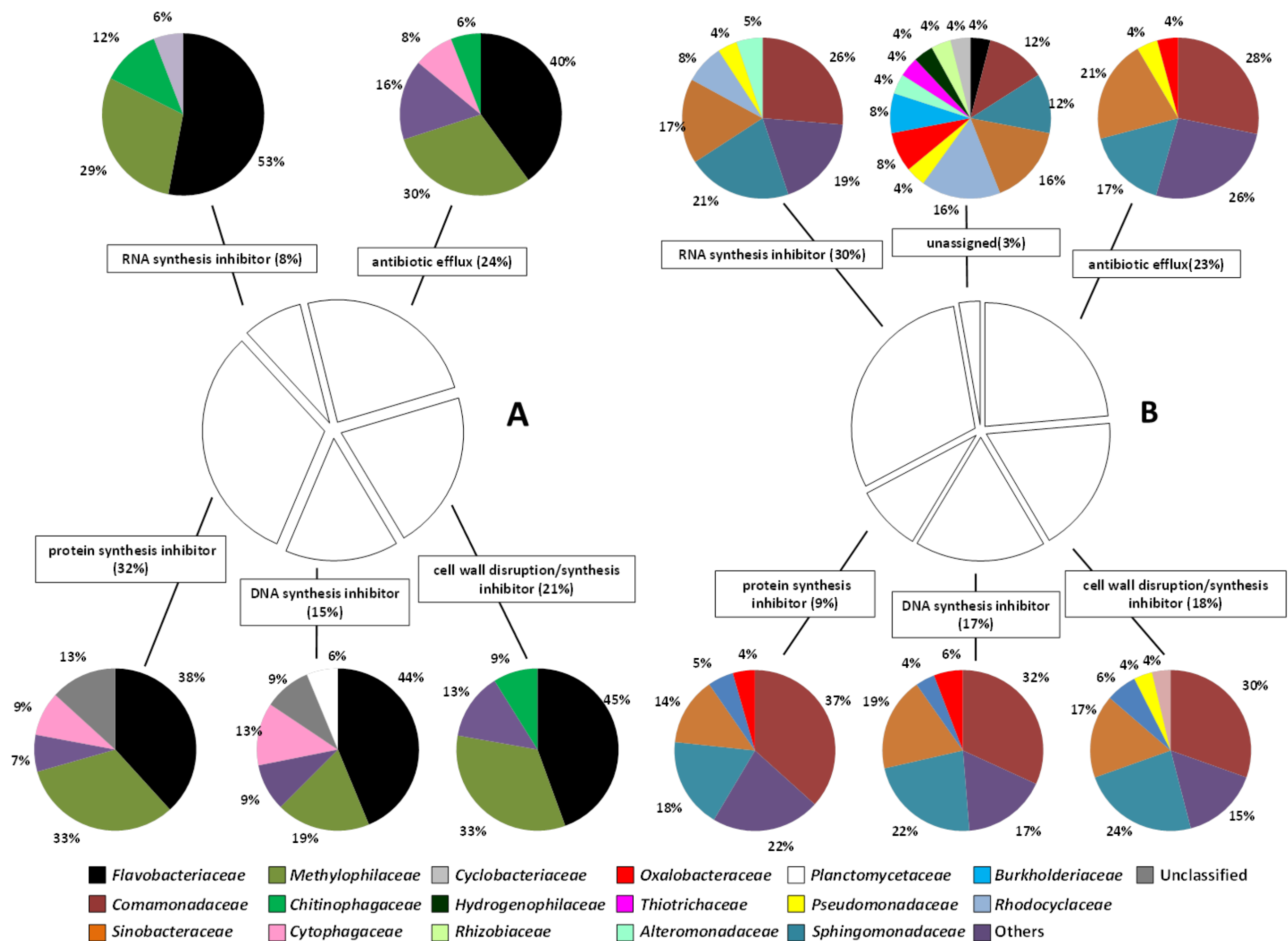


Figure 3: *Comparison of the antibiotic resistance related ORFs distribution among the viral (A) and the microbial (B) metagenome. ARGs have been divided in six different drug classes: RNA synthesis inhibitor, Protein synthesis inhibitor, DNA synthesis inhibitor, Cell wall disruption and synthesis inhibitor, Antibiotic efflux and Unassigned; taxonomy distribution at family level has been associated to each drug class.*

From a taxonomic point of view, it is interesting to note that the microbial ORFs identified in the virome did not reflect the microbiome taxonomy, as determined by *16S rRNA* gene profiling (Figure 4) or shotgun metagenomic analysis (Figure 3), suggesting that microbial genes mobilized in the genome of viruses could be considered to be remnants of past recombination events rather than a picture of the current microbial diversity.

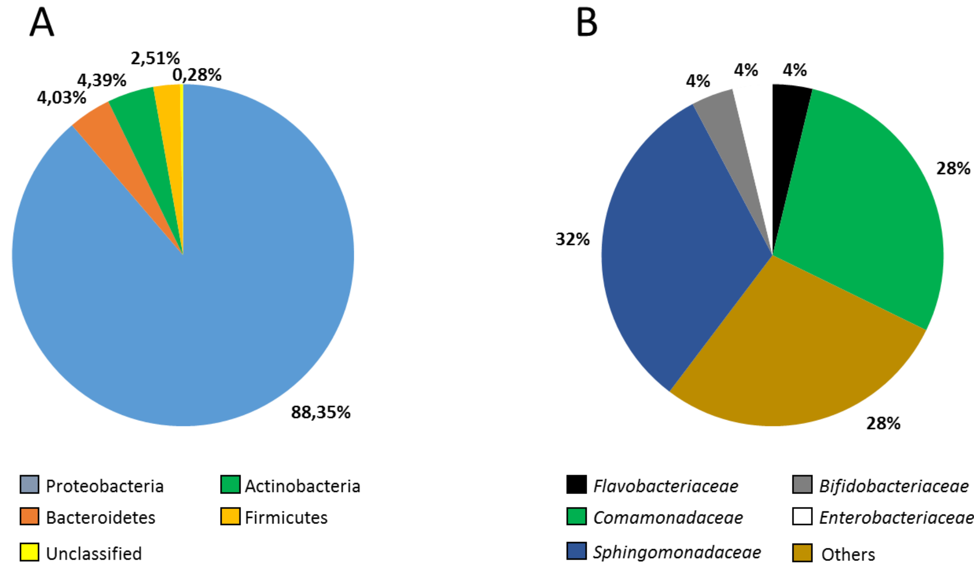


Figure 4: *Taxonomic affiliation based on the 16S rRNA profiling at phylum (A) and family (B) level; GreenGenes 16S database of reference OTUs at 97% identity*

3.2.3 The virome and microbiome share ARGs

In order to identify possible gene mobilization events, ORFs located both in the microbiome and in the virome were identified. A total of 213 different ORFs were shared between the two metagenomes, and most of them were related to microbial metabolism. With respect to the ARGs identified in the virome, 4% of them were shared with the microbiome and had an amino acid sequence identity higher than 95%. Half of the identified ORFs encode for genes modulating the antibiotic efflux (i.e., transcriptional regulator); ORFs that encode for glycopeptide, tetracycline and beta-lactam resistance genes were also identified. Among these genes, one was classified as a *Sphingopyxis alaskensis* β-lactamase coding gene, which was putatively mobilized together with four other genes including a pseudogene coding for a β-lactamase of *Sphingomonas* sp (Table S2). The other five genes shared between the virome and the microbiome are putatively involved in tetracycline and glycopeptide resistance and antibiotic efflux mechanisms. In all cases, shared ARGs were not transferred alone but with other genes hosted by the same contig (Table S2). Our data indicate that ARGs are present in the virome and some of these ARGs are mobilized from bacteria to phages or *vice versa*. These data suggest that ARGs may be mobilized even in the absence of selective pressure, *i.e.,* in the absence of antibiotic treatment in aquaculture. Moreover, we showed that ARGs were mobilized together with other bacterial genes encoding products involved in more general metabolic functions, thus confirming the presence of a complex phage-bacterial network in the aquaculture environment.

3.2.4 ARGs present in publically available viromes

In order to compare our data with publically available data, we calculated the relative amount of ARGs in the 17 viromes tested (Table S3). Our sample showed a relative ARG abundance of 0.85%, the second highest value after that of El Barbera lake (Fancello *et al.,* 2012). Interestingly, the relative amount of ARGs did not correlate with the presence of anthropic activities geographically located near the tested aquatic environments, because the highest ARG abundance was found in the El Barbera sample, which was collected from the Mauritanian desert.

3.3 Discussion

In this paper, we describe an aquaculture sample by means of three different high throughput sequencing analysis: *16S rRNA* gene profiling, shotgun sequencing of the microbial community and shotgun sequencing of VLPs. Shotgun sequencing of viruses revealed that more than 50% of the predicted ORFs belonged to *Caudovirales* order and in particular to the *Myoviridae* family (28%). *Caudovirales* are bacteriophage classified as Group I dsDNA viruses, which are characterized by the presence of a tail. The large abundance of *Caudovirales* ORFs can be plausibly explained by the absolute majority of bacterial ORFs compared with those belonging to Eukarya and Archea (Table S1). It is worth of mention the relative abundance of *Mimiviridae* and *Phycodnaviridae* viruses (41%-43%) that infect amoebae and algae (28%). In contrast to the clear majority of bacterial ORFs in the virome, those related to algae and amoebae were represented below the 0.2%, probably due to the outnumber of bacterial cells with respect to the eukaryotic ones. The largest variety of phages identified interacted with *Proteobacteria* coherently with the abundance of this phylum in the microbiome (Figure 4). Inside *Proteobacteria* the two main families represented in the water sample, *Sphingomonadaceae* and *Comamonadaceae*, and their related genera are typical constituents of freshwater environments, constituting the bacterioplankton or in some cases being recognized as opportunistic pathogens (Charity and Foukas, 2005; Chen *et al.,* 2013). On the other hand, the most abundant ORFs identified in virome had best match score with phages that infect *Bacillus*, *Synechococcus* and *Mycobacterium*, bacterial genera that have not been identified in the microbiome, thus suggesting that these genera were under-represented in the water sample probably due to a phage infection resulting in a lytic cycle. In this context, the sharply reduction or local extinction of microbial taxa by viral predation is a phenomenon that would be expected according to the Kill-the-Winner dynamics hypothesis (Thingstad 2000), that postulates a repetitive cycle in

which an increase in prey population leads to an increase in the predator population that, in turn, decreases the prey population, thus causing its own subsequent decline. From a taxonomic point of view, it is interesting to note that microbial ORFs identified in virome did not reflect the microbiome taxonomy, thus suggesting that microbial genes mobilized in the genome of viruses could be considered as a reminiscence of the past recombination events rather than a picture of the current microbial diversity. One of the aims of this study was the assessment of the entity of ARGs in the aquatic virome in order to understand the relevance of viruses in the spreading of antibiotic resistances in the ecosystem. Our data indicated that ARGs are present in virome and some of these ARGs appeared to be mobilized from bacteria to phages or vice versa. The role of viruses in the mobilization of ARGs by transduction mechanisms have been recently described in a murine model (Modi *et al.,* 2013). The authors showed that the phageome became broadly enriched for functionally beneficial genes under stress-related conditions (e.g. an antibiotic treatment) and that antibiotic treatment expanded the interactions between phage and bacterial cells, leading to a more highly connected phage–bacterial network for gene exchange (Modi *et al.,* 2013). In our data, we showed that ARGs can be mobilized even in the absence of selective pressure, i.e. in the absence of antibiotic treatment in the aquaculture. Moreover, we showed that ARGs were mobilized together with other bacterial genes coding for more general metabolic functions, thus confirming the presence of a complex phage-bacterial network in the aquaculture environment. In order to compare our data with those publically available, the relative amount of ARGs in 17 viromes obtained from freshwater samples was calculated. The sample analyzed in this work was basically groundwater pumped in aquaculture tanks containing approximately 7 kg/m^3 of salmonid (the fish density in a commercial aquaculture is close to 25 kg/m^3) with a water flow of approximately 83 m^3/h. Nevertheless, our sample showed a relative abundance of ARGs of 0.85%, the second highest value after the El Barbera lake (Fancello *et al.,* 2012). Analyzing the data reported in Table S3, it is worth of mention that the relative amount of ARGs did not seems to correlate with the presence of anthropic activities geographically located near to the aquatic environments analyzed, being the highest in ARGs relative abundance found in the El Barbera sample, which was collected from the Mauritanian desert. The presence of ARGs in uncontaminated wild environment have been previously reported. Several studied, in fact, identified ARGs in different ecosystems, from soil and permafrost (D'Costa *et al.,* 2006; Hallen *et al.,* 2010) to human gut (Hu *et al.,* 2013) as recently observed in the microbiome of members of an isolated Yanomami Amerindian village (Clemente *et al.,* 2015). Therefore, the diversity, the taxonomic distribution and the ecological role of antibiotic resistant genes in the environment is still far unknown. Moreover, it should be underlined that the presence of ARGs in metagenomes did not directly represent a risk for human health because a

proper ranking of these risks should be carried out (Martinez *et al.,* 2015). In conclusion, this study addresses for the first time a complete description of microbiome and virome in an aquaculture sample, giving information on the presence of ARGs and their mobilization by transduction mechanisms. In this context, the monitoring of the ARGs in a microbial community could be a useful tool to follow environmental perturbations, particularly in aquaculture where prophylactic use of antibiotics is a common practice.

4. Exploring microbiomes and viromes for antibiotic-resistance genes in aquatic environments

4.1 Introduction

Among the aquatic environment, rivers are the least studied and the main works have analyzed mainly the bacterial community using the *16S rRNA* gene profiling. Water has been described as the major sink for bacteria on earth (Taylor *et al.,* 2011) and bacteriophages outnumber bacteria by a factor ranging from 1 to 10 (Muniesa *et al.,* 2013a). In rivers, a diverse mixture of antibiotics and other pollutants, their metabolites and resistant bacteria, can reach the aquatic environment through treated and untreated sewage, hospital waste, aquaculture discharges, and agricultural run-off. These aquatic compartments, such as water and sediment, may therefore have a significant role in driving ARGs transfer, ecology, and evolution. In this work we developed a method for the comprehensive characterization of viral and microbial communities by shotgun metagenomics focusing the attention on the presence and relative abundance of ARGs across the Lambro River flow, from the spring to the highly urbanized/industrialized Milan area. The Lambro River catchment is located in the north of Milan with a total drainage area of approximately 1,950 km^2. Lambro River's spring is situated in the Pre-Alps (1,450 m above sea level) in the Magreglio area and it flows into a confluence with the Po River (50 m above sea level) with an estimated length of 130 km. The Lambro River catchment average annual rainfall varies between 900 and 1,500 mm (Deksissa *et al.,* 2004). The Lambro River is one of the main tributaries of the Po River, the largest Italian river (71,000 km^2).

The high diversity of land that the river crosses, both agricultural and urbanized, means that its water comes into contact with heavy loads of nutrients, organic matter and toxicants from municipal, industrial and agricultural sources (Viganò *et al.,* 1999; Viganò *et al.,* 2008).

The Lambro River was declared a high environmental risk area in 1987 because of contamination with heavy metals and persistent organic pollutants (Binelli *et al.,* 2010) and in the recent years the situation has not improved. In February 2010, hydrocarbons were spilled into the Lambro River with a high release of oil and its refined products (Giari *et al.,* 2012). For these reasons the river was chosen in the past years as a model of a polluted environment.

Past studies focused their attention on the presence of pollutants, such as pesticides, metals and organochlorines (Pettine *et al.,* 1996; Viganò *et al.,* 1999; Viganò *et al.,* 2001), on the androgenic

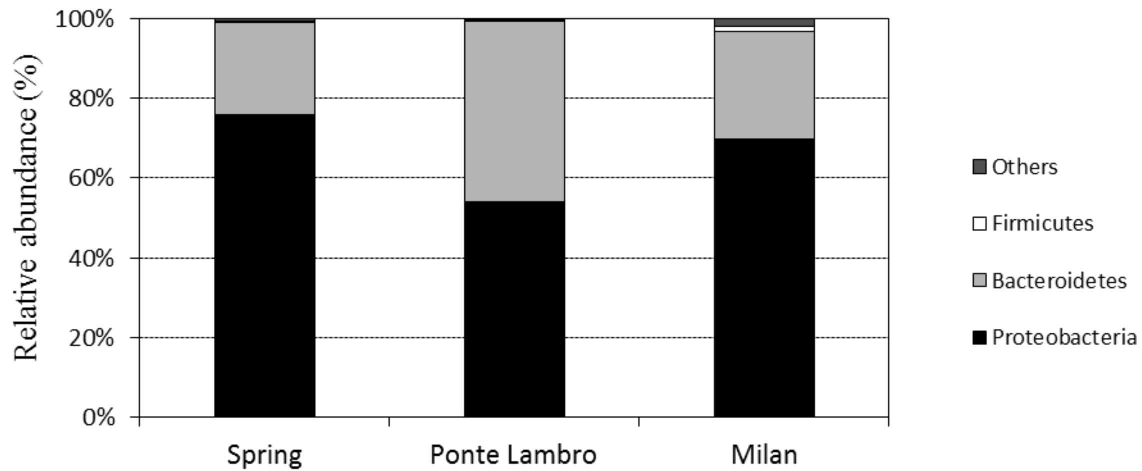
and antiandrogenic activities (Urbatzka *et al.,* 2007), on the effect of accumulation of polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs) and 1,1,1-trichloro-2,2′bis(p-chlorophenyl) ethane (DDT) (Viganò *et al.,* 2007). To carry out these analyses, different model organisms were applied such as *Ceriodaphnia dubia* (Viganò *et al.,* 2008), *Dreissena polymorpha* (Bivalvia) (Binelli *et al.,* 2010) or *Tubifex tubifex* (Oligochaeta) (Crottini *et al.,* 2008). No study has ever taken into consideration the investigation of the microbial and viral communities in this particular model. Here we showed that viromes data collected represents a huge reservoir of microbial genes, ARGs included, that could be a tracer of microbial community fluctuation determined by environmental changes that happened across the river flow.

4.2 Results

4.2.1 16s RNA gene profiling analysis

The *16s RNA* gene profiling analysis has revealed a similar taxonomical distribution at the phylum level between the spring and Milan samples characterized by a higher presence of *Proteobacteria* (75% and 69%, respectively) and *Bacteroidetes* (23% and 27%, respectively) as shown in the Figure 5A. In contrast, the Ponte Lambro sample showed a lower abundance of *Proteobacteria* (53%) in favor of a higher amount of *Bacteroidetes* (45%). 1% of *Firmicutes* was observed in the Milan sample, while other phyla represented less than 1% of abundance. Looking more in depth, the most represented family among the three samples was *Pseudomonadaceae*, up to 50% in the spring sample compared to Ponte Lambro (15%) and Milan (35%). The second most abundant family observed was *Flavobacteriaceae*, mainly represented in the Ponte Lambro sample (41%) compared to the spring (19%) and Milan (14%) samples. Notably, the *Oxalobacteraceae* family was highly represented (9%) in spring sample, whereas *Sphingomonadaceae* (12%) and *Comamonadaceae* (10%) families were both highly represented in the Ponte Lambro sample. In the Milan sample we detected the highest value of families represented below the 2% threshold (Fig. 5B).

A



B

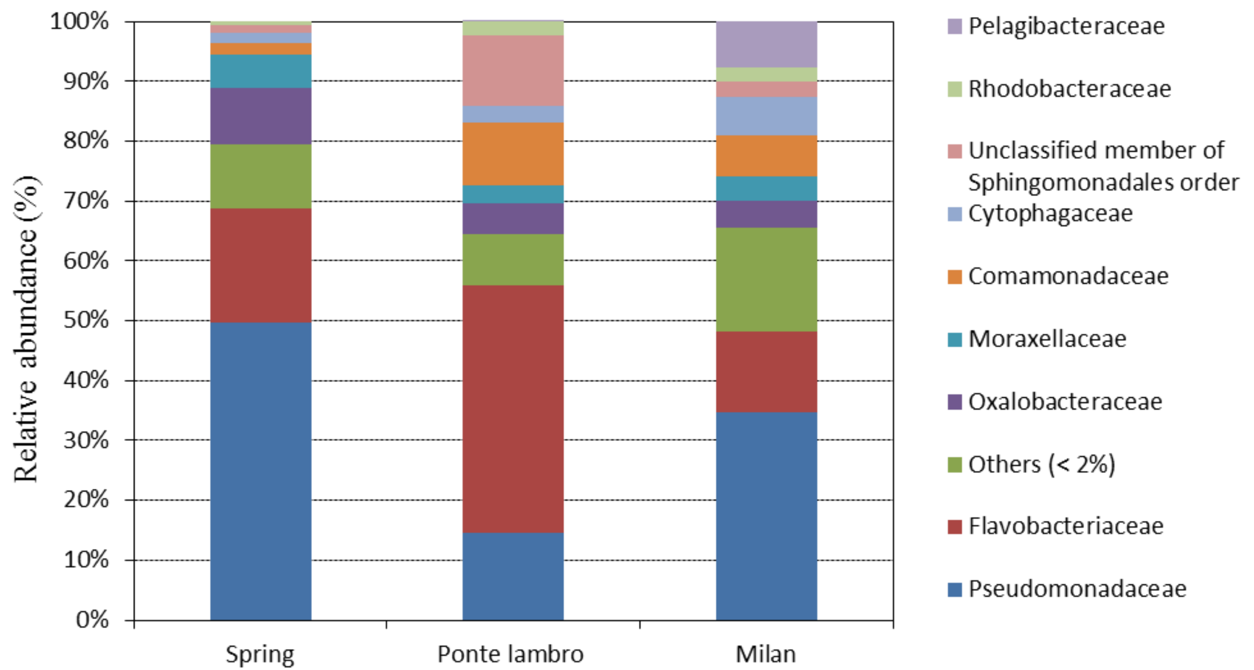


Figure 5: *Taxonomic affiliations based on the 16S rRNA profiling at phylum (A) and family (B) levels, using the GreenGenes 16S database reference OTUs at 97% identity. Data are expressed as relative abundance.*

4.2.2 Taxonomic description of microbial communities using the shotgun metagenomic approach

Shotgun metagenomic sequencing of the microbial populations revealed a slightly different picture compared to that obtained using the *16S rRNA* gene profiling approach. Shotgun metagenomic sequencing of the three samples resulted in a similar and almost complete presence of prokaryotic reads (95%) and a minor presence of eukaryotic reads (< 5%) (Fig. 6A). The spring sample was described by a large portion (53%) of reads identified as generic bacterial sequences without any other taxonomic description. Similar sequences were also observed in the other two samples but in lower amounts (< 15%). The spring sample also showed a homogeneous distribution among *Oxalobacteraceae*, *Sphingomonadaceae*, *Comamonadaceae* and *Flavobacteraceae*, which were represented ranging between 4% and 8% of the total populations. In the Ponte Lambro sample, an equal amount of reads from *Pseudomonadaceae* and *Comamonadaceae* was observed (16%); *Flavobacteraceae* (12%) and *Xanthomonadaceae* (7%) were also detected at meaningful amounts. An increased presence of *Pseudomonadaceae* (26%) was observed in the Milan sample, and bacterial families present at less than 4% each accounted in total for up to 34% of the total population. These data were higher than that observed in the spring (19%) and Ponte Lambro (24%) samples as observed in Figure 6B.

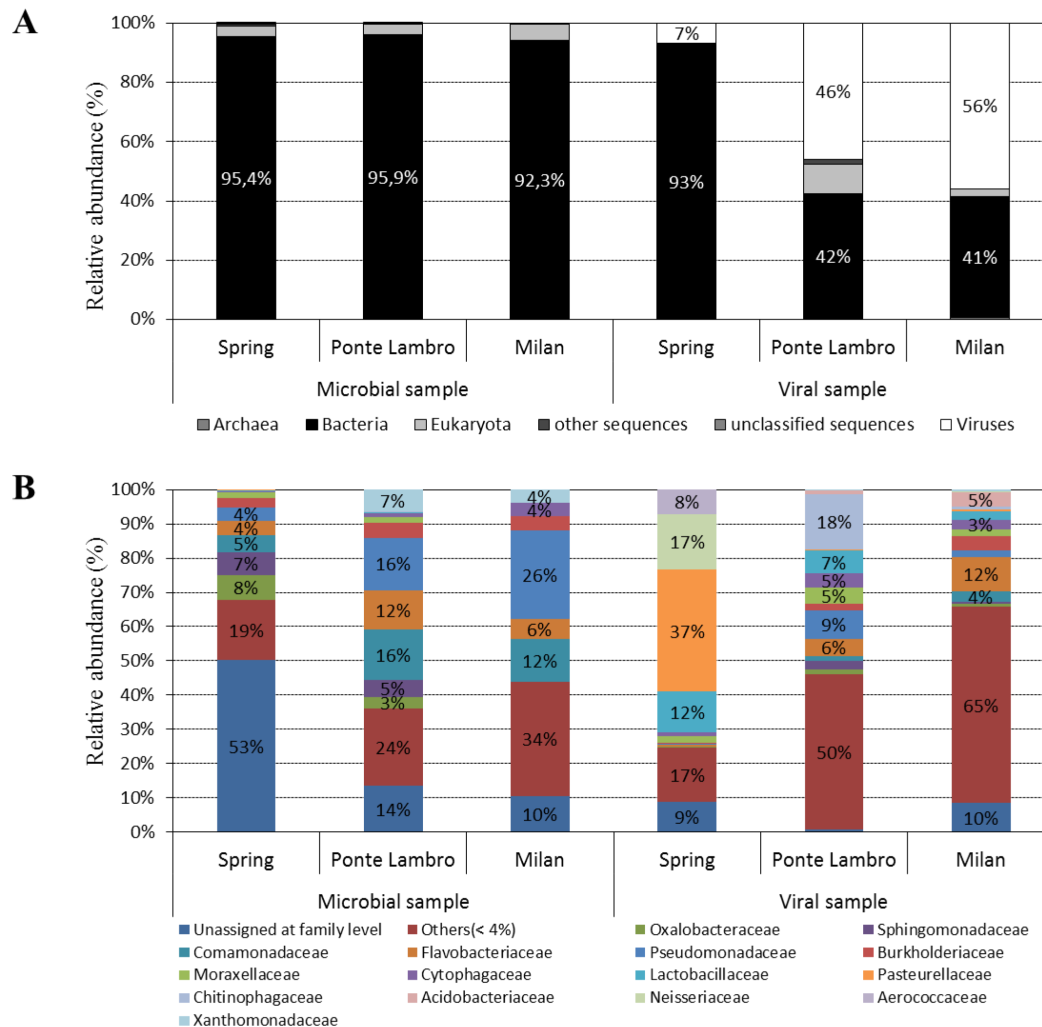


Figure 6: *Taxonomic affiliation at kingdom (A) and family (B) level of the reads from metagenomic shotgun sequencing in microbial and viral metagenomes, assigned by BLAST with a threshold of* 10^{5} *for the E value against NCBI taxonomy database*

4.2.3 Taxonomic description of viral communities using the shotgun metagenomic approach

Identification of viral reads in the viromes of the three samples resulted in two different scenarios (Fig. 6A). The spring sample was characterized by the largest abundance of bacterial reads (93%) and only 7% were identified as viral reads. On the contrary, the other two samples showed a lower amount of bacterial reads (41-42%) and a higher level of the viral reads (46-56%). In addition, we would like to underline that ribosomal sequences were detected with a frequency (n. of ribosomal reads on the total of reads in virome) largely below the 0,2‰ threshold in two of the three viral metadata. Currently, when >0.2 ‰ of reads match to 16S ribosomal DNA, a non-viral marker of bacterial DNA, the viromes are considered to have high bacterial DNA content (Roux *et al.*, 2013).

We obtained level of 0.006 ‰ and 0.009 ‰ of reads match to 16S ribosomal DNA for Ponte Lambro and Milan samples, and 0.35 ‰ for the spring sample. Moreover, in all three samples qPCR quality control assays revealed that the levels of bacterial contaminating DNA in virome samples accounted in less than 0.2% of the total viral DNA extracted, a threshold considered acceptable as suggested by Modi *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 by freshwater samples (Fancello *et al.,* 2013; Roux *et al.,* 2012; Rosario *et al.,* 2009; López-Bueno *et al.,* 2009; Modi *et al.,* 2013). Unfortunately, these studies analyzed the microbial reads only in terms of viral taxonomy without considering the taxonomy and functionality of the microbial gene phage-encoded. In our sample taxonomy was first associated with the viral reads. Among them approximately 30% of the reads were classified only at the kingdom level. The remaining reads showed a taxonomy distribution as follows: *Microviridae* was the most abundant family identified both in the spring (51%) and Milan (52%) while only 7% in the Ponte Lambro sample. Similarly, *Circoviridae* were present in a large amount in the Spring and Ponte Lambro samples (38% and 44%, respectively) while only the 8% of the identified viral reads were recognized in the Milan sample. To a lesser extent *Myoviridae* and *Siphoviridae* were present in a similar abundance in the Ponte Lambro and Milan samples (17 and 12%; 14 and 15%, respectively) while only at a low amount in the spring sample (4% and 2%), Figure 7.

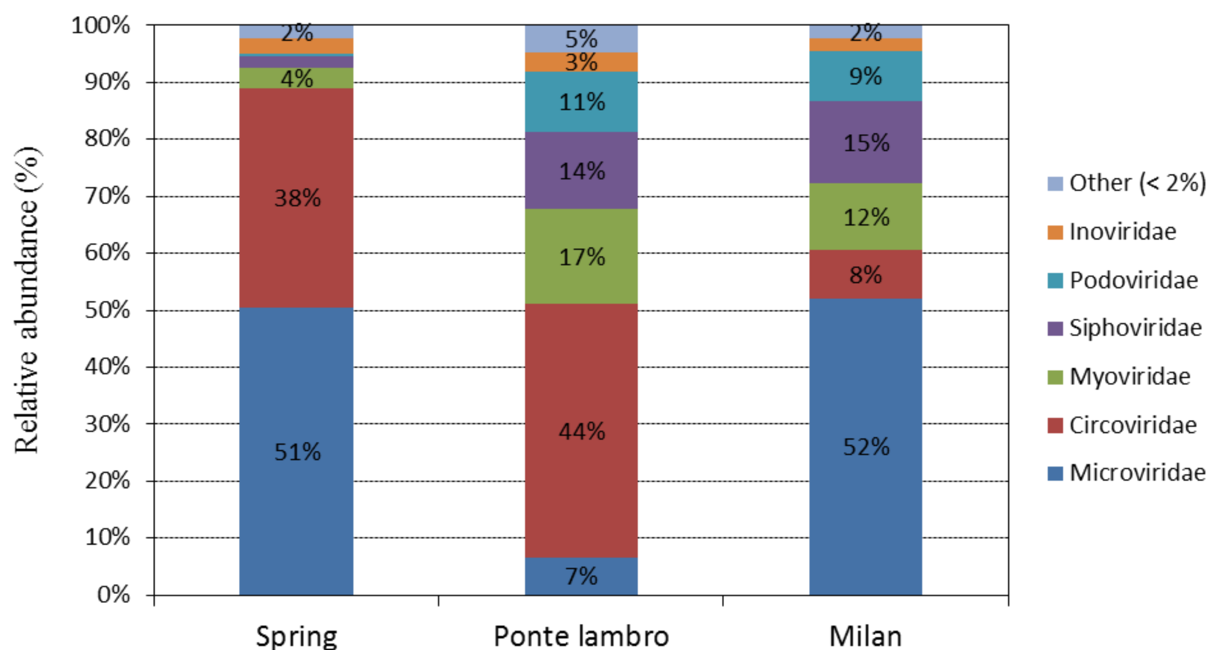


Figure 7: *Taxonomic affiliation at family level of the reads from metagenomic shotgun sequencing of the three viral samples, assigned by BLAST against our customized viral NCBI taxonomy database.*

A separate taxonomic identification of the microbial reads in the three viromes has been performed (Fig. 6B). As observed by other authors (Fancello *et al.,* 2013; Roux *et al.,* 2012; Rosario *et al.,* 2009; López-Bueno *et al.,* 2009; Modi *et al.,* 2013) bacterial reads in viromes are abundant and accounted for up to 90% of the microbial reads. We therefore focused on their taxonomic identification. *Flavobacteraceae* was the only family that stood out of the three samples (37% in the spring sample with respect to 6% and 12% in Ponte Lambro and Milan). Other notable families identified were *Chitinophagaceae* (18% in Ponte Lambro), *Neisseriaceae* (17% in the spring) and *Lactobacillaceae* (12% in the spring). Interestingly, the bacterial families identified in viromes showed a richness dependent on the sampling area. Milan-virome showed the highest richness with 294 different bacterial families represented, followed by Ponte Lambro and the spring showing 290 and 236 bacterial families respectively.

4.2.4 Functional characterization of microbiome data

Functional classification was assigned to each reads identified using the COG database. No changes were observed for each category among the different samples. Metabolism related reads were the most represented (37-42%), reads associated with "cellular processes and signaling" were the second most observed (25-31%), and reads related to "information storage and processes" were the third most observed (15-19%) (Fig. 8A). Reads included in the "mobilome" category accounted for 2-3% of the total: among these, up to 65% were identified as transposase-related reads both in the spring and Ponte Lambro samples (only 37% in Milan), while from 12% to 25% were classified as phage-related reads. Among the metabolism-related category, none was predominant with respect to the others: amino acid transport and metabolism (19-21%), carbohydrate transport and metabolism (14-18%), inorganic ions transport and metabolism (14-18%), lipid transport and metabolism (12-13%) and energy production (12-17%) (Fig. 8B).

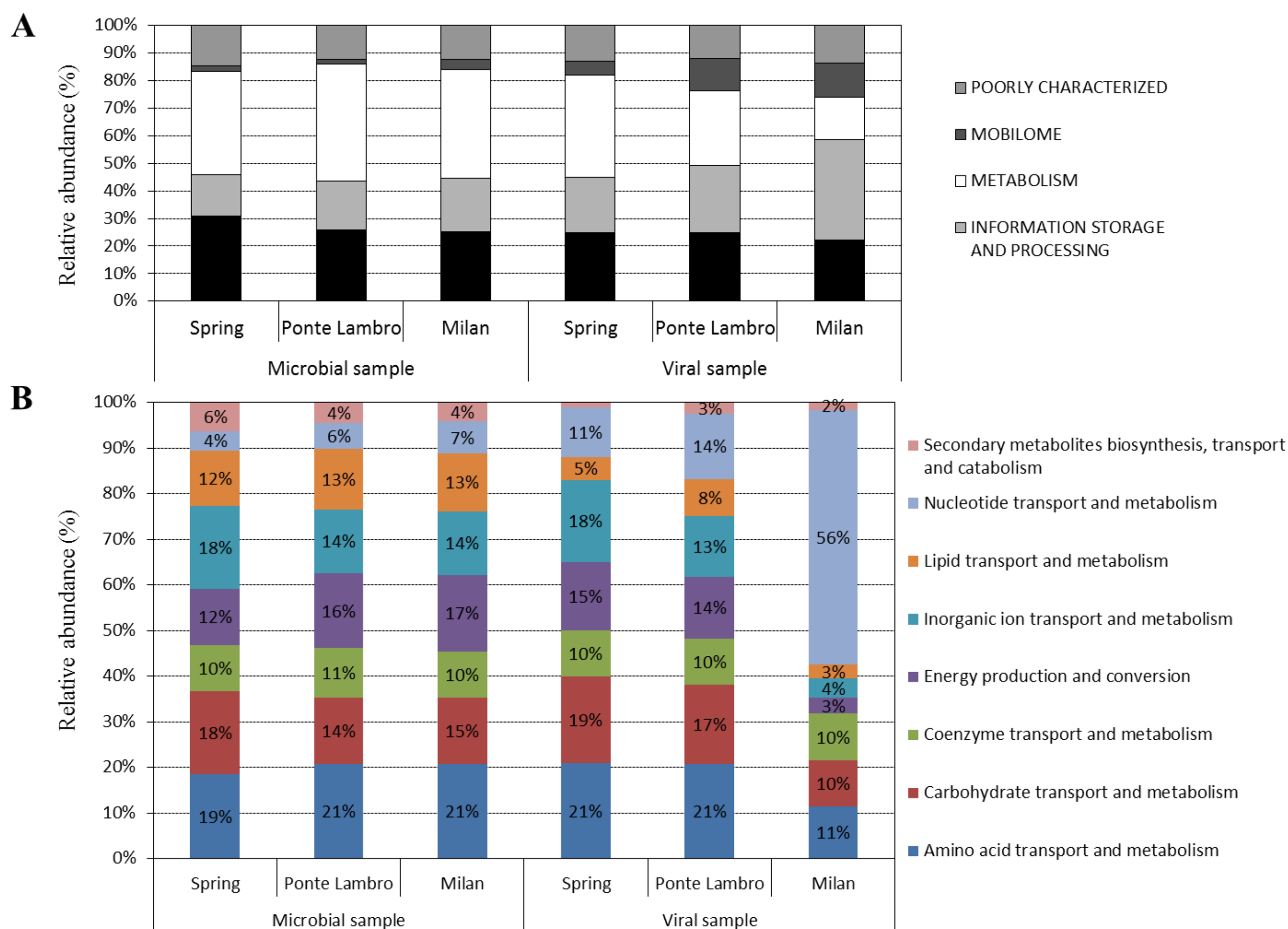


Figure 8: *Functional classification of the reads from metagenomic shotgun sequencing of the microbial and viral samples. Data were first assigned to general classes with COG database (A) then attention was focused on metabolism related classes (B). Data were expressed as relative abundance with respect to the total reads (A) and with respect to only reads belonging to metabolism related class (B)*

4.2.5 Functional characterization of viromes data

Conversely to what was observed for the microbiome data, the functional classification of viromes reads (Fig. 8A) revealed variability between the three Lambro River samples. Interestingly, metabolism-related reads decreased their abundance from the spring to Milan (37%-27%-15%), whereas information storage and processing related reads increased from the spring to Milan (20%-24%-36%). The abundance of reads assigned to cellular processes and signaling was similar in all three samples (25-25-22%). The distribution of metabolism-related reads (Fig. 8B) was similar in the spring and Ponte Lambro samples, no subcategory was overrepresented and the differences for each subcategory in the two samples were lower than 5%. A different scenario was observed in the Milan sample: nucleotide transport and metabolism-related reads accounted for up to 56% (11% -

14% in the other two samples); abundance of reads associated with amino acid and carbohydrate transport and metabolism accounted for 10% (17% and 19% in the other two samples).
Mobilome-associated reads increased their abundance in the Ponte Lambro and Milan samples (12%-12%) compared to the spring (5%). This category, as defined by NCBI-COG, mainly groups reads that encode for phage related proteins as terminase, capisd, tail and other as transposase and protein involved in initiation of plasmid replication. These results were probably due to the increased number of viral reads in the viromes of the Ponte Lambro and Milan samples compared to those detected in the spring sample. According to these observations, it is worth of mentioning that phage-related ORFs in the Ponte Lambro and Milan samples accounted for 32% and 59% of the mobilome, respectively. However, the spring sample showed the highest abundance of reads related to transposase and protein involved in plasmid replication (33% and 37% of the mobilome, respectively), while virus-associated ORFs accounted for 20% of the mobilome.

4.2.6 Identification of antibiotic resistance genes in the microbiome and virome

Relative abundance of reads associated with ARGs increased from the spring microbiome (8%) to Ponte Lambro (11%) and Milan (12%) (calculated the as the relative abundance of reads associated with ARGs over the total reads identified with NCBI). Nevertheless, the three microbiomes showed a similar distribution of the five drug classes as shown in Figure 9A (here the relative abundance is calculated for the total amount of reads associated with ARGs): ARGs that encodes efflux pumps were the most abundant (approximately 47%), then cell wall synthesis and disruption ARGs (23%), and finally DNA synthesis inhibitors (17%). Looking specifically at the antibiotic resistance genes identified from the microbial metagenome, an identical situation was observed: up to 45% of the ARGs identified were genes that encode for efflux pumps or gene modulating the efflux systems (Fig. 9B). Aminocumarin and glycopeptide resistance genes were observed in all three samples in amounts ranging from 12% to 16%. Sequences belonged to beta lactam coding genes, and chloramphenicol and tetracycline resistance genes were also found (7% - 2% - 4%). A different and more heterogeneous scenario has been observed in the virome distribution of ARGs. More generally, in viromes we observed a decrease in the relative abundance of reads associated with ARGs from the spring (1.92%) to Ponte Lambro (1.58%) and Milan (0.48%) samples (calculated the as the relative abundance of reads associated with ARGs over the total reads identified with NCBI).
Interestingly, the gene coding DNA synthesis inhibitors and protein synthesis inhibitor classes increased their relative abundance from the spring to Milan samples (Fig. 9A), whereas ARGs

belonging to efflux pumps and genes modulating antibiotic efflux showed an opposite trend, decreasing from a relative abundance of 57% in the spring sample to 24% in the Milan sample (Fig. 9A). Once again Aminocumarin and glycopeptide resistance genes were two of the most represented in the three samples (Fig. 9B). Interestingly tetracycline resistance genes were observed at 12% in Ponte Lambro, well above the relative abundance of the same gene category observed in the other two viromes and in the three microbiomes (Fig. 9B).

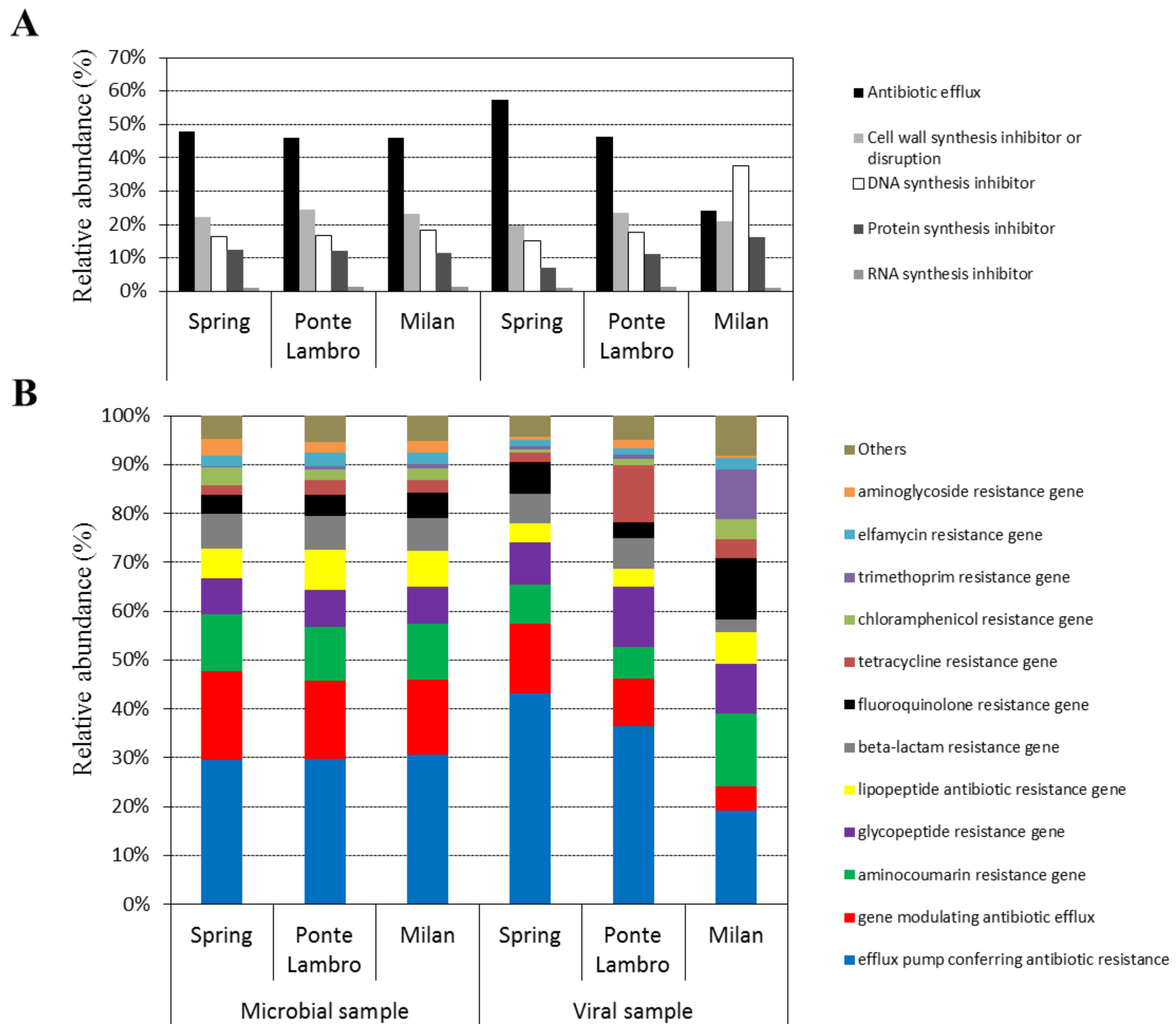


Figure 9: *Comparison of the antibiotic resistance related reads distribution among the microbial and the viral fraction of the three samples. (A) Distribution of the drug classes based on the inhibition target of the antibiotic. (B) Distribution of reads identified as ARGs expressed as relative abundance with respect to the total ARGs-related reads identified.*

4.2.7 Identification of heavy metal and biocide resistance genes

From the spring to Milan, the Lambro River crosses increasingly industrialized areas that negatively affect the point-source pollution levels in its waters, specifically pesticides, metals and organochlorines and organic compounds. For these reasons, microbiomes and viromes were analyzed for the identification of genes related to resistance mechanisms against heavy metals (MRGs). While in the microbiomes the relative abundance of MRGs increased from the spring (9% of the total reads identified with NCBI) to Ponte Lambro (13%) and Milan (14%), an opposite trend was observed in the three viromes from the spring (3%) to Milan (1%). Fifty-six different resistance entries were identified with a similar pattern of relative abundance among the six different metagenomes (Table S4). Among the genes associated with metal resistance the highest abundances identified were against copper (approximately 13% in the three microbiomes while up to 20% in the Ponte Lambro virome), zinc (an average of 13% in the microbiomes and from 8% to 14% in the viromes), iron (8% in the microbiomes; up to 16% in the spring virome) and nickel (7% in the microbiomes; from 5% to 13% in the viromes). Biocide resistance genes were identified in a much lower amount with respect to the MRGs: among them the most abundant identified were resistance genes against hydrogen peroxide (2% in the microbiome and up to 4% in the viromes of the spring and Milan samples).

4.2.8 Identification of gene coding bacteriocin in microbiomes and viromes

The six metagenomes were then analyzed to find sequences that encode for bacteriocines (Table S5). Relative abundance of the reads encoded for bacteriocines observed, ranged from 0.06% (the spring) to 0.12% (Milan). In the three microbiomes the large majority of the reads of interest, are related to Zoocin A (87% to 97%). Other bacteriocines identified at appreciable amount were Pyocin S1 (6% in the spring and 3% in Milan) and Colicin lb (3% in the spring). Another 17 different bacteriocines were identified in at least one of the three samples even if at very low levels (> 1%). No appreciable presence of reads encoded for bacteriocines was observed in the three viromes: the abundances of these reads accounted for less than 0.04% of the total.

4.2.9 Identification of antibiotic resistance genes co-resident with phage-genes in the same sequence contig

Due to the high number of bacterial genes in viromes, only few ARGs co-resident with phage genes in the same contig have been identified. Virome-contigs harboring phage genes and ARGs are listed in Table S6. Several ARGs located in the same contig together with Enterobacteria-phage-genes, have been identified as fluoroquinolone resistance genes. Other fluoroquinolone resistance genes have been identified in contigs together with Cyanobacteria-phage genes. Moreover, with the exception of an aminocoumarin reistance gene co-resident with a *Bacillus* sp. phage gene, the remaining the ARGs have been found co-resident with eukaryotes-viral genes. The presence of bacterial genes in viruses associated to eukaryotes confirmed lateral gene transfer events occurring between bacteria and eukaryotes in aquatic environments as previously documented between bacteria and rumen ciliates (Ricard *et al.,* 2006).

4.2.10 Frequency of antibiotic resistant colonies in water samples

We quantified the total cultivable microbial abundance by plating on two different media: plate PCA and MC which is selective for Gram-negative bacteria. The total bacterial count obtained in the two media revealed an increase of cultivable bacteria from the spring to the water samples collected close to Milan city. On PCA, the bacterial count ranged between 7 x 10^3 CFU/L (3.8 log10) ± 4 x 10^2 in the spring sample and 1 x 10^6 Log_{10}CFU/L (6.1 log10) ± 7 x 10^4 in the Milan sample (Fig. 10A).

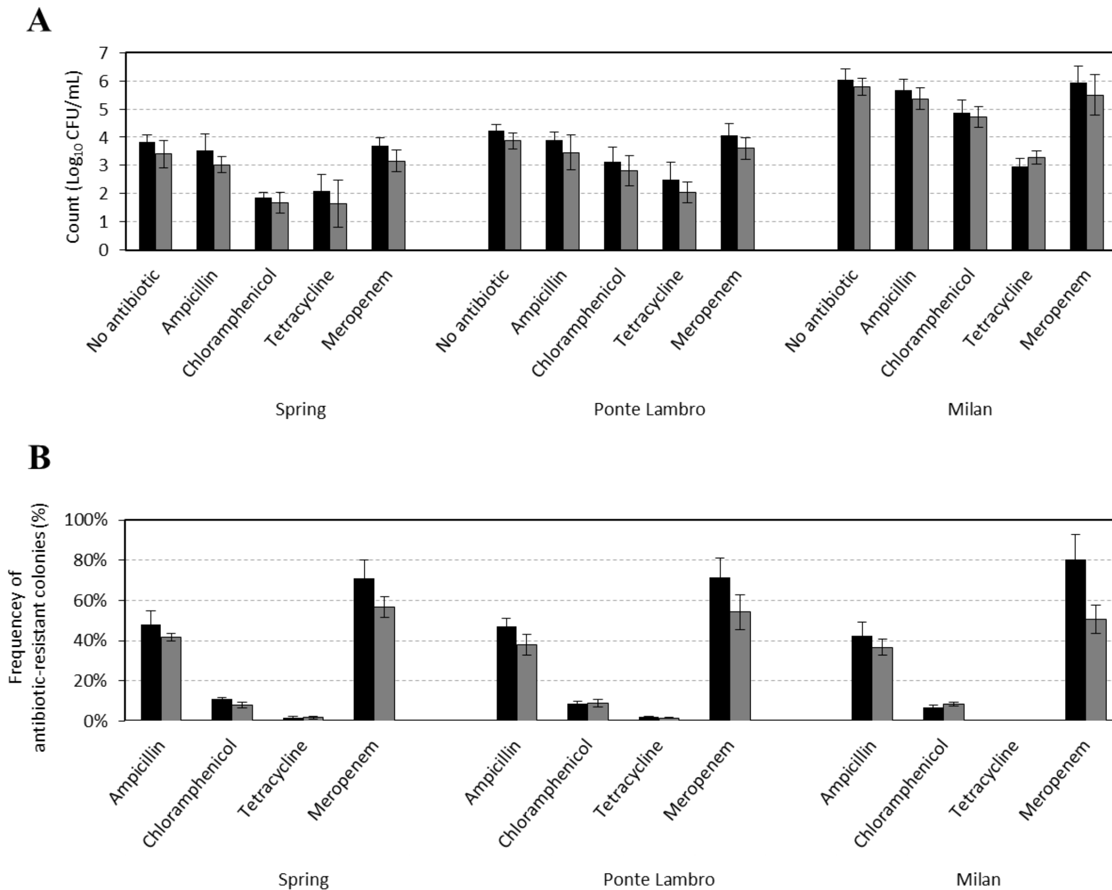


Figure 10: *Microbial count in water samples, expressed as* Log_{10} *(CFU/mL). Microbial count was carried out in PCA media (black bars) and Mac Conkey media (grey bars) after incubation at 20 °C for 48 h in absence of antibiotics (A), and with the addition of addition of antibiotics as reported in the figure at a final concentration of 20 µg/mL. Frequency of the antibiotic resistance colonies (B) was calculated as the ratio of the colony growth in presence of a specific antibiotic over the amount of colonies growth in absence of it. Number expressed as percentage ± the standard deviation.*

Similarly, on MC agar the Gram-negative bacterial count ranged between 2 x 10^3 CFU/L (3.4 log10) ± 4 x 10^2 in the spring sample and 6 x 10^5 CFU/L (5.8 log10) ± 3 x 10^4 in the Milan sample. Focusing on the frequencies of the antibiotic resistant colonies, surprisingly no significant differences were observed among the different samples. In PCA media the highest frequencies were observed in presence of beta lactamic antibiotics: meropenem-resistant colonies reached up to 70% of the total, and ampicillin-resistant colonies were approximately 47%. However, chloramphenicol-resistant colonies were from 7 to 11%, and tetracycline-resistant colonies were under 2% in all three samples (Fig. 10B). We therefore conclude that the frequency of antibiotic resistant colonies was not affected by the increasingly industrialized and urbanized areas crossed by the River from the spring to Milan city. However, the absolute number of antibiotic-resistant colonies increased significantly from the spring to the Milan area as a consequence of the increase in the total bacterial count in the water samples.

4.3 Discussion

The role of viruses in controlling and shaping microbial communities is getting more and more evident (Danovaro *et al.,* 2016; Mody *et al.,* 2013; Muniesa *et al.,* 2013a; Muniesa *et al.,* 2013b). In most of these studies viromes were characterized in terms of abundance of virus-like particles, and in terms of relative abundance of the taxonomical classes of viruses, without taking into consideration the taxonomy and the functional classes of microbial genes harbored by viral genomes. Some authors hypothesize that phageome may represent a community-based mechanism for protecting microbial community, preserving its functional robustness during antibiotic stress (Mody *et al.,* 2013). Mody *et al.* (2013) explored the murine gut phageome as a potential genetic reservoir for bacterial adaptation following antibiotic treatment. The authors observed that antibiotic treatment leads to the enrichment of phage-encoded genes that confer resistance via disparate mechanisms to the administered drug. Based on all these considerations, we decide to explore the quality and abundance of ARGs in viromes and microbiomes isolated from river samples in order to set up a protocol suitable to describe the phage-bacterial ecological network. Lambro River was chosen as a model because it is known as a high environmental risk area for more than twenty years due to contamination with heavy metals and persistent organic pollutants. In all water samples collected we were able to isolate and concentrate the microbiota and the viral populations adopting a TFF-based procedure. A CsCl gradient centrifugation followed by a DNAse treatment of concentrated virus-like particles limited under the 0.2% threshold the contamination (qPCR evaluated) of bacterial DNA in viral DNA extraction. This extremely lower level of contamination was reflected, after viral genome amplification and shot-gun sequencing, in values lower than 0.2 ‰ of abundance of reads associated to ribosomal genes on the total of reads in two of the three viromes analyzed. These value of contamination is comparable/lower than those measured in viromes already available in public database (Table S7) (López-Bueno *et al.,* 2009; Rosario *et al.,* 2009; Fancello *et al.,* 2012; Zablocki *et al.,* 2014).

As reported by other authors, viromes harbored a huge amount of phage-encoded bacterial sequences whose relative abundance was sample-dependent (Fig. 6A). The functional classification analysis carried out on viromes and microbiomes highlights interesting differences in the relative abundance of the different gene classes. Whereas the three microbiomes showed a very similar distribution for each gene functional class, the three viromes showed remarkable differences. The reads associated with genes coding for functions related to the microbial metabolism drastically decreased from the spring (37%) to the Milan (15%) sample, which was enriched with reads associated with genes coding for nucleotide transport and metabolism category (56% compared to

11% and 14% in the spring and Ponte Lambro samples respectively). The ecological reasons for these differences are not known, but the bacterial genes harbored in the viromes surely reflect the gene content harbored in the genomes of the bacterial taxa subjected to viral lytic cycles.

In contrast to what was observed for metabolic genes, the three microbiomes show an increased abundance of ARG-associated reads (abundance of ARGS on the total amounts of reads) (Fig.11), while from the spring to Milan sample the relative abundance of each ARGs class (calculated over the total amount of ARG reads) was kept constant (Fig. 9A), showing no changes among the three samples. A completely different scenario was represented by the ARGs analysis in the three viromes. Here the relative abundance of reads associated with ARGs decreases from the spring to Milan samples (Fig. 11), and the relative abundance of ARGs classes showed a consistent reduction of genes associated with the antibiotic efflux systems, whereas genes associated with protein synthesis and DNA synthesis inhibitors increased consistently (Fig. 9A). Specifically, we observed an increase of reads associated with: i) aminocumarin resistance (from 8 to 15%), ii) fluoroquinolone resistance (from 6 to 13%), and iii) trimethoprim resistance (from 1 to 10%). Concerning MRG frequencies, we observed again an opposite trend depending on the genetic background analyzed, microbiomes or viromes (Fig. 11).

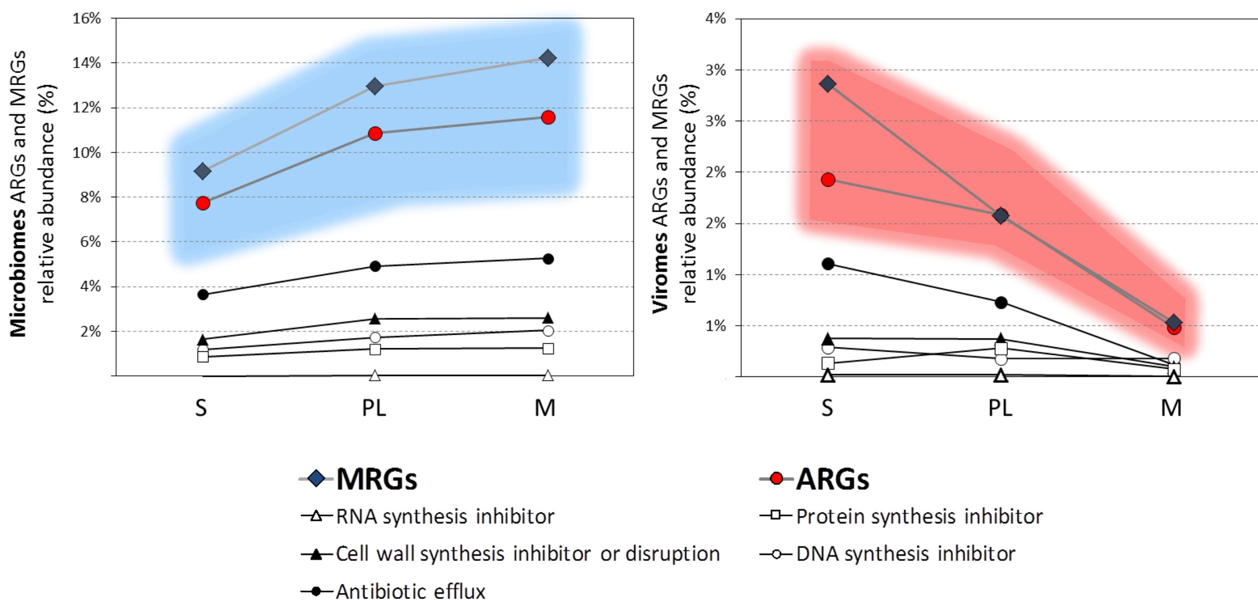


Figure 11: *Frequency of ARGs and MRGs on the total amount of reads in microbiomes (highlighted in the blue shaded area) and viromes (highlighted in the red shaded area). The frequency of each ARGs classes is also shown.*

An increased abundance of MRG-associated reads (abundance of MRGs on the total amount of reads) was observed from the spring to Milan microbiome samples, whereas the opposite occurred in the viromes. Interestingly, both ARGs and MRGs frequencies increased from the spring to Milan

microbiomes coinciding with the increased industrialization/urbanization level of the geographical area where samples were collected. This association is in agreement with a previously formulated hypothesis that considers the exposure of environmental landscapes to pollutants (antibiotics, metals, disinfectants) a risk for the selection and appearance of multidrug resistant strains (Berendonk *et al.,* 2015). Moreover, a link between heavy metals and the maintenance/acquisition of antibiotic resistances has been documented (Knapp *et al.,* 2011; Chen *et al.,* 2015; Teixeira *et al.,* 2016). This correlation is caused by cross- and co-resistance phenomena. Cross-resistance occurs when the same mechanism reduces the susceptibility to antibiotics and metals simultaneously, whereas co-resistance happens when different resistance mechanisms (ARGs and MRGs) are co-localized in the same genetic locus and therefore they are co-transferred in case of horizontal gene transfer mechanisms (Baker-Austin *et al.,* 2006). The contamination of the Lambro River with heavy metals is well documented (Camusso *et al.,* 1995; Galli *et al.,* 1998; Barghigiani *et al.,* 2001; Viganò *et al.,* 2008; Farkas *et al.,* 2007) and the toxic effects of these pollutants are also well known (Tchounwou *et al.,* 2012). In biological systems, heavy metals have been reported to affect cellular structures, several physiological activities (Wang and Shi 2001), and metal ions have been found to cause DNA damage. Therefore, the presence of heavy metals in water environment represents a selective pressure that could force the co-selection of MRGs and ARGs in the microbiome in the absence of antibiotic molecules or in the presence of low concentrations of antibiotic molecules. In this context, it is worth mentioning the increased relative abundance of mobilome-associated reads in the viromes of the Milan sample, thus increasing the probability of transduction events in that environment. However, the observed increase of ARGs and MRGs in microbiome from the spring to Milan area, accompanied by the opposite trend in virome (Fig. 11), need to be validated through the increase of the number of replicates.

In a previous study carried out on an aquaculture water sample, we observed that microbial ORFs identified in the virome did not reflect the microbiome taxonomy, as determined by *16S rRNA* gene profiling or shotgun metagenomic analysis, suggesting that microbial genes mobilized in the genome of viruses could be considered to be remnants of past recombination events rather than a picture of the current microbial diversity (Colombo *et al.,* 2016). Likewise, in this study we observed the same phenomena (Fig. 6B) in all three samples analyzed, thereby confirming that viromes represent a picture of past microbial populations. This is in accordance with the “Kill-the-Winner” dynamics model (Thingstad 2000), which postulates a repetitive cycle in which an increase in prey population leads to an increase in the predator population that in turn decreases the prey population, thus causing its own subsequent decline.

In conclusion, we have developed an integrated approach suitable for the exploration of microbiomes and viromes for ARGs and MRGs starting from fresh water river samples. This approach could represent a useful tool for the characterization of phage-bacterial ecological network in water environments, and for the understanding of environmental constrains which are relevant in the mobilization of ARGs.

5. Evaluation of airborne viruses from two different dairy plants with metagenomic approach

5.1 Introduction

In many cheese-making activities microorganisms in the air represent a hidden and dangerous risk factor. The air present in the processing areas can be a source of pathogenic or spoilage microorganism that could contaminate foods causing problems both on the safety and the quality of the foods itself. Among foods dairy products are particularly susceptible to contamination by airborne microorganisms. 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; Kang *et al.,* 1990; Vickers *et al.,* 1986; Salustiano *et al.,* 2003)
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).
Because significant amounts of bacterial cells are necessary in cheese making production, most of these industries have experienced problems with phage contamination. Phage outbreaks represent an industrial concern because they can slow down the fermentation process and negatively impact the production, resulting in extra costs and delays (Emond and Moineau, 2007). In order to solve this serious problem great attention has been focused on the study and the control of phage populations (Verreault *et al.,* 2011).
Many can be the sources for virulent phages: raw milk, can be contaminated at a low concentration (McIntyre *et al.,* 1991). Moreover, some lactococcal phages can withstand pasteurization (Atamer *et al.,* 2009; Madera *et al.,* 2004). Reservoirs of phages include the materials and equipment used in the manufacturing process as well as the fermentation products and by-products. Even if more difficult to detect, aerosolization is another way for the spreading of phages. Up to 10^8 PFU/m^3 of air have been detected in some areas, in a German cheese manufacturing plant (Neve *et al.,* 1995). Aerosol contamination represents a major microbiological challenge, especially when the aerosol sources and contents are not known. Studies of the airborne distribution of phages could help

identify the sources and determine the impact of aerosols on failed milk fermentations. (Verreault *et al.,* 2011).
Steps taken in the dairy industry to reduce the impact of phage infection on starter activity include: rotation of phage-unrelated strains, the use of mixed strain cultures, separation of the bulk starter preparation area from cheese production and whey handling areas, the use of phage inhibitory media for bulk starter preparation, direct inoculation of cheese vats with frozen concentrated starter cultures, sanitization regimes, air filtration as well as many other strategies (Cogan *et al.,* 1993; Mc Grath *et al.,* 2007).
Avoiding the generation of bioaerosols as well as limiting the air microbial count by using spray systems with appropriate disinfectants should help in controlling infections. Additionally, the air used for positive pressure applications must be filtered to remove dust particles, which may bind phages, and the air inlet for the filters should be located as far as possible from the milk silos and whey tanks. Finally, the efficiency of the filters should be checked regularly (Marcò *et al.,* 2012)

In this study two very different cheeses and cheese-making factory have been considered: Gorgonzola and Grana Padano.
Gorgonzola is an Italian blue-veined cheese. It is one of the most popular Italian cheese 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*. LAB work as starter cultures during the beginning of ripening. 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. Gorgonzola is also susceptible to contamination by *Listeria monocytogenes*, a pathogenic bacteria that can cause infections of the central nervous system and gastroenteritis in people who are immunocompromised.
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 24 months. 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).

Samples of air were collected 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. Viruses have been identified through metagenomic shotgun sequencing. In addition samples from different food matrices (whey, curd, rind) have been collected nearby the air collection process. Microbial DNA from these matrices have been extracted and sequenced with *16S rRNA* gene profiling analysis in order to identify the bacteria present.

5.2 Results

5.2.1 16S rRNA gene profiling analysis

Description of the bacterial communities in the different cheese matrices has been performed with the *16S rRNA* gene profiling analysis (Figure 12A). At phylum level, bacterial profile from GP whey strongly differentiate from the two Gorgonzola related profiles; in fact while GP whey is composed almost entirely by *Firmicutes*, Gorgonzola curd shows the presence of only one main phylum (*Proteobacteria*, 97%). The third sample, 50 days Gorgonzola rind, is characterized by the presence of both *Proteobacteria* (70%) and *Firmicutes* (30%) (Figure 12A).

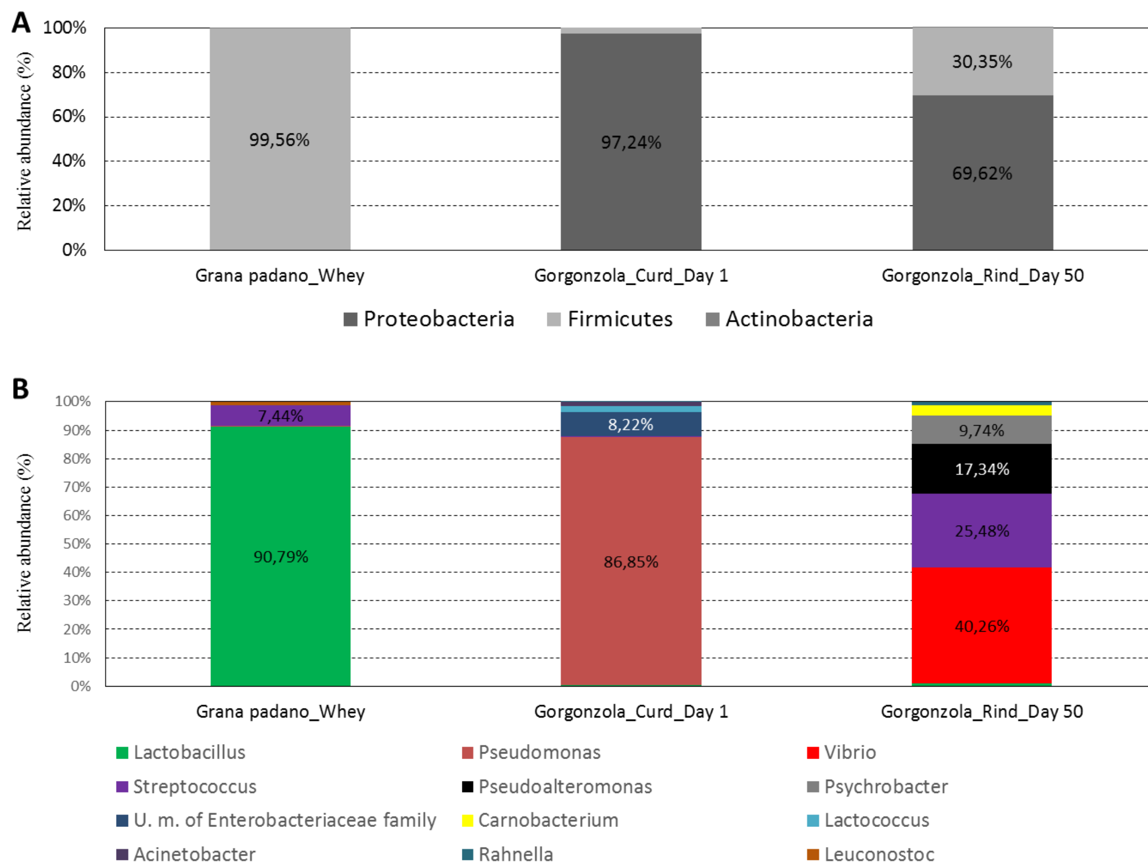


Figure 12: *Taxonomic affiliations based on the 16S rRNA profiling at phylum (A) and genus (B) levels, using the GreenGenes 16S database reference OTUs at 97% identity. Data are expressed as relative abundance.*

Moving to a deeper level, *Firmicutes* identified in GP whey have been recognized as *Lactobacilli* (90%) and *Streptococci* (7%)(Figure 12B). At genus level the profiles of two Gorgonzola samples looks widely different: in fact while the *Proteobacteria* observed in Gorgonzola curd refers mainly to *Pseudomonas* (87%) and *Enterobacteriace* not classified at genus level (8%), a more various profile describes 50 days Gorgonzola rind, where *Vibrio* (40%), *Pseudoalteromonas* (17%) and *Psychrobacter* (10%) are identified. *Streptococcus* (25%) is the main genus related to *Firmicutes* identified%)(Figure 12B). In the last sample have been identified the presence of *Carnobacterium*, a bacterium added during the washing steps as protective culture against *Listeria Monocytogenens*.

5.2.2 Viral shotgun metagenomic – Taxonomy

Viral communities present in the three air samples were sequenced with a shotgun metagenomic approach. BLAST analysis resulted in a larger presence of virus related ORFs in the air surrounding the Gorgonzola curd (13%) with respect to the other two (4% and 5%). As expected, most of ORFs were assigned to bacteria kingdom, from 62% to 91% (Figure 13A).

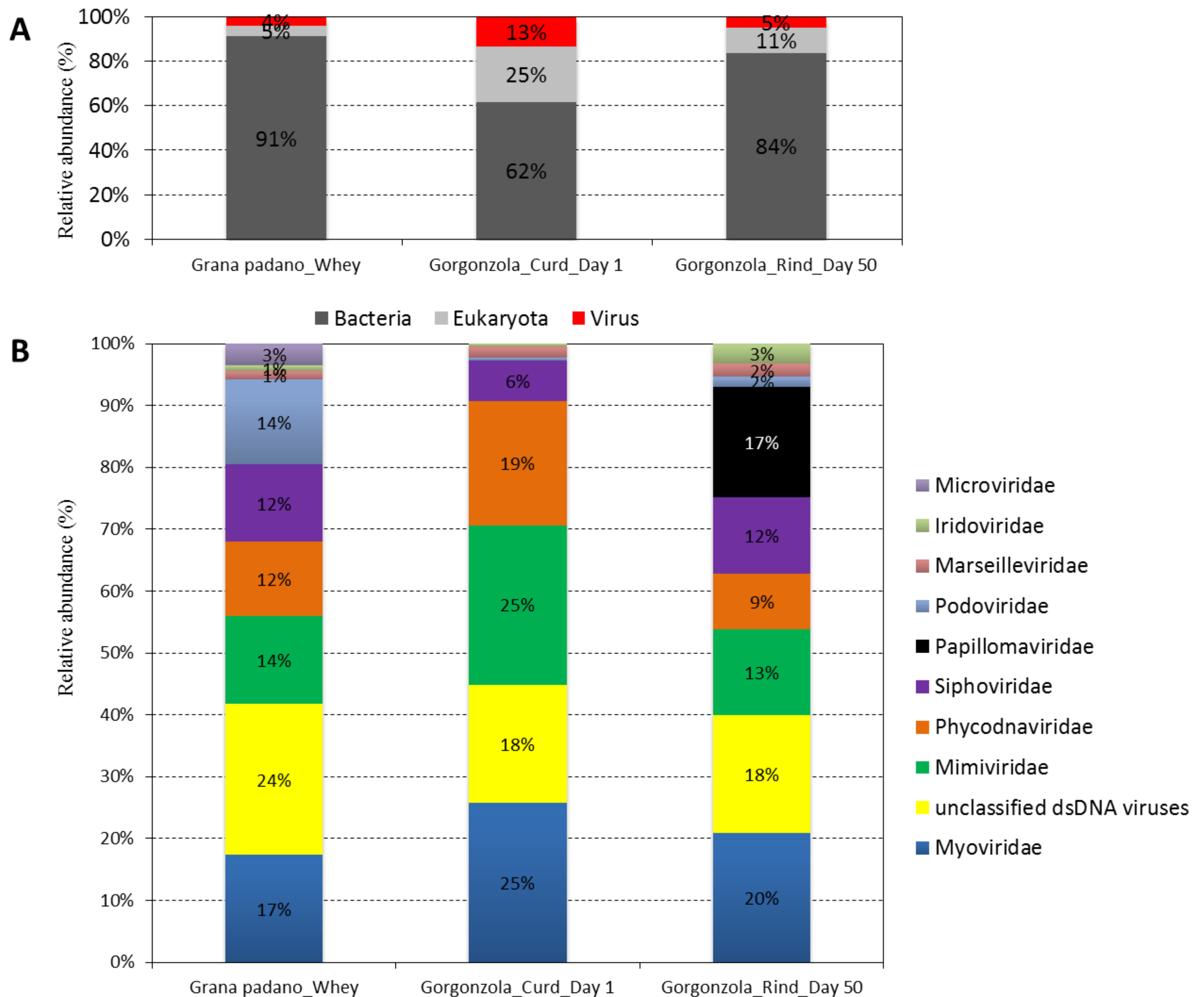


Figure 13: *Taxonomic affiliation at kingdom (A) and family (B) level of the ORFs from metagenomic shotgun sequencing of the three viral samples, assigned by BLAST against NCBI (A) and our customized viral NCBI taxonomy database (B).*

The three samples showed taxonomic profiles very similar among them (Figure 13B). Grana padano related virome was the one showing the highest amount of ORFs unclassifiable at family level and generically describes 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. Gorgonzola curd related virome showed highest abundance with respect to the other two samples for both *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. Gorgonzola 50 days rind related virome showed a similar profile with respect to the two other samples with main difference the presence of *Papillomaviridae* family at high abundance (17%) (Figure 13B).

While the viral taxonomical profiles look similar among the three viromes, the taxonomical profiles of the bacterial ORFs identified in them can be clearly classified based on the dairy plant origin (Figure 14); in fact in the two profiles from the Gorgonzola dairy plant can be observed ORFs with the same bacterial class origins. Even if ORFs from same bacterial class are present in both the samples they differs in terms of abundance; so virome from air surrounding Gorgonzola curd showed higher presence of ORFs from *Sphingobacteria Flavobacteriia* and *Bacilli* families while virome from the 50 days aging room were enriched in ORFs from *Cytophagia*, *Chlamydiia* and unclassified *Alphaproteobacteria* classes.

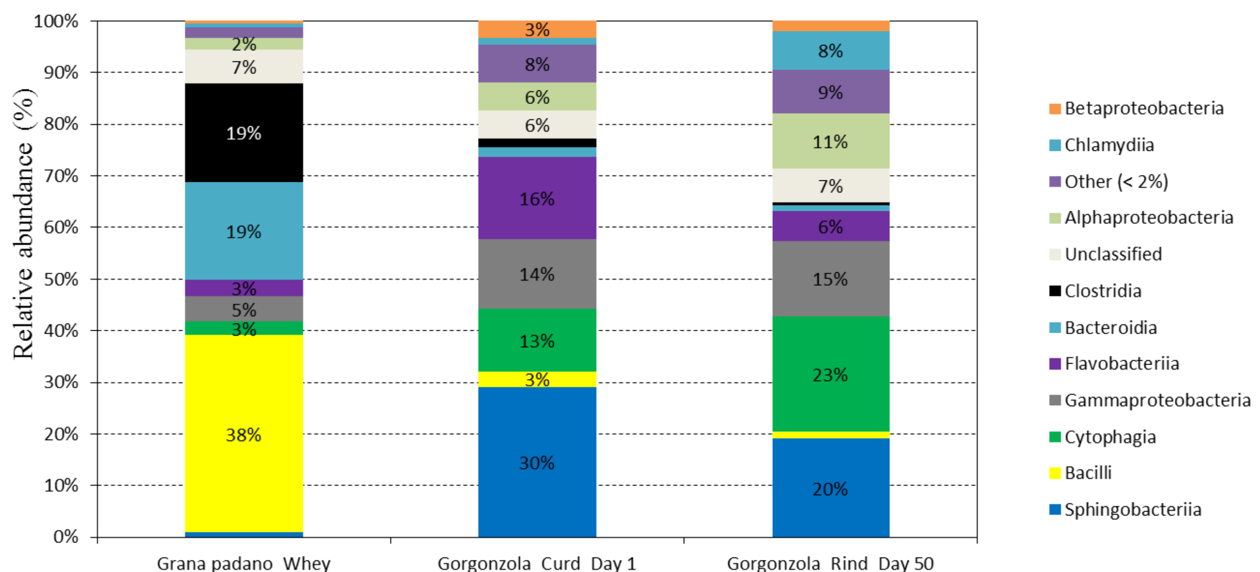


Figure 14: *Taxonomic affiliation of microbial ORFs at family (B) level of the ORFs from metagenomic shotgun sequencing in viral metagenomes, assigned by BLAST with a threshold of* 10^{5} *for the E value against NCBI taxonomy database*

Opposite situation has been observed in air sample from grana padano production plant: ORFs from families highly abundant in the two Gorgonzola related samples, were barely present while other never identified over 3% threshold, were extremely abundant in grana padano as ORFs related to *Bacilli* (38%), *Clostridia* (19%) and *Bacteroidia* (19%) class (Figure 14).
In all the viromes has been possible to identify the presence of bacterial ORFs that share the same taxonomy with bacteria identified in the food matrices with the *16s rRNA* gene profiling. In fact in Grana padano related virome, were observed ORFs from *Lactobacillaceae* family, analogously in Gorgonzola curd related virome were observed ORFs associated to *Streptococcaceae*, *Lactobacillaceae*, *Pseudomonadaceae* and *Chromatiaceae* while in Gorgonzola 50 days rind related virome ORFs associated to *Pseudomonadaceae* and *Vibrionaceae*.

5.2.3 Viral shotgun metagenomic – Functional classification

Functional classification to the bacterial ORFs has been assigned using COG database (Figure 15A). In accordance with the two similar taxonomical profiles seen in the two Gorgonzola related samples, it can be observed that the same samples show similar profiles even in terms of functional classes (Figure 15A). Metabolism related class was the most abundant in both the viromes (up to 39% of the ORFs) followed by class that collect ORFs involved in Cellular process and Signaling (29%). Fewer ORFs were assigned to Information storage and Processing class (19%) while only 2% belonged to the Mobilome one.

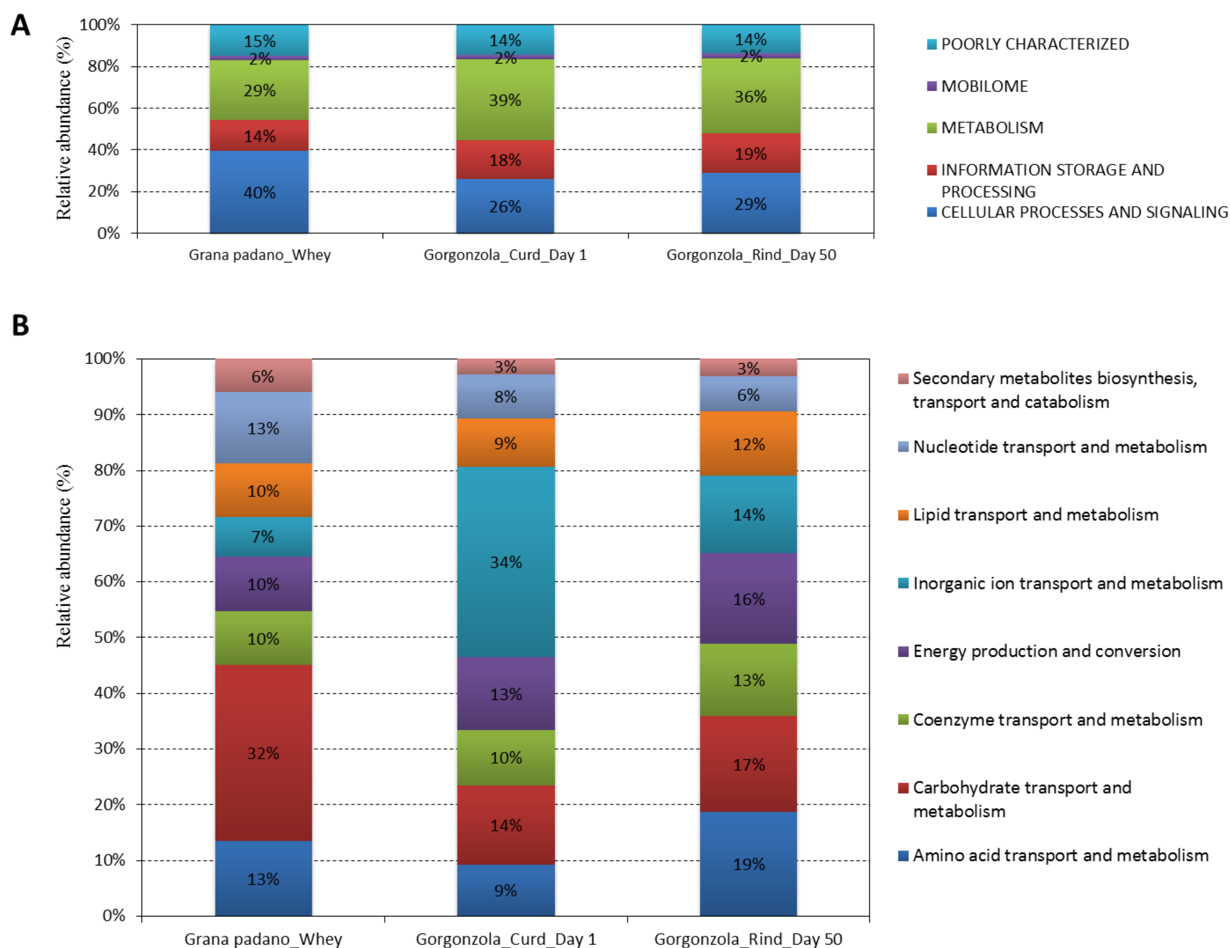


Figure 15: *Functional classification of the ORFs from metagenomic shotgun sequencing of the viral samples. Data were first assigned to general classes with COG database (A) then attention was focused on metabolism related classes (B). Data were expressed as relative abundance with respect to the total ORFs (A) and with respect to only ORFs belonging to metabolism related class (B)*

Taking in consideration metabolism related class, since was observed as the most abundant, a deeper classification has been performed (Figure 15B). It resulted in a higher abundance of ORFs related to Inorganic ion transport and metabolism (+20%) in curd related sample with respect to the 50 days aging room. Opposite situation was observed considering the Amino acid metabolism related ORFs that are more abundant in the aging room (+10%) compared to the other Gorgonzola related sample. All the other categories can be considered equally represented in both the sample (differences under 2%).

Virome related to grana padano whey, was enriched in ORFs associated to Cellular process and signaling category (+10% with respect to the other two profiles) while on the other hand metabolism related ORFs were reduced (-10%); within metabolism category, can be recognized a significant number of carbohydrate related ORFs (32%).

5.3.4 Identification of the antibiotic resistance genes in air viromes

Antibiotic resistance related ORFs were identified through comparison with CARD database. Virome from 50 days aging room showed the highest frequency (2.97%) of ARGs calculated as number of ARGs related ORFs over the number of ORFs identified in NCBI with an aminoacidic similarity over 30%. Lower frequency was observed in curd related virome (1.87%) while only 0.42% was observed in the grana padano related one. Beyond the highest frequency, in the sample from 50 days aging room were also identified ARGs related ORFs against 19 different antibiotics, while against 12 antibiotics in the other Gorgonzola sample and only 10 in the grana padano one.
Once identified ARGs, classification based on the antibiotic drug target was performed (Figure 16A). Virome associated to Gorgonzola curd resulted to be the most balanced sample among the different classes with values ranging from 17% (DNA synthesis inhibitor) to 34% (Antibiotic efflux). In the other two samples the distribution was less balanced among the classes since most of ARGs were collected in the Antibiotic efflux one, 57% and 48% in grana padano and Gorgonzola 50 days related samples. Noteworthy both DNA and Protein synthesis inhibitor showed values ranging from 13% to 26% in both the samples. Looking even more in detail can be seen a meaningful presence of genes against lipopetide, aminocoumarin and chlorampenicol (all at 14%) in Gorgonzola curd related sample, genes against lipopetide, aminocoumarin glycopeptide and beta-lactam (7%-9%)in Gorgonzola rind related sample while only elfamycin and tetracycine resistance genes were identified in grana padano related sample over the 5% threshold excluding genes involved in antibiotic efflux (Figure 16B).

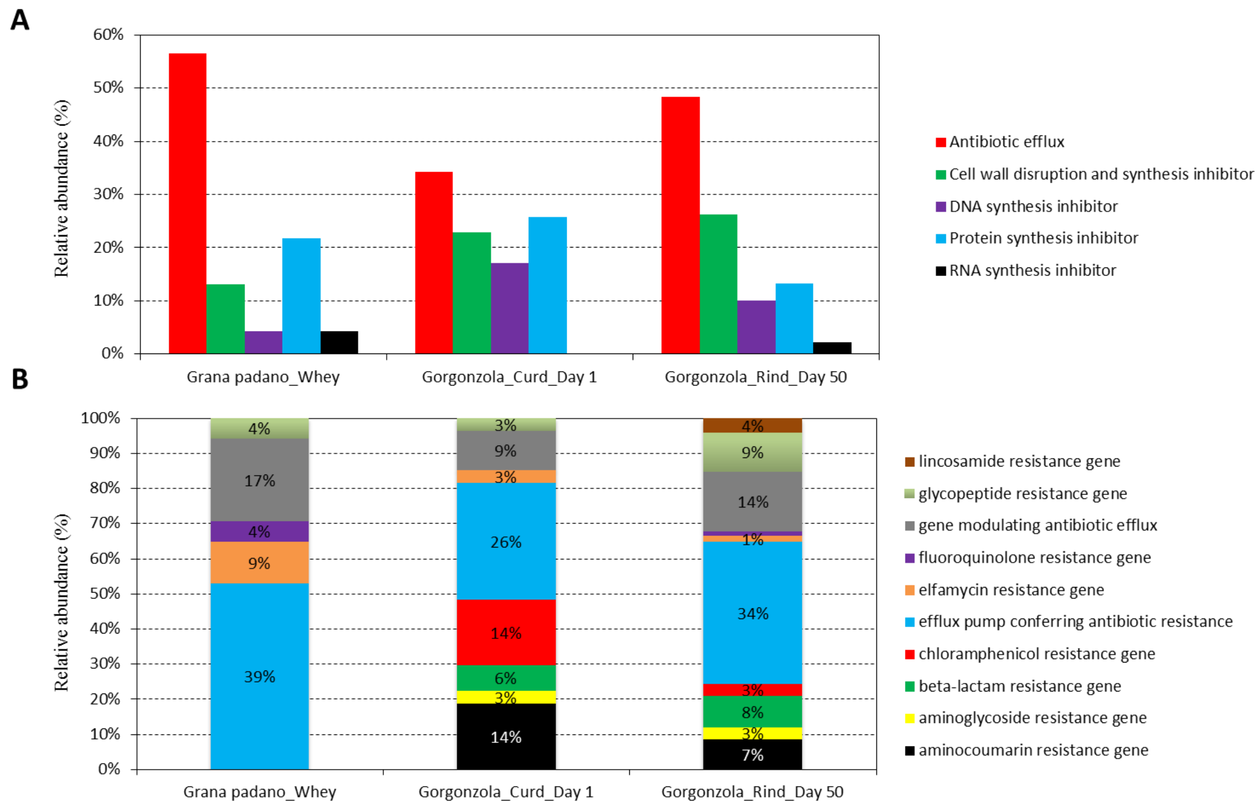


Figure 16: *Comparison of the antibiotic resistance related ORFs distribution among the microbial and the viral fraction of the three samples. (A) Distribution of the drug classes based on the inhibition target of the antibiotic. (B) Distribution of ORFs identified as ARGs expressed as relative abundance with respect to the total ARGs-related ORFs identified*

5.3.5 Identification of Legionella related ORFs

The shotgun metagenomic analysis on the virome let us to identify the presence of ORFs from pathogenic bacteria; in fact in both samples from the Gorgonzola plant present ORFs from different species of Legionella. In particular sample from 50 days aging room was particularly enriched in this ORFs (3.40% of the total ORFs identified with NCBI) that originates from 20 different species: *L. anisa, L. cherrii, L. drancourtii, L. fairfieldensis, L. fallonii, L. geestiana, L. hackeliae, L. lansingensis, L. longbeachae, L. massiliensis, L. moravica, L. norrlandica, L. oakridgensis, L. oakridgensis, L. parisiensis, L. pneumophila, L. sainthelensi, L. shakespearei, L. tunisiensis, L. wadsworthii.* The second Gorgonzola air sample appear clearly less abundant since just the 0.9% of the total ORFs where assigned to Legionella and only 7 different species where identified: *L. drancourtii, L. fallonii, L. hackeliae, L. oakridgensis, L. pneumophila, L. shakespearei, L. tunisiensis.*

Focusing the attention on the 50 days aging room virome, has been observed that the 5% of the ARGs identified were associated to *Legionella* and in particular these ARGs have as target different common antibiotics as chloramphenicol, beta-lactamic, tetracycline and aminoglicosyde and other ORFs have a more wide function so they are generically involved in antibiotic efflux.
Looking at the totality of the ORFs related to Legionella, can be observed that 37% of them was associated to Information Storage And Processing category, while both for Metabolism and Cellular Processes And Signaling categories values were around 25% of the ORFs.

5.3.6 Phage – host interaction

Starting from the data from metagenomic shotgun sequencing, a prediction of the phage-host interaction has been performed: the prediction was based on the phages identified in the three samples. The prediction shows that the *Proteobacteria* is the most stricken phylum in all the samples: 12 different genera were identified to be target of phages. In particular, in grana padano sample highest value were observed with *Vibrio, Acinetobacter* and *Pseudomonas* genera (Figure 6).
In both the Gorgonzola samples highest values were associated to the *Firmicutes* phylum and in particular the *Bacillus* genera (up to 18%). Meaningful the values observed for *Streptococcus*, *Cronobacter* and *Cellulophaga* in the air associate to Gorgonzola curd while *Leptospira* in the air associate to Gorgonzola 50 days rind (Figure 17).

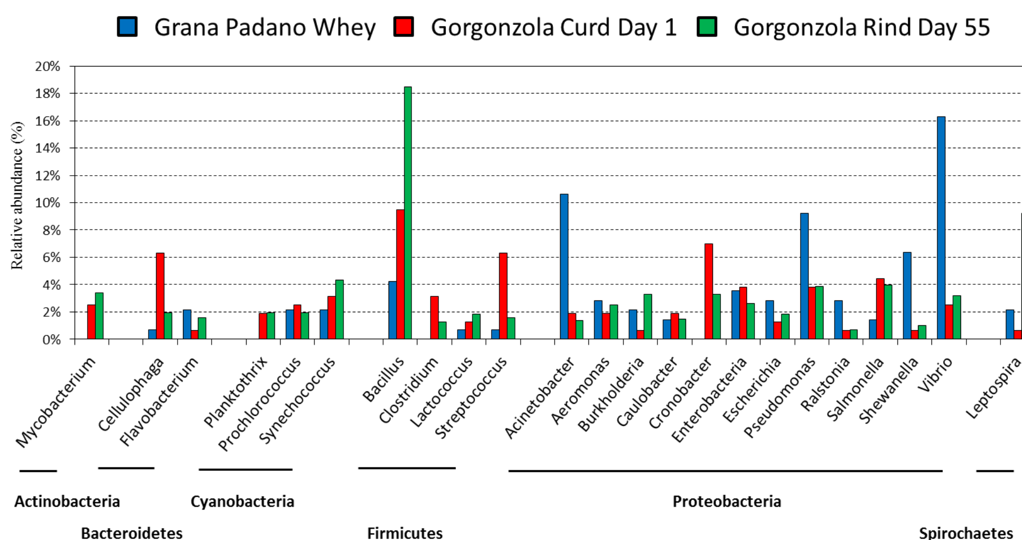


Figure 17: *Distribution of phage-hosts association. Percentages were respect to the total amount of phage-related ORFs identified. Bacteria were grouped based on taxonomic level of phyla.*

5.3 Discussion

The study of food contamination used to focus the attention on the possible risk coming from row materials, food handling and contaminate tools. On the other hand, the potential of the air in the contamination problems is often underestimated. Air can easily move microorganisms, spores and dust to reach food in preparation and limited technologies are available for an efficient sanitization of it. Identification of microorganisms present in the air cannot be limited to outdated techniques as dilution and plating but must take place with the newest technologies such as the next generation sequencing.

Cheese making plants are an optimal case of study since contaminants present in the air could dramatically affect the fermentation steps leading to food safety and economical problems.

Metagenomic shotgun sequencing put in evidence that both in different cheese making plants and in different areas of the same plant were present the same families of viruses even if they showed slight differences in their relative abundance. Even if they are so comparable across the three samples, looking at the bacterial genes harbored in the different viromes is it possible to clearly differentiate the samples based on the cheese making plant: the two profile from the gorgonzola production plant result almost identical to each others while totally different from the GP production plant. Interestingly the two samples from Gorgonzola production plant result incredibly similar even if they belong to production areas that totally differ in terms of temperature, humidity, light, presence of workman and so on. This situation could be consequence of the movement of viruses from an area to another or consequence of the presence of bacteria from analogous families in both the working areas. Despite the clusterization based on the cheese-making plant, evidences of the variability of the virome in different areas of the same plant are observed looking at the presence of *Legionella* related ORFs. *Legionella* are Gram-negative bacteria that cause Legionnaires' disease (LD) and Pontiac fever in humans. *Legionella* infections are regularly traced in systems characterized by dark, high humidity and low temperature such as room humidifiers, cooling towers and air conditioning systems. These systems often exhibit favorable growth conditions for *Legionella* and transmission to humans occurs via contaminated water aerosolization (Van Heijnsbergen *et al.,* 2015). The same characteristic are common to the aging room where Gorgonzola was kept for 50 days. Presence of *Legionella* related ORFs in the air virome in not enough to affirm that there was a *Legionella* contamination during the sampling time but is an evidence of a past presence of bacteria from *Legionellaceae* family justified by phage lytic cycle events.

The parallel analysis of the air viromes and the bacteria present in the food matrices allowed to identified the presence of other possible event of genetic exchange. The continuous presence of the same microorganisms at meaningful concentration and direct exposed to the air is a favorable condition for the transfer of genetic material from cheese related bacteria and phages in the air. In the light of these considerations we should not be surprised to have identified in the three viromes, the presence of ORFs associated to bacteria characteristic of the respective food matrix. These evidences underline that there have been a direct interaction between bacteria and their respective phages, justified by the presence of genes transferred from the food associated bacteria to the phages but, consequently, it is also possible that other genes followed the opposite route moving from phages and integrating in food related bacteria with an hypothetic increase of the risk for the human health.

6. Products

Disclosing the phage-mediated antibiotic resistances in the food chain (diPHARE-FOOD)

Stefano Colombo (stefano.colombo@unimi.it)
Dept. of Food, Environmental and Nutritional Sciences, University of Milan, Italy
Tutor: Prof. Diego Mora

State of art

Antibiotic resistance (AR) may be obtained by spontaneous mutations or acquired by the incorporation of antibiotic resistance genes (ARGs). ARGs spread between cells by using genetic platforms known as mobile genetic elements: plasmids, transposons, integrons and bacteriophages. Bacteriophages have been confirmed as the most abundant beings in many natural environments, and their concentrations use to overnumber those of bacteria by a factor most frequently ranging from 1 to 10. Due to their structural characteristics in extracellular phase, bacteriophages persist quite successfully in the environment and are quite resistant to natural or man generated stressors [1]. Recently the idea that bacteriophage transduction plays a role in the HGT from environmental to human and animal body associated biomes is gaining momentum. It is clear the importance of the characterisation of the phage communities in water ecosystems linked to the food chain, such as fresh fruits, vegetables and fish, which could be contaminated with antimicrobial and ARGs in many ways (soil, environment, irrigation, sewage discharges, food handling) [2]. Fruits and vegetables consumed raw are increasingly being recognized as important vehicles for transmission of human pathogen. Moreover, in the aquaculture farming therapeutic and prophylactic usages of antimicrobials is a very common practice. The aquatic ecosystem clearly appear as a key environment that may provide an ideal setting for the acquisition and dissemination of AR, because they are frequently impacted by anthropogenic activities. A diverse mixture of antibiotics and other pollutants, their metabolites and resistant bacteria, reaches the aquatic environment through treated and untreated sewage, hospital waste, aquacultures discharges, and agricultural runoff. These aquatic compartments may therefore have a significant role in driving ARGs transfer, ecology, and evolution [3][4].

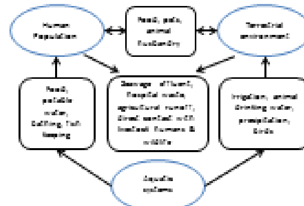


Figure 1: Routes by which antimicrobial resistant bacteria and resistance genes can cycle through human population and terrestrial and aquatic systems [4].

Aim of the project

This PhD research project is aimed to characterise the microbial and viral communities of environmental samples and to assess the impact of their phage populations in the spreading of antibiotic resistance genes (ARGs) through the food-chain. Water samples will be treated with a tangential flow filtration (TFF) system to separate viral particles and microbial cells. Then phageome will be investigated through next generation sequence approach for the evaluation of phage-carried ARGs, allowing to establish the link between the AR associated to the bacterial communities and the phageome.

Objectives and mile stones

The project activities will be organized in four main work packages (WPs):

- **WP1 - Setup and protocols optimization**
  - A1) separation of the microbial and viral communities (Tangential Flow Filtration system).
  - A2) purification of viral particles by CsCl density ultracentrifugation.
  - A3) viral DNA extraction.
  - A4) viral DNA amplification.
- **WP2 - Microbial communities characterization**
  - A5) taxonomic composition of the microbial communities by 16S rRNA profiling.
  - A6) quantification of the microbial communities abundance by flow cytometry.
  - A7) frequency of antibiotics resistant colonies by standard plating procedure.
- **WP3 - Phage communities characterization**
  - A8) taxonomic composition of phage communities through a metagenomic approach.
  - A9) ARGs associated to phage communities by comparing the metagenomic data with the Antibiotic-Resistance SPM database and the Antibiotic Resistant gene Data Base (ARBD) [5].
- **WP4 - Data analysis**
  - A10) estimation of the relevance of the phage-mediated antibiotic resistances in the food chain.

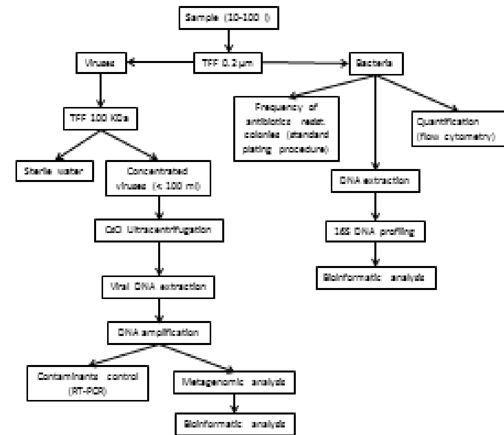


Figure 2: schematic representation of each step of the protocol. This scheme will be applied to each sample to be analyzed.

Gantt diagram

Activity \ Month	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
Setup and protocols optimization (WP1)	■	■	■	■	■	■	■	■																												
Sampling and analysis of water samples (WP2 + WP3)						■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■							
Analysis of microbiome and virome data (WP4)																											■	■	■	■	■	■	■	■	■	■
Preparation of manuscripts and thesis	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Table 1: Planning of the 3 years activities

References

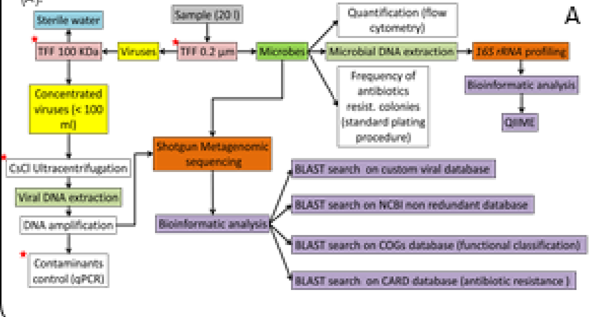
[1] Muniesa, M., Colomer-Lluch, M., & Jofre, J. (2013). Could bacteriophages transfer antibiotic resistance genes from environmental bacteria to human-body associated bacterial populations? *Mobile Genetic Elements* 3:4
[2] Claire Verraes, Sigrid Van Boxstael, Jeroen Dewulf and Lieve Herman (2013). Antimicrobial Resistance in the Food Chain: A Review. *International Journal of Environmental Research and Public Health* ISSN 1660-4601
[3] Elisabet Marti, Eleni Variatza, and Jose Luis Balcazar (2014). The role of aquatic ecosystems as reservoirs of antibiotic resistance. *Trends in Microbiology*. Vol. 22, No. 1
[4] Nick G. H. Taylor, David W. Verner-Jeffreys and Craig Baker-Austin (2011) Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends in Ecology and Evolution*, Vol. 26, No. 6
[5] Liu, B. & Pop, M. (2009). ARDB—Antibiotic Resistance Genes Database. *Nucleic Acids Research*. 37, D443–D447

Virome-associated Antibiotic Resistance Genes In Freshwater Aquaculture

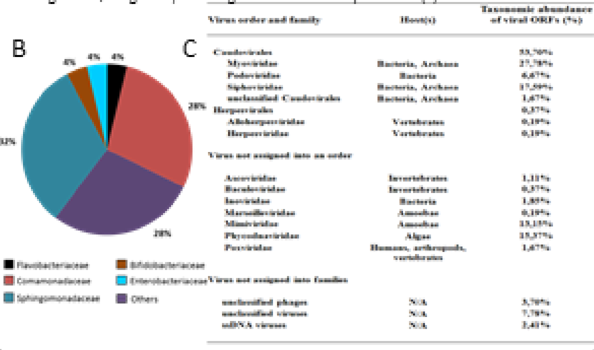
Stefano Colombo (stefano.colombo@unimi.it) Dept. of Food, Environmental and Nutritional Sciences, University of Milan, Italy
Tutor: Prof. Diego Mora

State of Art: Antibiotic resistance (AR) may be obtained by spontaneous mutations or by the acquisition of antibiotic resistance genes (ARGs). Due to their structural characteristics in extracellular phase, bacteriophages persist quite successfully in the environment and are resistant to natural or man generated stressors [1]. The idea that bacteriophage transduction plays a role in the HGT from environmental to human and animal body associated biomes is gaining momentum. The aquatic ecosystem clearly appear as a key environment that may provide an ideal setting for the acquisition and dissemination of AR. According to the newly released data (FAO), world aquaculture production of food fish has exceeded 60 million tonnes in 2011 and aquaculture contributed 40.1% to the world total fish production [2]. This widespread growth of aquaculture has been accompanied by an increased use of a wide range of antibiotics [3] to prevent and treat bacterial infection in fish. This work is aimed to characterize the microbial and viral communities of aquaculture samples and to assess the impact of their viral populations in the spreading of ARGs.

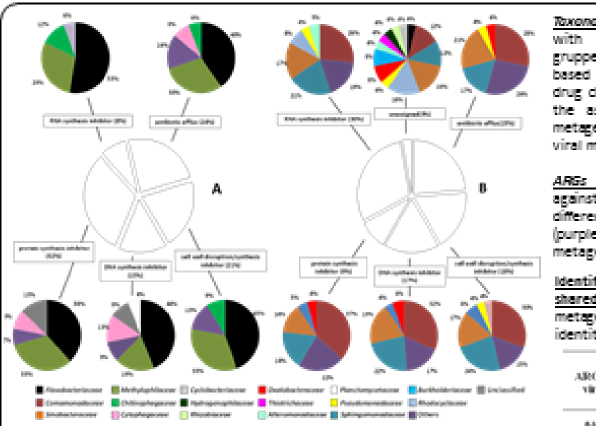
Pipeline: 20 l of breeding tank wastewater were collected at Fondazione Edmund Mach – Technology Transfer Centre (San Michele all'Adige – TN – Italy); The tank was used for the breeding of different salmonids as Carpione del Garda (*Salmo carpio*), Trota marmorata (*Salmo trutta marmoratus*) and Salmerino alpino (*Salvelinus alpinus*). Experiment pipeline shown in (A).



Taxonomy: Taxonomic affiliation of bacteria (family level - based on the results of *16S rRNA* profiling analysis (B)) and based on the predicted open reading frames (ORFs) from viral metagenome, assigned by BLAST against NCBI taxonomy database (C).



Virus order and family	Host(s)	Taxonomic abundance of viral ORFs (%)
Caudovirales		53,70%
Myoviridae	Bacteria, Archaea	27,78%
Podoviridae	Bacteria	6,67%
Siphoviridae	Bacteria, Archaea	17,59%
unclassified Caudovirales	Bacteria, Archaea	1,67%
Herpesvirales		0,37%
Alloherpesviridae	Vertebrates	0,19%
Herpesviridae	Vertebrates	0,19%
Virus not assigned into an order		
Ascoviridae	Invertebrates	1,11%
Baculoviridae	Invertebrates	0,37%
Inoviridae	Bacteria	1,85%
Marseilleviridae	Amoebae	0,19%
Mimiviridae	Amoebae	13,15%
Phycodnaviridae	Algae	15,37%
Poxviridae	Humans, arthropods, vertebrates	1,67%
Virus not assigned into families		
unclassified phages	N/A	3,70%
unclassified viruses	N/A	7,78%
ssDNA viruses	N/A	2,41%



Taxonomy of ARGs: ARGs identified with CARD database have been grupped in four different drug classes based on the inhibition target. Each drug class have been described with the associate taxonomy. Microbial metagenome (A) was compared with viral metagenome (B)

ARGs Identification: BLAST analysis against CARD database identified 19 different ARGs in the microbial (purple bars) and viral (yellow bars) metagenomes (C).

Identification of ARGs genes shared between viral and microbial metagenomes (> 95% aminoacidic identity)(D).

ARGs shared between viral and microbial metagenome	Coding function *	Mobilized ORFs hosted in the same contig
β-lactam resistant gene (*Sphingopyxis alaskensis*)	beta-lactamase (class C family)	β-lactam resistant gene partial (*Sphingomonas* sp.) hypothetical protein (*Sphingomonas* sp.) peptidase (*Sphingopyxis alaskensis*) membrane protein (*Sphingopyxis alaskensis*)
Tetracycline resistant gene (*Caulobacter* sp.)	Transcriptional regulator, subgroup of AraC transcriptional regulators having an N-terminal Type I glutamine amidotransferase (GATase1)-like domain	integrase (*Novosphingobium* sp.) sarcosine oxidase subunit beta (*Novosphingobium* sp.)
Gene modulating antibiotic efflux (Oxalobacteraceae bacterium IMCC9480)	Transcriptional regulator (repressor function), TetR family	cyclohexanone monooxygenase (*Brevundimonas subvibrioides*) acyl-CoA ligase (*Brevundimonas subvibrioides*) TonB-dependent receptor (*Brevundimonas subvibrioides*) L-carnitine dehydratase/bile acid-inducible protein F (*Caenispirillum salinarum*) alpha/beta hydrolase (*Deinococcus phoenicis*) putative flavin-containing monooxygenase (*Rhodococcus wratislaviensis*)
Glycopeptide resistance gene (*Sphingopyxis baekryungensis*)	AAA+ (ATPases Associated with a wide variety of cellular Activities)	GTPase HflX (*Sphingopyxis baekryungensis*) RNA-binding protein Hfq (*Sphingopyxis baekryungensis*) ATPase AAA (*Sphingopyxis baekryungensis*)
Glycopeptide resistance gene (*Sphingobium baderi*)	HATPase_C (Histidine kinase-like ATPases)	
Gene modulating antibiotic efflux (*Limnohabitans* sp. Rim28)	REC (cheY-homologous signal receiver domain) HATPase_C (Histidine kinase-like ATPases)	/ /

Comparison with publically available viral metagenomes: relative abundance of ARGs, with respect to the total amount of ORFs identified in 17 freshwater viromes (E).

Freshwater sample	ARGs	% ARGs	Freshwater sample	ARGs	% ARGs
Elbarbera	638	0,9036%	RW_nursery_DNA	80	0,0311%
Aquaculture (present study)	214	0,8463%	Lake Bourget	175	0,0223%
Antarctic lake summer	136	0,4723%	RW_effluent_DNA	49	0,0205%
FTR_0707	441	0,4309%	Tilapia_11	6	0,0152%
FTR_01_08	267	0,3642%	Antarctic lake spring	4	0,0103%
Ilij	186	0,3189%	Tilapia0406	4	0,0090%
Hamdoun	118	0,2893%	Tilapia_08	3	0,0085%
FTR_0807	234	0,2641%	Lake Pavin	34	0,0085%
Molomhar	138	0,1831%	RW_park	9	0,0054%

Conclusion: This study addresses for the first time a complete description of microbiome and virome in an aquaculture sample, giving information on the presence of ARGs and their mobilization by transduction mechanisms. Shotgun sequencing of viruses revealed that more than 50% of the predicted ORFs belonged to *Caudovirales* order and in particular to the *Myoviridae* family (28%). *Caudovirales* are bacteriophage classified as Group I dsDNA viruses. The large abundance of *Caudovirales* ORFs can be plausibly explained by the absolute majority of bacterial ORFs compared with those belonging to Eukarya and Archea. One of the aims of this study was the assessment of the entity of ARGs in the aquatic virome: our data indicated that ARGs are present in virome and some of these ARGs appeared to be mobilized from bacteria to phages or *vice versa*. We also show that ARGs could be mobilized even in the absence of selective pressure, *i.e.* in the absence of antibiotic treatment in the aquaculture. Moreover, we showed that ARGs were mobilized together with other bacterial genes coding for more general metabolic functions, thus confirming the presence of a complex phage-bacterial network in the aquaculture environment

References:
[1] Muniesa M, Colomer-Lluch M, Jofre J. 2013. Could bacteriophages transfer antibiotic resistance genes from environmental bacteria to human-body associated bacterial populations? Mobile Genetic Elements 3:4
[2] FAO Fisheries and Aquaculture Department. 2013. Global Aquaculture Production Statistics for the year 2011.
[3] Armstrong SM, Hargrave BT, Haya K. 2005. Antibiotic Use in Finfish Aquaculture: Modes of Action, Environmental Fate, and Microbial Resistance. Hdb Env Chem, 5M, pp 341-357.

XX WORKSHOP
ON THE DEVELOPMENTS IN THE ITALIAN PHD RESEARCH ON FOOD SCIENCE, TECHNOLOGY AND BIOTECHNOLOGY
PERUGIA, SEPTEMBER 23rd-25th, 2015

Virome-associated Antibiotic Resistance Genes In Freshwater Aquaculture

Stefano Colombo[a], Stefania Arioli[a], Simone Guglielmetti[a], Giulia Della Scala[a], Eros Neri[a], Fernando Lunelli[b], Diego Mora[a]*

[a]Department of Food, Nutritional, and Environmental Sciences, University of Milan, Italy; [b]Fondazione Edmund Mach, Dipartimento Filiere Agroalimentari, Unità Acquacoltura e Idrobiologia, San Michele all'Adige, Trento, Italy

State of Art: Antibiotic resistance (AR) may be obtained by spontaneous mutations or acquired by the incorporation of antibiotic resistance genes (ARGs). Due to their structural characteristics in extracellular phase, bacteriophages persist quite successfully in the environment and are resistant to natural or man generated stresses [1]. The idea that bacteriophage transduction plays a role in the HGT from environmental to human and animal body associated biomes is gaining momentum. The aquatic ecosystem clearly appear as a key environment that may provide an ideal setting for the acquisition and dissemination of AR. According to the newly released data (FAO - Fisheries and Aquaculture Department), world aquaculture production of food fish has exceeded 60 million tonnes in 2011, an increase of 6.2% over the previous year. The value of farmed food fish has been calculated in USD 130 billion. Aquaculture contributed 40.1% to the world total fish production, and almost all the seaweeds production [2]. This widespread growth of aquaculture has been accompanied by an increased use of a wide range of antibiotics [3] to prevent and treat bacterial infection in fish; once again the prophylactic use of antibiotics is a common practice. The project is aimed to characterise the microbial and viral communities of aquaculture samples and to assess the impact of their viral populations in the spreading of ARGs.

Pipeline: 20 l of the sample IN (before the contact with fishes) and the sample OUT (breeding tank wastewater) have been collected at Fondazione Edmund Mach – Technology Transfer Centre (San Michele all'Adige – TN – Italy); The tank was used for the breeding of different salmonids as Carpione del Garda (*Salmo carpio*), Trota marmorata (*Salmo trutta marmoratus*) and Salmerino alpino (*Salvelinus alpinus*).

Water samples underwent a double tangential flow filtration (TFF) step in order to separate the microbial community and concentrate the viral one. From both communities DNA have been extracted. Bacterial DNA have been analysed with 16S rRNA profiling. Viral DNA have been sequenced with a metagenomic shotgun approach. Data from 16S rRNA profiling has been analysed using QIIME pipeline (Quantitative Insight Into Microbial Ecology).
Data from metagenomic shotgun sequencing have been analysed using BLAST alignment against different database in order to characterise the viral community.

★ Indicate the steps for the removal and checking of the presence of microbial contaminats

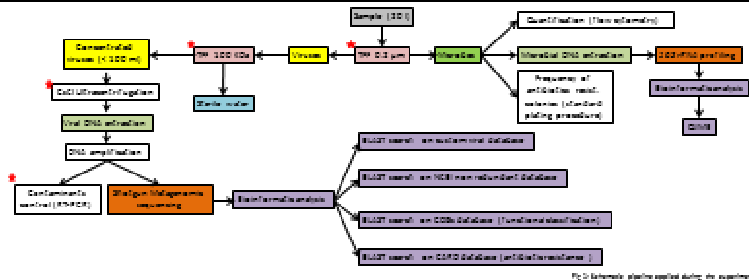
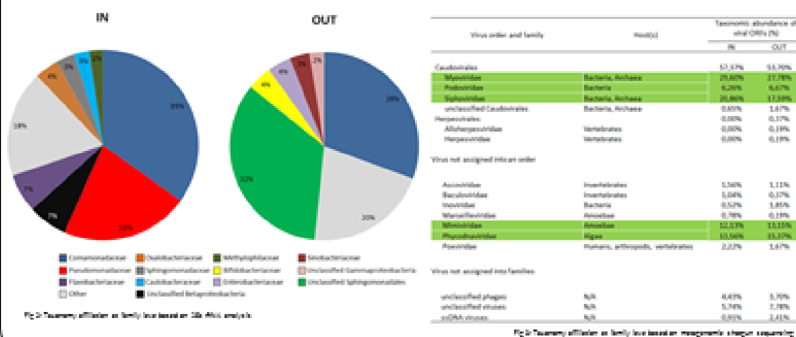
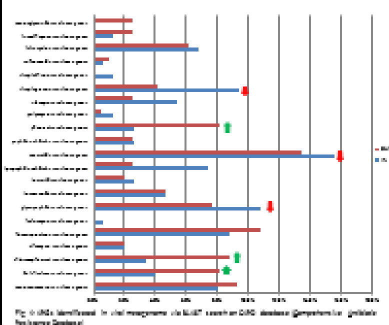


Fig 1: Schematic pipeline applied during the experiment

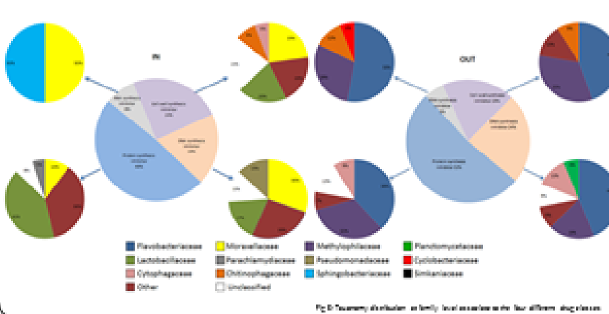
Taxonomy: Taxonomic affiliation at family level based on the results of 16S rRNA profiling analysis (Fig 2) and based on the predicted open reading frames (ORFs) from viral metagenomic, assigned by BLAST against NCBI taxonomy database (Fig 3). 16S rRNA profiling revealed a complete different scenario between the two samples while no substantial differences can be observed in the viral communities.



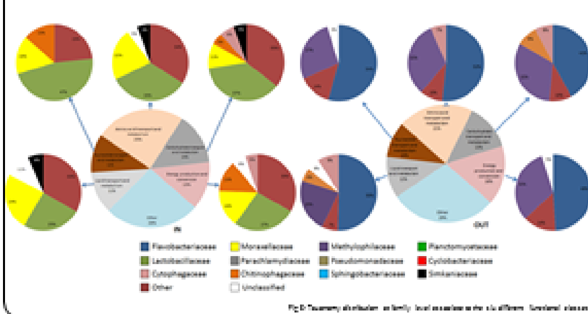
ARGs Identification: BLAST analysis against CARD database identified 21 different ARGs in the two samples. After the contact with fishes relative abundance of chloramphenicol, beta-lactam, aminoglycoside and phenicol resistance genes increase



Taxonomy of ARGs: ARGs identified with CARD database have been grupped in four different drug classes based on the inhibition target. Each drug class have been described with the associate taxonomy. Even if the relative abundance of each drug class does not significantly change, the taxonomy of the relative ORFs was totally different between the two samples.



Taxonomy of metabolic functional classes : ORFs identified with nr NCBI database have been grupped in six different functional classes based on the COGs classification. Each functional class have been described with the associate taxonomy. Even if the relative abundance of each functional class did not significantly change, the taxonomy of the relative ORFs was totally different between the two samples.

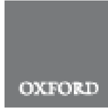


Conclusion: Small changes (<2%) have been observed between the two samples in viral community distribution. More than 50 % of the identified ORFs were associated to *Caudovirales*. A different scenario was observed in the bacterial community where large populations (*Pseudomonadaceae*, 22%) disappear and others (*Sphingomonadaceae*, 32%) significantly increased. The relative abundance of drug classes did not significantly change between the samples but the distribution of the associated taxonomy dramatically changed. This phenomenon has been confirmed also for the metabolic classes division, gaining importance due to the large amount of metabolic related ORFs (~ 50%). These data underline the changes in the microbial genes into the viral genomes, which show no differences between the two conditions.

References

[1] Muniesa M, Colomer-Lluch M, Jofre J. 2013. Could bacteriophages transfer antibiotic resistance genes from environmental bacteria to human-body associated bacterial populations? *Mobile Genetic Elements* 3:4
[2] FAO Fisheries and Aquaculture Department. 2013. *Global Aquaculture Production Statistics for the year 2011.*
[3] Armstrong SM, Hargrave BT, Haya K. 2005. Antibiotic Use in Finfish Aquaculture: Modes of Action, Environmental Fate, and Microbial Resistance. *Handb. Environ. Chem.*, 5M, pp 341–357.

* For Correspondence: diego.mora@unimi.it



FEMS Microbiology Ecology, 92, 2016, fiw003

doi: 10.1093/femsec/fiw003
Advance Access Publication Date: 5 January 2016
Research Article

RESEARCH ARTICLE

Virome-associated antibiotic-resistance genes in an experimental aquaculture facility

Stefano Colombo[1], Stefania Arioli[1], Simone Guglielmetti[1], Fernando Lunelli[2] and Diego Mora[1,*]

[1]Department of Food, Nutritional, and Environmental Sciences, University of Milan, 20133 Milan, Italy and [2]Fondazione Edmund Mach, Dipartimento Filiere Agroalimentari, Unità Acquacoltura e Idrobiologia, San Michele all'Adige, 38010 Trento, Italy

*Corresponding author: Department of Food, Nutritional, and Environmental Sciences, University of Milan, Via Celoria 2, 20133 Milan, Italy. Tel: +39-0250319133; Fax: +39-0250319238; E-mail: diego.mora@unimi.it
One sentence summary: Aquaculture-associated ARGs are mobilized, together with other bacterial genes, from bacteria to phages or vice versa even in the absence of antibiotic treatment.
Editor: Pascal Simonet

Downloaded from http://femsec.oxfordjournals.org/ by guest on November 24, 2016

ABSTRACT

We report the comprehensive characterization of viral and microbial communities within an aquaculture wastewater sample, by a shotgun sequencing and 16S rRNA gene profiling metagenomic approach. *Caudovirales* had the largest representation within the sample, with over 50% of the total taxonomic abundance, whereas approximately 30% of the total open reading frames (ORFs) identified were from eukaryotic viruses (*Mimiviridae* and *Phycodnaviridae*). Antibiotic resistance genes (ARGs) within the virome accounted for 0.85% of the total viral ORFs and showed a similar distribution both in virome and in microbiome. Among the ARGs, those encoding proteins involved in the modulation of antibiotic efflux pumps were the most abundant. Interestingly, the taxonomy of the bacterial ORFs identified in the viral metagenome did not reflect the microbial taxonomy as deduced by 16S rRNA gene profiling and shotgun metagenomic analysis. A limited number of ARGs appeared to be mobilized from bacteria to phages or *vice versa*, together with other bacterial genes encoding products involved in general metabolic functions, even in the absence of any antibiotic treatment within the aquaculture plant. Thus, these results confirm the presence of a complex phage–bacterial network in the aquaculture environment.

Keywords: virome; microbiome; aquaculture; antibiotic resistance genes; horizontal gene transfer

INTRODUCTION

According to recently released data (FAO - Fisheries and Aquaculture Department), world aquacultural production of fish for consumption has exceeded 60 million tons in 2011, with an increase of 6.2% over the previous year (FAO Fisheries and Aquaculture Department 2013). This large growth has been accompanied by an increased usage of a wide range of antibiotics (Armstrong, Hargrave and Haya 2005). In aquaculture, the main use for antibiotics is the prevention and treatment of bacterial infections in fish; the prophylactic use of antibiotics is a common practice as well. The necessity of antibiotic use in aquaculture is a consequence of lowered host defenses associated with high-density breeding under suboptimal hygienic conditions (Grave *et al.* 1999; Defoirdt *et al.* 2007; Cabello *et al.* 2013). Horizontal gene transfer among bacteria occurs by one of the three following mechanisms: conjugation, free DNA transformation or transduction through bacteriophages. Despite the potential importance of bacteriophages in transferring resistance genes

Received: 14 August 2015; Accepted: 29 December 2015
© FEMS 2016. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

from the environment to human and animal microbiomes, studies on this topic are limited (Modi et al. 2013). With respect to the virome associated with freshwater environments, studies only describe the taxonomic diversity and composition (López-Bueno et al. 2009; Rosario et al. 2009; Roux et al. 2011; Fancello et al. 2012; Tseng et al. 2013) or the viral and microbial community dynamics (Rodriguez-Brito et al. 2010).

In this work, we report a comprehensive characterization of viral and microbial communities in an experimental aquaculture sample using a metagenomics approach. The contemporary study of both communities resulted in the identification of different genes that are mobilized in the virome, with particular attention paid to antibiotic resistance genes (ARGs).

VIRAL TAXONOMIC CHARACTERIZATION

A water sample (20 L) was collected at the Edmund Mach Foundation – Technology Transfer Center (San Michele all'Adige, TN, Italy). The sample was collected from the wastewater of a fish breeding tank. Groundwater was initially pumped into a degasser tower for the removal of excess carbon dioxide prior to its use in the fish tanks. Tanks contained approximately 7 kg m^{-3} of salmonid (the fish density in a commercial aquaculture is close to 25 kg m^{-3}) with a water flow of approximately 83 m^3 h^{-1}. Antibiotic treatments had not been applied. The tank was used for the breeding of salmonids such as *Salmo carpio*, *Salmo trutta marmoratus* and *Salvelinus alpinus*. Water was treated to isolate virus-like particles (VLPs), from which DNA was extracted. The complete procedure is described in the Supplementary Materials and methods section. Metagenomic shotgun sequencing of viral DNA was determined using an Ion Torrent PGM platform (Life Technologies, Carlsbad, CA, USA). The MIRA program was used for *de novo* assembly of contigs, and ORFs were predicted by PRODIGAL (Supplementary Materials and methods).

According to several other studies (Fancello et al. 2012; Roux et al. 2012; Zablocki et al. 2014), the majority of the identified ORFs (accession number SAMEA3334740) originated from bacteria (86%) whereas those from viruses were the second most represented (13%). Thirteen families were identified (Supplementary Table S1), and prokaryotic viruses were the most abundant in the sample. *Caudovirales* accounted for more than 50% of the total taxonomic abundance. *Myoviridae* and *Siphoviridae* were the most represented families with 28% and 18% of the total ORFs, respectively. Analysis revealed the presence of viruses whose hosts include amoebae and algae; *Mimiviridae* and *Phycodnaviridae* were the main representatives of this group, at 13–15% of the total. Furthermore, the largest variety of phages identified interacts with *Proteobacteria*, which is consistent with the abundance of this phylum in the microbiome (Supplementary Fig. S1). The two main families of *Proteobacteria* represented in the water sample, *Sphingomonadaceae* and *Comamonadaceae*, and their related genera are typical constituents of freshwater environments, because bacterioplankton are opportunistic pathogens (Kallman et al. 2006; Kilic et al. 2007; Lin et al. 2010; Vaz-Moreira, Nunes and Manaia 2011; Chen et al. 2013). However, the most abundant ORFs identified in the virome were best matched with phages that infect *Bacillus*, *Synechococcus* and *Mycobacterium*, bacterial genera that have not been identified in the microbiome, thus suggesting that these genera were probably under-represented in the water sample because of a phage infection that resulted in a lytic cycle. In this context, the sharp reduction or local extinction of microbial taxa by viruses is a phenomenon that would be expected according to the 'Kill-the-Winner' dynamics hypothesis (Thingstad 2000). This dynamic model postulates a repetitive cycle in which an increase in prey population leads to an increase in the predator population that in turn decreases the prey population, thus causing its own subsequent decline.

DISTRIBUTION AND TAXONOMY OF ARGS IN THE VIROME AND MICROBIOME

Microbial cells were isolated to extract their DNA (Supplementary Materials and methods). 16S rRNA gene profiling (accession number SAMEA3333506) and metagenomics shotgun analysis (accession number SAMEA3334741) were performed on this sample. The ORFs obtained from shotgun metagenomic sequencing were analysed using the CARD database. The total number of ARGs identified was 950 (4.13% of the total ORFs identified with NCBI) in the microbiome and 214 (2.64%) in the virome. The distribution of the antibiotic resistance gene (ARG) identified in the two metagenomes (Supplementary Fig. S2) showed that the most abundant genes were those encoding proteins involved in the modulation of antibiotic efflux pumps (CARD nomenclature), with values ranging from 16 to 19% of the total identified ARGs. Among the different ARG classes, antibiotic efflux pumps had the most general function, acting on different target molecules and having a cell detoxifying activity (Pao, Paulsen and Saier 1998). Macrolide-resistance genes were the second most abundant group and included erythromycin, telithromycin and clarithromycin resistance genes. With few exceptions, the distribution of ARGs was similar between microbial and viral metagenomes. Glycopeptide efflux pump and lincosamide ARGs had higher values in the virome than in the microbiome (Supplementary Fig. S2), whereas gene modulating antibiotic efflux pumps and fluoroquinolone resistance genes were the most abundant in the microbiome. For each ARG identified, the taxonomy of the respective ORF was assigned based on the results obtained from the BLAST comparison with the NCBI refseq database. The taxonomy of the ARGs in the microbiome (Fig. 1) showed the presence of three main families (*Comamonadaceae* 26–37%, *Sphingomonadaceae* 18–24% and *Sinobacteraceae* 14–21%) equally distributed among the different drug classes, and a large number (15–26%) of the ARGs were shared among different microbial families, accounting for an amount lower than 4%. The virome analysis revealed that ARGs within this metagenome were mainly distributed between *Flavobacteriaceae* (38–53%) and *Methylophilaceae* (19–33%) families and that other bacterial families (7–13%) had a relative abundance below the 5% threshold (Fig. 1). None of the families present in both the virome and the microbiome, accounted for a relative abundance higher than 5%. From a taxonomic point of view, it is interesting to note that the microbial ORFs identified in the virome did not reflect the microbiome taxonomy, as determined by 16S rRNA gene profiling (Supplementary Fig. S3) or shotgun metagenomic analysis (Fig. 1), suggesting that microbial genes mobilized in the genome of viruses could be considered to be remnants of past recombination events rather than a picture of the current microbial diversity.

THE VIROME AND MICROBIOME SHARE ARGS

In order to identify possible gene mobilization events, ORFs located both in the microbiome and in the virome were identified (Supplementary Materials and methods). A total of 213 different ORFs were shared between the two metagenomes, and

Downloaded from http://femsec.oxfordjournals.org/ by guest on November 24, 2016

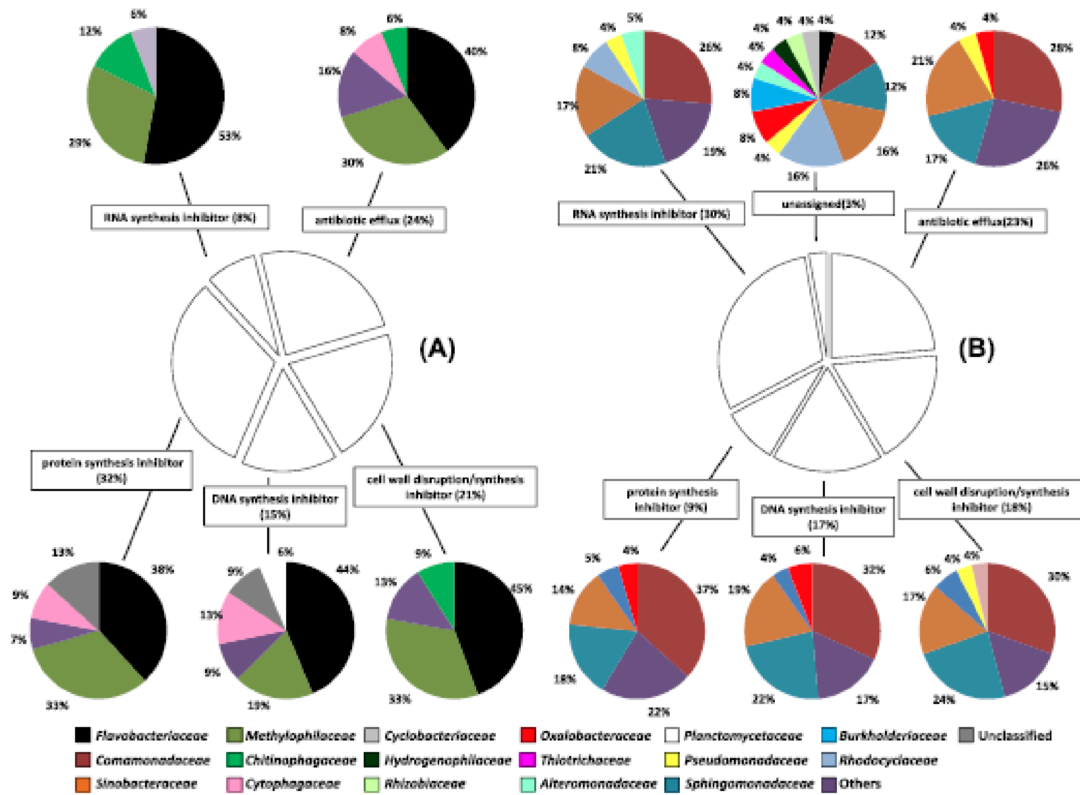


Figure 1. Comparison of the antibiotic resistance-related ORF distribution among viral (A) and microbial (B) metagenomes. ARGs have been divided into six different drug classes: RNA synthesis inhibitors, protein synthesis inhibitors, DNA synthesis inhibitors, cell wall disruption and synthesis inhibitors, antibiotic efflux modulators, and unassigned. Relative abundance (%) of each drug class was calculate with respect to the total of the classes. The taxonomic distribution at family level has been assigned to each drug class: data were expressed as relative abundance (%) with respect to the single drug class.

most of them were related to microbial metabolism. With respect to the ARGs identified in the virome, 4% were shared with the microbiome and had an amino acid sequence identity higher than 95%. Half of the identified ORFs encode genes modulating the antibiotic efflux (i.e. transcriptional regulator); ORFs that encode glycopeptide, tetracycline and β-lactam resistance genes were also identified. Among these genes, one was classified as a *Sphingopyxis alaskensis* β-lactamase encoding gene, which was putatively mobilized together with four other genes including a pseudogene coding for a β-lactamase of *Sphingomonas* sp. (Table 1). The other five genes shared between the virome and the microbiome are putatively involved in tetracycline and glycopeptide resistance and antibiotic efflux mechanisms. In all cases, shared ARGs were not transferred alone but with other genes hosted by the same contig (Table 1). Our data indicate that ARGs are present in the virome and some of these ARGs are mobilized from bacteria to phages or *vice versa*. These data suggest that ARGs may be mobilized even in the absence of selective pressure, i.e. in the absence of antibiotic treatment in aquaculture. Moreover, we showed that ARGs were mobilized together with other bacterial genes encoding products involved in more general metabolic functions, thus confirming the presence of a complex phage–bacterial network in the aquaculture environment.

ARGS PRESENT IN PUBLICLY AVAILABLE VIROMES

In order to compare our data with publicly available data, we calculated the relative amount of ARGs in the 17 viromes tested. Our sample showed a relative ARG abundance of 0.85%, the second highest value after that of El Barbera lake (Fancello *et al.* 2012). Interestingly, the relative amount of ARGs did not correlate with the presence of anthropic activities geographically located near the tested aquatic environments, because the highest ARG abundance was found in the El Barbera sample, which was collected from the Mauritanian desert. The presence of ARGs in an uncontaminated wild environment has been reported previously. Several studies have identified ARGs in different ecosystems, including soil and permafrost (D'Costa *et al.* 2006; Hallen *et al.* 2010), the human gut (Hu *et al.* 2013), and the microbiome of members of an isolated Yanomami Amerindian village (Clemente *et al.* 2015). Therefore, the diversity, taxonomic distribution and ecological role of antibiotic resistant genes in the environment are still far from being fully understood. Moreover, it is worth noting that the presence of ARGs in metagenomes does not directly represent a risk for human health, and a proper analysis of such risks should be carried out (Martinez, Coque and Baquero 2015).

Downloaded from http://femsec.oxfordjournals.org/ by guest on November 24, 2016

Table 1. ARGs shared between viral and microbial metagenomes.

ARGs shared between viral and microbial metagenome	Coding function[a]	Identity[b] (%)	Mobilized ORFs hosted in the same contig	contig no.	
				Virome	Microbiome
β-Lactam resistant gene (*Sphingopyxis alaskensis*)	β-Lactamase (class C family)	100	β-Lactam resistant gene partial (*Sphingomonas* sp.) Hypothetical protein (*Sphingomonas* sp.) Peptidase (*Sphingopyxis alaskensis*) membrane protein (*Sphingopyxis alaskensis*)	c4174	c208
Tetracycline resistant gene (*Caulobacter* sp.)	Transcriptional regulator, subgroup of AraC transcriptional regulators having an N-terminal Type 1 glutamine amidotransferase (GATase1)-like domain	100	Integrase (*Novosphingobium* sp.) Sarcosine oxidase subunit β (*Novosphingobium* sp.)	c4836	c92
Gene modulating antibiotic efflux (Oxalobacteraceae bacterium IMCC9480)	Transcriptional regulator (repressor function), TetR family	100	Cyclohexanone monooxygenase (*Brevundimonas subvibrioides* Acyl-CoA ligase (*Brevundimonas subvibrioides*) TonB-denpendent receptor (*Brevundimonas subvibrioides*) L-Carnitine dehydratase/bile acid-inducible protein F (*Caenispirillum salinarum*) α/β-Hydrolase (*Deinococcus phoenicis*) Putative flavin-containing monooxygenase (*Rhodococcus wratislaviensis* NBRC 100605)	c4172	c321
Glycopeptide resistance gene (*Sphingopyxis baekryungensis*)	AAA+ (ATPases Associated with a wide variety of cellular Activities)	100	GTPase HflX (*Sphingopyxis baekryungensis*) RNA-binding protein Hfq (*Sphingopyxis baekryungensis*)	c4235	c143
Glycopeptide resistance gene (*Sphingobium baderi*)	HATPase_C (Histidine kinase-like ATPases)	100	ATPase AAA (*Sphingopyxis baekryungensis*)		
Gene modulating antibiotic efflux (*Limnohabitans* sp. Rim28)	REC (cheY-homologous signal receiver domain) HATPase_C (Histidine kinase-like ATPases)	97	—	c73	c713
		98	—		

[a]Function assigned by NCBI conserved domains database.
[b]Amino acidic identity between the ORFs in the virome and the ortholog in the microbiome.

Downloaded from http://femsec.oxfordjournals.org/ by guest on November 24, 2016

In conclusion, this study addresses, for the first time, a complete description of the microbiome and virome in an aquaculture plant sample, relaying information on the presence of antibiotic resistance genes and their mobilization via transduction mechanisms.

SUPPLEMENTARY DATA

Supplementary data are available at FEMSEC online.

ACKNOWLEDGEMENTS

All bioinformatics analyses were performed using the facilities of CINECA SCAI (SuperComputing Applications and Innovation) (CINECA, Bologna, Italy) as part of the ISCRA VArFaWtr project.

Conflict of interest. None declared.

REFERENCES

Armstrong SM, Hargrave BT, Haya K. Antibiotic use in finfish aquaculture: modes of action, environmental fate, and microbial resistance. *Handb Environ Chem* 2005;**5M**:341–57.

Cabello FC, Godfrey HP, Tomova A et al. Antimicrobial use in aquaculture re-examined: Its relevance to antimicrobial resistance and to animal and human health. *Environ Microbiol* 2013;15:1917–42.

Chen H, Jogler M, Rohde M et al. *Sphingobium limneticum* sp. nov. and *Sphingobium boeckii* sp. nov., two freshwater planktonic members of the family *Sphingomonadaceae*, and reclassification of *Sphingomonas suberifaciens* as *Sphingobium suberifaciens* comb. nov. *Int J Syst Evol Microbiol* 2013;63:735–43.

Chevreux B, Wetter T, Suhai S. Genome sequence assembly using trace signals and additional sequence information. In: *Computer Science and Biology: Proceedings of the German Conference on Bioinformatics (GCB)*. 1999, 45–56.

Clemente JC, Pehrsson EC, Blaser MJ et al. The microbiome of uncontacted Amerindians. *Sci Adv* 2015;e1500183.

D'Costa VM, McGrann KM, Hughes DW et al. Sampling the antibiotic resistome. *Science* 2006;**311**:374–7.

Defoirdt T, Boon N, Sorgeloos P et al. Alternatives to antibiotics to control bacterial infections: luminescent vibriosis in aquaculture as an example. *Trends Biotechnol* 2007;25:472–9.

Fancello L, Trape S, Robert C et al. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J* 2012;7:359–69.

FAO Fisheries and Aquaculture Department. *Global Aquaculture Production Statistics for the year 2011*. 2013. ftp://ftp.fao.org/fi/news/GlobalAquacultureProductionStatistic2011.pdf (4 February 2016, date last accessed).

Grave K, Lingaas E, Bangen M et al. Surveillance of the overall consumption of antibacterial drugs in humans, domestic animals and farmed fish in Norway in 1992 and 1996. *J Antimicrob Chemother* 1999;**43**:243–52.

Hallen HK, Donato J, Wang HH et al. Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 2010;8:251–59.

Hu Y, Yang X, Qin J et al. Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 2013;4:2151.

Kallman O, Lundberg C, Wretlind B et al. Gram negative bacteria from patients seeking medical advice in Stockholm after the tsunami catastrophe. *Scand J Infect Dis* 2006;**38**:448–50.

Kilic A, Senses Z, Kurekci AE et al. Nosocomial outbreak of *Sphingomonas paucimobilis* bacteremia in a hemato/oncology unit. *Jpn J Infect Dis* 2007;60:394–6.

Lin JN, Lai CH, Chen YH et al. *Sphingomonas paucimobilis* bacteremia in humans: 16 case reports and a literature review. *J Microbiol Immunol Infect* 2010;43:35–42.

López-Bueno A, Tamames J, Velázquez D et al. High diversity of the viral community from an Antarctic lake. *Science* 2009;326:858–61.

Martinez JL, Coque TM, Baquero F. Waht is a resistant gene? Ranking risk in resistome. *Nat Rev Microbiol* 2015;13:116–23.

Modi SL, Lee HH, Spina CS et al. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. *Nature* 2013;499:219–23.

Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. *Microbiol Mol Biol Rev* 1998;62:1–34.

Rodriguez-Brito B, Li L, Wegley L et al. Viral and microbial community dynamics in four aquatic environments. *ISME J* 2010;4:739–51.

Rosario K, Nilsson C, Lim YW et al. Metagenomic analysis of viruses in reclaimed water. *Environ Microbiol* 2009;11:2806–20.

Roux S, Enault F, Robin A et al. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* 2012;7:e33641.

Roux S, Faubladier M, Mahul A et al. Metavir: A web server dedicated to virome analysis. *Bioinformatics* 2011;27:3074–75.

Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr* 2000;45:1320–8.

Tseng CH, Chiang PW, Shiah FK et al. Microbial and viral metagenomes of a subtropical freshwater reservoir subject to climatic disturbances. *ISME J* 2013;7:2374–86.

Vaz-Moreira I, Nunes OC, Manaia CM. Diversity and antibiotic resistance patterns of *Sphingomonadaceae* isolates from drinking water. *Appl Environ Microbiol* 2011;77:5697–706.

Zablocki O, Van Zyl L, Adriaenssens EM et al. High-level diversity of tailed phages, eukaryote-associated viruses, and virophage-like elements in the metaviromes of antarctic soils. *Appl Environ Microb* 2014;80:6888–97.

Downloaded from http://femsec.oxfordjournals.org/ by guest on November 24, 2016

7. Conclusion

My PhD project aims to investigate two great questions of the microbiology: the role and ecology of the viruses and their possible role in the spreading of the antibiotic resistance genes among the ecosystems.

The approach applied in the three projects is the most complete nowadays available to describe environmental samples from a microbiological point of view. The major limits in the description of the viral community are attributable to the incompleteness of the viral databases together with the absence of a viral analogs of the microbial ribosomal genes which allow a taxonomic gene profiling approach for bacteria or eukarya (i.e. 16S or ITS).

The three different projects showed how antibiotic resistance genes are wide spread in different environments irrespective of the presence of associated anthropic activities. In this context, we hypothesize that the release of antibiotics molecules in the environment by the microbiota is a driving force able to maintain ARGs in the microbiome and, consequently, mobilize them in the virome. In addition, we should considered the effect of the intense environmental release of antibiotics by humans activities. How can this release impact the mobilization of ARGs in a defined environment? Working directly on environmental samples, it becomes extremely difficult to separate these two effects due to the complexity of the samples beyond the presence of incalculable environmental and matrix variables. For these reasons, it is necessary to make a step back from the analysis of the final sample and to focus on: i) how the transduction mechanisms work; ii) how the presence of a high antibiotic selective pressure can modulate the transduction mechanisms. In this context, the development of mesocosms can be the turning point since this system let us to work with complex environmental samples but focusing the attention only at one variable at time as, for example, the supplementation of defined amount of antibiotics.

8. Materials And Methods

8.1 Sampling

8.1.1 Sampling of wastewater from an aquaculture plant (Chapter 3)

A water sample (20 liters) was collected on March 2014 at the Edmund Mach Foundation – Technology Transfer Center (San Michele all'Adige – TN – Italy; 46°11'42.9"N 11°08'10.1"E). The sample was collected from fish breeding tank wastewater. Groundwater was initially pumped into a degasser tower for the removal of excess carbon dioxide and then used for the fish tanks. Tanks contained approximately 7 kg/m^3 of salmonid (the fish density in a commercial aquaculture is close to 25 kg/m^3) with a water flow of approximately 83 m^3 /h.

Antibiotic treatments were not applied. The tank was used for the breeding of different salmonids, such as Salmo carpio, Salmo trutta marmoratus and Salvelinus alpinus.

8.1.2 Sampling and evaluation of antibiotic-resistant colony frequencies in water samples (Chapter 4)

Three water samples (30 L) were collected in February 2015 along the Lambro River. The first sample was collected at the river's spring in the Magreglio area (45.935270, 9.264576). The second sampling point was near the city of Ponte Lambro (45.49349, 9.13274) approximately 15 km from the origin. The third sampling point was in Milan in Lambro Park (45.495399, 9.247936) approximately 55 km from the origin. Samples were maintained at 6 °C for 4 h and then processed. For the evaluation of antibiotic resistant colony frequencies, samples from the 0.2 µm tangential flow filtration system retentate (see next paragraph) were used for dilution and plating on two selective media: plate count agar (PCA) and Mac Conkey (MC) agar. Both media were also tested with the addition of antibiotics such as ampicillin, tetracycline, chloramphenicol and meropenem. The antibiotics were present at a final concentration of 20 µg/mL. Samples were incubated at 20 °C for 48 h. The microbial count was carried out in triplicate.

8.1.3 Sampling of air an cheese related matrices in two dairy plant (Chapter 5)

Samples were collected in two dairy plants, one for the production of Grana Padano and one for the production of Gorgonzola. In particular, were collected Grana Padano whey, Gorgonzola curd and Gorgonzola rind after 50 days of ripening. Beyond that, 30 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 with 0.01% Tween 20: the cyclonic effect developed by the Coriolis system trapped microorganisms and particles in the liquid.

8.2 Virus-like particle isolation

Water samples were filtered at 0.2 µm (Pall, Life Sciences, Milan, Italy) with a tangential flow filtration system (TFF) (Quattro systems, Parma, Italy), as previously described in Colombo *et al.* (2016). Briefly, the permeate containing all the virus-like particles (VLPs) was later filtered at 100 kDa (Pall, Life Sciences, Milan, Italy) using the same TFF system; pressure was maintained under 62 kPa in order to avoid VLP damage (Thurber *et al.,* 2009). The retained material containing the VLPs (400 mL) was then precipitated overnight at 4 °C using PEG 8000 at a final concentration of 10% (wt/vol) 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 then deposited on top of a 2.5 mL step gradient composed of multiple 0.8 mL CsCl solutions with respective 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 by 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.

8.3 Extraction and amplification of VLP DNA for metagenomic analysis

Viral DNA was extracted according to a protocol previously described in 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 M Tris HCl/0.2 M 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 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 M NaCl and 80 µL of a solution of 10% cetyltrimethylammonium bromide/0.7 M 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 was 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).

In order to increase the amount of DNA available, whole-genome amplification was performed using reagents and protocols in the Illustra GenomiPhi V2 kit (GE Healthcare, Milan, Italy) to generate sufficient material for metagenomic analysis. Purified VLP DNA (5–15 ng) was mixed with 9 µL of sample buffer provided by Illustra GenomiPhi V2 kit and heat-denatured at 95 °C for 3 min. Next, 9 µL of kit reaction buffer and 1 µL of kit enzyme mix were added, and the solution was incubated for 90 min at 30 °C. The amplified products were subsequently pooled and purified (UltraClean PCR Clean-UP Kit, MoBio Laboratories, Solana Beach, CA, USA) (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).

8.4 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., Milano, 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 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 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.

8.5 Extraction of microbial DNA for metagenomic analysis

The permeate from 0.2 μm tangential flow filtration, used in the first step of VLP recovery as described above, was centrifuged for 15 min at 9000 x g and 4 °C and then suspended in 400 μL of TE buffer (pH 8). To this, 15 μL of lysozyme (20 mg mL^{-1}) was added before a 60 min incubation at 37 °C. Next, 15 μL of SDS (20%) and 15 μL of proteinase K (20 mg mL^{-1}) (Fisher Scientific) were added, and the mixture was incubated at 55 °C for 15 min. Phenol (250 μL) was added, and the mixture was centrifuged at 22,000 x g for 5 min. The phenol phase was removed and 400 μL of chloroform was added to the aqueous supernatant. The centrifugation step was repeated as before. The supernatant was transferred to a new tube with 40 μL of sodium acetate (3 M pH 5.2) and 800 μL of ethanol (-20 °C). The mixture was centrifuged for 30 min at 22,000 x g and 4 °C. The liquid was discarded and the pellet was resuspended in 300 μL of 70% ethanol (-20 °C). The centrifugation step was repeated as before, and the liquid was discarded. The DNA was dried and suspended in 40 μL of sterile distilled water. The PowerWave XS Microplate Spectrophotometer at 260 nm (BioTek) and the Take3 Multi-Volume Plate (BioTek) were used to quantify 2 μL of the DNA solution.

8.6 16S rRNA gene amplification

The bacterial composition of the water samples was determined by assessing the distribution of *16S rRNA* gene sequences by Ion Torrent PGM sequencing technology. Partial *16S rRNA* gene amplification (with the primer pair Probio_Uni and Probio_Rev, which targets the V3 region) and sequencing reactions were performed by GenProbio, according to the optimized protocol described by Milani *et al.* (2013). These primers were designed to include at their 5′ end one of the two

adaptor sequences used in the Ion Torrent sequencing library preparation protocol, which links a unique 10-bp tag barcode to identify different samples. The PCR conditions used were 5 min at 95°C, 35 cycles of 30 s at 94°C, 30 s at 55°C and 90 s at 72°C, followed by 10 min at 72°C. Amplification was carried out by a Verity Thermocycler (Applied Biosystems). The integrity of the PCR amplicons was analyzed by electrophoresis on an Experion workstation (BioRad, UK) (Milani *et al.,* 2013).

8.7 Ion Torrent PGM sequencing of *16S rRNA* gene-based amplicons

The PCR products derived from amplification of *16S rRNA* gene sequences were purified by electrophoretic separation on a 2% agarose gel, and purified with the Wizard SV Gen PCR Clean-Up System (Promega), followed by a further purification step involving Agencourt AMPure XP DNA purification beads (Beckman Coulter Genomics GmbH) to remove primer dimers. The DNA concentration of the amplified sequence library was estimated using the Experion system (BioRad Laboratories). From the concentration and the average size of each amplicon library, the amount of DNA fragments per microliter was calculated, and libraries for each run were diluted to 3 x 10^9 DNA molecules prior to clonal amplification. Emulsion PCR was carried out using the Ion OneTouch™ 200 Template Kit v2 DL (Life Technologies) according to the manufacturer's instructions. Sequencing of the amplicon libraries was carried out on a 316 chip using the Ion Torrent PGM system and employing the Ion Sequencing 200 kit (Life Technologies) according to the supplier's instructions (Milani *et al.,* 2013). After sequencing, the individual sequence reads were filtered by the PGM software to remove low quality and polyclonal sequences. Sequences matching the PGM 3′ adaptor were also automatically trimmed. All PGM quality-approved, trimmed and filtered data were exported as BAM files (Milani *et al.,* 2013).

8.8 16S rRNA-based microbiota analysis

The BAM files were processed using QIIME (Caporaso *et al.,* 2010). Quality control retained sequences with a length of 140 to 400 bp, homopolymers < 8 bp and sequences with a mean sequence quality score > 25, with truncation of a sequence at the first base if a low quality rolling 10-bp window was found. In addition, primer sequences were removed and reads, in which at least a single primer was missing, were discarded to retain only full-length amplicons. In order to reach the genus taxonomic level, the standard QIIME pipeline was modified using *de novo* OTU

(Operational Taxonomic Unit) at 97% identity, the GreenGenes *16S rRNA* gene database of reference OTUs at 97% identity and a modified GreenGenes ID taxonomy tabular file (DeSantis *et al.,* 2006).

8.9 Metagenomic shotgun sequencing of microbial and 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, Wetter and Suhai 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 iontor_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)

8.10 Quality control of metagenomic reads.

Reads from metagenomic shotgun sequencing were initially checked for quality using the FASTQC program (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Fastq files were trimmed and filtered with FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) and homerTool (http://homer.salk.edu/homer/ngs/homerTools.html). The fastq output was filtered for reads with a quality of < 25, as well as reads of 80 bp. Bases were removed from the ends of the reads until the average quality in a window of 5 bp was > 25. Contaminant reads (i.e. human reads) were removed with CS-SCORE (Haque *et al.,* 2015). Sequencing metadata, assembly metrics and annotation statistics are summarized in Table S1 of the Supporting Information.

Identification of contaminant reads in viral metagenomes has been performed with a BLASTn analysis on the "16SMicrobial" database provided by NCBI (ftp://ftp.ncbi.nlm.nih.gov/blast/db/). BLASTn was performed with a threshold of 10-5 on the E value and only results with identity over 90% were kept. Levels of contamination were then expressed in percentage (%) as number of reads identified with the "16SMicrobial" database over the total number of reads.

8.11 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.

Both viral and microbial data underwent a BLAST analysis against the both the non-redundant nucleotide and the non-redundant protein database 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.

Identification of bacteriocines related reads was performed with BLASTp and BLASTn against BACTIBASE database (Hammami *et al.,* 2007). Identification of metal resistance reads was performed with BLASTp and BLASTn against BacMet database (Pal *et al.,* 2014). All bioinformatics analyses were performed using the CINECA SCAI (SuperComputing Applications and Innovation) (CINECA, Bologna, Italy), part of the ISCRA VArFaWtr project.

9. Supplementary Tables

Table S1. Taxonomic abundance of ORFs (BLASTp with a threshold of 10^5 for the E value) identified by MetaVir in the aquaculture virome

Virus order and family	Host(s)	Taxonomic abundance of viral ORFs (%)
Caudovirales		53,70%
Myoviridae	Bacteria, Archaea	27,78%
Podoviridae	Bacteria	6,67%
Siphoviridae	Bacteria, Archaea	17,59%
unclassified Caudovirales	Bacteria, Archaea	1,67%
Herpesvirales		0,37%
Alloherpesviridae	Vertebrates	0,19%
Herpesviridae	Vertebrates	0,19%
Virus not assigned into an order		
Ascoviridae	Invertebrates	1,11%
Baculoviridae	Invertebrates	0,37%
Inoviridae	Bacteria	1,85%
Marseilleviridae	Amoebae	0,19%
Mimiviridae	Amoebae	13,15%
Phycodnaviridae	Algae	15,37%
Poxviridae	Humans, arthropods, vertebrates	1,67%
Virus not assigned into families		
unclassified phages	N/A	3,70%
unclassified viruses	N/A	7,78%
ssDNA viruses	N/A	2,41%

N/A not assigned

Table S2. ARGs genes shared between viral and microbial metagenomes

ARGs shared between viral and microbial metagenome	Coding function [a]	Identity [b] (%)	Mobilized ORFs hosted in the same contig	contig n. virome	contig n. microbiome
β-lactam resistant gene (Sphingopyxis alaskensis)	beta-lactamase (class C family)	100%	β-lactam resistant gene partial (Sphingomonas sp.) hypothetical protein (Sphingomonas sp.) peptidase (Sphingopyxis alaskensis) membrane protein (Sphingopyxis alaskensis)	c4174	c208
Tetracycline resistant gene (Caulobacter sp.)	Transcriptional regulator, subgroup of AraC transcriptional regulators having an N-terminal Type 1 glutamine amidotransferase (GATase1)-like domain	100%	integrase (Novosphingobium sp.) sarcosine oxidase subunit beta (Novosphingobium sp.)	c4836	c92
Gene modulating antibiotic efflux (Oxalobacteraceae bacterium IMCC9480)	Transcriptional regulator (repressor function), TetR family	100%	cyclohexanone monooxygenase (Brevundimonas subvibrioides acyl-CoA ligase (Brevundimonas subvibrioides) TonB-denpendent receptor (Brevundimonas subvibrioides) L-carnitine dehydratase/bile acid-inducible protein F (Caenispirillum salinarum) alpha/beta hydrolase (Deinococcus phoenicis) putative flavin-containing monooxygenase (Rhodococcus wratislaviensis NBRC 100605)	c4172	c321
Glycopeptide resistance gene (Sphingopyxis baekryungensis) Glycopeptide resistance gene (Sphingobium baderi)	AAA+ (ATPases Associated with a wide variety of cellular Activities) HATPase_C (Histidine kinase-like ATPases)	100% 100%	GTPase HflX (Sphingopyxis baekryungensis) RNA-binding protein Hfq (Sphingopyxis baekryungensis) ATPase AAA (Sphingopyxis baekryungensis)	c4235	c143
Gene modulating antibiotic efflux (Limnohabitans sp. Rim28)	REC (cheY-homologous signal receiver domain) HATPase_C (Histidine kinase-like ATPases)	97% 98%	/ /	c73	c713

[a], function assigned by NCBI conserved domains database; [b], amino acidic identity between the ORFs in the virome and the orthologue in the microbiome

Table S3. Relative abundance of ARGs, with respect to the total amount of ORFs identified, in several freshwater viromes. (BLASTx with a threshold of 10^5 for the E value against CARD database).

Freshwater sample	ARGs	% ARGs
Elbarbera (Fancello *et al.,* 2012)	638	0,9036%
Aquaculture (this study)	214	0,8463%
Antarctic lake summer (López-Bueno *et al.,* 2009)	136	0,4723%
FTR_0707 (Tseng *et al.,* 2013)	441	0,4309%
FTR_01_08 (Tseng *et al.,* 2013)	267	0,3642%
Ilij (Fancello *et al.,* 2012)	186	0,3189%
Hamdoun (Fancello *et al.,* 2012)	118	0,2893%
FTR_0807(Tseng *et al.,* 2013)	234	0,2641%
Molomhar (Fancello *et al.,* 2012)	138	0,1831%
RW_nursery_DNA (Rosario *et al.,* 2009)	80	0,0311%
Lake Bourget (Roux *et al.,* 2012)	175	0,0223%
RW_effluent_DNA (Rosario *et al.,* 2009)	49	0,0205%
Tilapia_11 (Rodriguez-Brito *et al.,* 2010)	6	0,0152%
Antarctic lake spring (López-Bueno *et al.,* 2009)	4	0,0103%
Tilapia0406 (Rodriguez-Brito *et al.,* 2010)	4	0,0090%
Tilapia_08 (Rodriguez-Brito *et al.,* 2010)	3	0,0085%
Lake Pavin (Roux *et al.,* 2012)	34	0,0085%
RW_park (Rosario *et al.,* 2009)	9	0,0054%

1 **Table S4 Relative abundance of reads associated to MRGs and biocide**
2 **resistant genes**

	Microbial sample			Viral sample		
	Spring	Ponte L.	Milan	Spring	Ponte L.	Milan
2-nitroimidazole [class: imidazole]	0,48%	0,70%	0,69%	0,02%	0,43%	0,61%
6-dichloroindophenol [class: Phenolic Compounds]	0,10%	0,07%	0,06%	0,03%	-	-
Aluminium (Al)	0,13%	0,19%	0,20%	0,05%	0,33%	4,36%
Antimony (Sb)	0,72%	0,74%	0,99%	0,57%	1,12%	5,69%
Arsenic (As)	2,83%	3,15%	3,23%	3,28%	3,62%	6,90%
Benzylkonium Chloride (BAC) [class: Quaternary Ammonium Compounds (QACs)]	0,02%	0,02%	0,02%	-	-	-
Bismuth (Bi)	0,08%	0,07%	0,07%	0,05%	0,18%	0,61%
Cadmium (Cd)	4,86%	4,76%	4,96%	3,00%	5,00%	2,66%
Cetylpyridinium Chloride (CPC) [class: Quaternary Ammonium Compounds (QACs)]	0,48%	0,70%	0,69%	0,02%	0,43%	0,61%
Chromium (Cr)	2,52%	3,24%	2,95%	0,27%	2,35%	3,75%
Cobalt (Co)	6,94%	6,66%	6,83%	2,27%	6,27%	3,75%
Copper (Cu)	12,80%	12,32%	12,75%	10,18%	20,46%	15,62%
Crystal Violet [class: Triarylmethane]	1,53%	1,11%	1,14%	0,33%	0,43%	0,24%
Cyclohexane [class: Cycloalkane]	0,22%	0,15%	0,14%	0,47%	0,07%	0,12%
Dequalinium [class: Quaternary Ammonium Compounds (QACs)]	-	-	-	0,18%	-	-
Diphenyl Ether [class: Phenyl]	0,05%	0,04%	0,04%	0,05%	-	0,12%
Dodine [class: Acetate]	0,51%	0,71%	0,72%	0,02%	0,43%	0,61%
Ethidium Bromide [class: Phenanthridine]	0,03%	0,59%	0,56%	0,34%	0,22%	-
Gallium (Ga)	2,09%	1,97%	2,17%	2,44%	1,09%	1,33%
Glycerol [class: Alcohol]	0,12%	0,08%	0,09%	0,12%	0,22%	-
Gold (Au)	1,65%	1,23%	1,30%	0,25%	0,47%	0,24%
Hydrochloric acid (HCl) [class: Acid]	0,30%	0,32%	0,35%	0,12%	0,07%	0,12%
Hydrogen Peroxide (H2O2) [class: Peroxides]	1,43%	1,58%	1,60%	3,82%	1,34%	4,12%
Iron (Fe)	8,13%	7,73%	8,01%	15,56%	5,90%	8,23%
Lead (Pb)	1,59%	1,75%	1,58%	0,67%	1,67%	0,73%
Magnesium (Mg)	1,18%	1,13%	1,25%	0,77%	1,01%	0,97%
Manganese (Mn)	2,44%	2,00%	2,15%	7,17%	1,23%	3,39%
Menadione [class: Naphthoquinone]	0,06%	0,11%	0,06%	-	-	0,12%
Mercury (Hg)	1,57%	1,93%	1,78%	4,05%	1,38%	2,30%

3
4

Methyl Viologen [class: Paraquat]	0,09%	0,12%	0,08%	0,18%	-	0,12%
Methylene Blue [class: Thiazinium]	1,57%	1,18%	1,19%	0,15%	0,47%	0,36%
Methylmercury Acetate [class: Organo-mercury]	0,06%	0,04%	0,04%	0,24%	0,04%	-
Molybdenum (Mo)	3,11%	3,13%	2,90%	5,09%	2,32%	2,18%
n-hexane [class: Alkane]	0,05%	0,04%	0,04%	0,05%	-	0,12%
Nickel (Ni)	7,37%	7,11%	7,07%	12,77%	7,68%	4,72%
Pentane [class: Alkane]	0,05%	0,04%	0,04%	0,05%	-	0,12%
Phenylmercury Acetate [class: Organo-mercury]	0,69%	0,77%	0,71%	1,74%	0,22%	0,97%
Plumbagin [class: Naphthoquinone]	0,01%	0,01%	0,01%	0,00%	0,04%	-
Proflavin [class: Acridine]	0,03%	0,04%	0,05%	0,18%	0,07%	-
Selenium (Se)	1,57%	2,02%	1,79%	1,03%	1,67%	5,45%
Silver (Ag)	2,86%	3,31%	3,12%	2,87%	8,51%	1,09%
Sodium acetate [class: Acetate]	0,13%	0,13%	0,12%	0,65%	0,14%	0,12%
Sodium azide [class: Azide]	0,17%	0,13%	0,18%	-	0,07%	0,12%
Sodium Chenodeoxycholate [class: Acid]	0,58%	0,58%	0,54%	0,15%	0,22%	-
Sodium Cholate [class: Acid]	0,58%	0,58%	0,54%	0,15%	0,22%	-
Sodium Deoxycholate (SDC) [class: Acid]	1,38%	1,29%	1,20%	1,08%	1,16%	0,36%
Sodium Deoxycholate [class: Acid]	0,58%	0,58%	0,54%	0,15%	0,22%	-
Sodium Dodecyl Sulfate (SDS) [class: Organo-sulfate]	0,58%	0,58%	0,54%	0,15%	0,22%	-
Sodium Taurocholate [class: Acid]	0,58%	0,58%	0,54%	0,15%	0,22%	-
Tellurium (Te)	2,37%	2,36%	2,46%	0,92%	2,06%	5,21%
Tetraphenylphosphonium (TPP) [class: Quaternary Ammonium Compounds (QACs)]	-	-	-	0,18%	-	-
Tributyltin (TBT) [class: Organo-tin]	0,01%	0,01%	0,01%	0,99%	-	-
Triclosan [class: Phenolic compounds]	0,54%	0,59%	0,60%	0,04%	0,29%	0,36%
Tungsten (W)	5,20%	5,29%	4,86%	6,98%	3,88%	2,91%
Vanadium (V)	0,65%	0,71%	0,77%	0,05%	0,33%	0,48%
Zinc (Zn)	14,31%	13,33%	13,33%	7,80%	14,02%	8,11%

5
6
7

8 **Table S5. Relative abundance of reads associated to bacteriocin genes**

9

10

	Microbial sample			Viral sample		
	Spring	Ponte L.	Milan	Spring	Ponte L.	Milan
Boticin B	-	0,5%	-	-	-	-
Butyrivibriocin AR10	-	-	0,4%	-	-	-
Carocin D	0,5%	0,2%	-	-	-	3%
Colicin-Ib	3,0%	0,5%	-	-	-	6%
Colicin-M	0,5%	-	-	-	-	-
Colicin-V (Microcin-V)	-	0,2%	-	-	-	-
Elgicins	0,3%	0,2%	-	-	-	-
Enterocin B	0,5%	-	-	-	-	-
Enterocin Xbeta	0,3%	-	-	-	-	-
Enterolysin A	-	-	1,2%	-	-	3%
Helveticin-J	-	0,2%	-	-	-	-
Hiracin JM79	-	0,2%	-	-	-	-
lactococcin-A	-	-	0,4%	-	-	-
Plantaricin ASM1	-	0,2%	-	-	-	-
Propionicin SM1	-	-	0,4%	-	-	-
Propionicin T1	-	0,2%	-	-	-	-
Pyocin S1	5,8%	0,2%	3,1%	25%	-	-
Pyocin S2	-	0,2%	-	-	-	-
Thermophilin A	-	0,2%	-	-	-	-
Zoocin A	89,0%	96,7%	94,5%	75%	100%	89%

11

12

13

14

15 **Table S6. Virome-contigs harboring phage genes and ARGs**

16

<table>
<tr><th>Taxonomy</th><th>Aminoacidic identiti (%)</th><th>Sample</th><th>ARGS</th><th>ORF name</th></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Myoviridae|unassigned Myoviridae|Bacillus phage G</td><td>64,62</td><td>Virome_spring</td><td>aminocoumarin resistance gene</td><td>_c101_7</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Phycodnaviridae|Phaeovirus|Ectocarpus siliculosus virus 1</td><td>32,46</td><td>Virome_spring</td><td>gene modulating antibiotic efflux</td><td>_c61_7</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Phycodnaviridae|Phaeovirus|Feldmannia species virus</td><td>33,55</td><td>Virome_ponte_lambro</td><td>gene modulating antibiotic efflux</td><td>_rep_c2366_2</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Phycodnaviridae|Chlorovirus|unclassified Chlorovirus|Paramecium bursaria Chlorella virus CviKI</td><td>64,52</td><td>Virome_Milan</td><td>chloramphenicol resistance gene</td><td>_rep_c11690_1</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Podoviridae|P22likevirus|unclassified P22likevirus|Enterobacteria phage ST104</td><td>52,27</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c10649_1</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Myoviridae|unclassified Myoviridae|Microcystis phage MaMV-DC</td><td>47,5</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c11061_1</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Myoviridae|unclassified Myoviridae|Microcystis phage MaMV-DC</td><td>52,59</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c11199_1</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Podoviridae|P22likevirus|unclassified P22likevirus|Enterobacteria phage ST104</td><td>39,07</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c11255_2</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Siphoviridae|unclassified Siphoviridae|Salmonella phage vB_SosS_Oslo</td><td>44,14</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c6658_1</td></tr>
<tr><td>root|Viruses|dsDNA viruses, no RNA stage|Caudovirales|Podoviridae|P22likevirus|unclassified P22likevirus|Enterobacteria phage ST104</td><td>40,94</td><td>Virome_Milan</td><td>fluoroquinolone resistance gene</td><td>_rep_c7264_1</td></tr>
</table>

17

18

19 **Table S7. Relative abundance of ARGs, with respect to the total amount of**
20 **ORFs identified, in several freshwater viromes. (BLASTx with a threshold**
21 **of 10^5 for the E value against CARD database).**
22

Freshwater sample	ARGs	% ARGs
Spring (present study)	**163**	**2,9543%**
Elbarbera (Fancello *et al.,* 2012)	638	0,9036%
Antarctic lake summer (López-Bueno *et al.,* 2009)	136	0,4723%
Ponte Lambro (present study)	**36**	**0,4563%**
FTR_0707 (Tseng *et al.,* 2013)	441	0,4309%
FTR_01_08 (Tseng *et al.,* 2013)	267	0,3642%
Ilij (Fancello *et al.,* 2012)	186	0,3189%
Hamdoun (Fancello *et al.,* 2012)	118	0,2893%
FTR_0807(Tseng *et al.,* 2013)	234	0,2641%
Molomhar (Fancello *et al.,* 2012)	138	0,1831%
Milan (present study)	**9**	**0,0523%**
RW_nursery_DNA (Rosario *et al.,* 2009)	80	0,0311%
Lake Bourget (Roux *et al.,* 2012)	175	0,0223%
RW_effluent_DNA (Rosario *et al.,* 2009)	49	0,0205%
Tilapia_11 (Rodriguez-Brito *et al.,* 2010)	6	0,0152%
Antarctic lake spring (López-Bueno *et al.,* 2009)	4	0,0103%
Tilapia0406 (Rodriguez-Brito *et al.,* 2010)	4	0,0090%
Tilapia_08 (Rodriguez-Brito *et al.,* 2010)	3	0,0085%
Lake Pavin (Roux *et al.,* 2012)	34	0,0085%
RW_park (Rosario *et al.,* 2009)	9	0,0054%

10. References

Abeles SR, Ly M, Santiago-Rodriguez TM, *et al.* Effects of long term antibiotic therapy on human oral and fecal viromes. *Plos One.*2015; 10:e0134941

Allen HK, Donato J, Wang HH *et al*. Call of the wild: antibiotic resistance genes in natural environments. *Nature Rev. Microbiol*. 2010; 8, 251-259.

Aminov RI. The role of antibiotics and antibiotic resistance in nature. *Environ. Microbiol*. 2009; 11, 2970–2988.

Andersen AA. New Sampler for the collection, sizing, and enumeration of viable airborne particles. *J Bacteriol*. 1958;76: 471-484

Armstrong SM, Hargrave BT, Haya K. Antibiotic Use in Finfish Aquaculture: Modes of Action, Environmental Fate, and Microbial Resistance. 2005. *Handb Environ Chem 5 M*, pp 341–357.

Atamer Z, Dietrich J, Müller-Merbach M, *et al.* Screening for and characterization of Lactococcus lactis bacteriophages with high thermal resistance. *Int. Dairy J.* 2009; 19:228-235

Baker-Austin C, Wright MS, Stepanauskas R, *et al*. Co-selection of antibiotic and metal resistance. *Trends Microbiol*. 2006;14(4):176-82.

Barghigiani C, Ristori T, Scerbo R. Assessment of water pollution and suitability to fish life in six Italian rivers. *Environ Monit Assess*. 2001;66(2):187-205.

Bentley DR, Balasubramanian S, Swerdlow HP, *et al*. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature*. 2008; 456, 53–59.

Berendonk TU, Manaia CM, Merlin C, *et al*. Tackling antibiotic resistance: the environmental framework. *Nat Rev Microbiol*. 2015;13(5):310-7

Bhullar K, Waglechner N, Pawlowski A, *et al.,* Antibiotic resistance is prevalent in an isolated cave microbiome. *PLoS One*. 2012;7(4):e34953.

Bidawid S, Bosch A, Cook N *et al*. Editorial. *Food Environ Virol*. 2009; 1: 1–2.

Binelli A, Cogni D, Parolini M, *et al*. A Multi-biomarker approach to investigate the state of contamination of the R. Lambro/R. Po confluence (Italy) by zebra mussel (*Dreissena polymorpha*). *Chemosphere*. 2010. 79(5):518-28

Bushman F. Lateral DNA Transfer. Cold Spring Harbor Laboratory, NY, USA (2002).

Cabello FC, Godfrey HP, Tomova A, *et al.* Antimicrobial use in aquaculture re-examined: Its relevance to antimicrobial resistance and to animal and human health. *Environ Microbiol* 2013;15:1917–42.

Camusso M, Viganò L, Balestrini R. Bioconcentration of trace metals in rainbow trout: a field study. *Ecotoxicol Environ Saf*. 1995; 31(2):133-41.

Caporaso JG, Kuczynski J, Stombaugh J**,** *et al.* QIIME allows analysis of high- throughput community sequencing data Intensity normalization improves color calling in SOLiD sequencing. *Nat Methods* 2010; *7*(5), 335–336.

Charity RM, Foukas AF. Osteomyelitis and secondary septic arthritis caused by Sphingomonas paucimobilis. *Infect*. 2005; 33:93e95.

Cheeseman K, Ropars J, Renault P, *et al.,* Multiple recent horizontal transfers of a large genomic region in cheese making fungi. *Nat Commun*. 2014;5:2876

Chen H, Jogler M, Rohde M *et al.,* . *Sphingobium limneticum sp. nov*. and *Sphingobium boeckii sp. nov*., two freshwater planktonic members of the family *Sphingomonadaceae*, and reclassification of *Sphingomonas suberifaciens* as *Sphingobium suberifaciens comb. nov. Int J Syst Evol. Microbiol*. 2013; 63, 735–743.

Chen S, Li X, Sun G *et al*. Heavy Metal Induced Antibiotic Resistance in Bacterium LSJC7. *Int J Mol Sci.* 2015; 16(10).

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

Chiura HX. Generalized gene transfer by virus like particles from marine bacteria. *Aquat. Microb. Ecol*. 1997;13, 75–83.

Clemente JC, Pehrsson EC, Blaser MJ, *et al.* The microbiome of uncontacted Amerindians. *Sci Adv* 2015, e1500183.

Cogan TM, Hill C, Cheese Starter Cultures, in: Fox, P. F. (Ed.), Cheese: chemistry, Physics and Microbiology. Chapman & Hall, London, 1993, pp. 193–255.

Colombo S, Arioli S, Guglielmetti S, *et al*. Virome-associated antibiotic-resistance genes in an experimental aquaculture facility. *FEMS Microbiol Ecol* 2016; 92(3)

Colomer-Lluch M, Calero-Caceres W, Jebri S, *et al*. Antibiotic resistance genes in bacterial and bacteriophage fractions of Tunisian and Spanish wastewatersas markers to compare the antibiotic resistance patterns in each population. *Environ.Int*. 2014a; 73,167–175.

Colomer-Lluch M, Jofre J, and Muniesa M. Quinolone resistance genes (*qnrA* and *qnrS*) in bacteriophage particles from wastewater samples and the effect of inducing agents on packaged antibiotic resistance genes. *J.Antimicrob Chemother*; 2014b; 69,1265–1274.

Colomer-Lluch M, Imamovic L, Jofre J, *et al*. Bacteriophages carrying antibiotic resistance genes in fecal waste from cattle, pigs, and poultry. *Antimicrob Agents Chemother*. 2011b; 55,4908–4911.

Crottini A, Marotta R, Barbuto M, *et al.* The world in a river? A preliminary analysis of the *16S rDNA* variability of Tubifex species (Clitellata: Tubificidae) from the Lambro River. *Mol Phylogenet Evol* 2008; (3):1189-203.

Danovaro R, Dell'Anno A, Corinaldesi C *et al.* Virus-mediated archaeal hecatomb in the deep seafloor. *Sci. Adv*. 2016; 2(10) e1600492

D'Costa VM, McGrann KM, Hughes DW, *et al.* Sampling the antibiotic resistome. *Science.* 2006; 311:374-77.

D'Costa VM, King CE, Kalan L, *et al.* Antibiotic resistance is ancient. *Nature*. 2011;477, 457–461.

Defoirdt T, Boon N, Sorgeloos P, *et al.* Alternatives to antibiotics to control bacterial infections: luminescent vibriosis in aquaculture as an example. *Trends Biotechnol.* 2007;*25*:472–79.

Deksissa T, De Pauw D, Vanrolleghem PA. Dynamic in-stream fate modeling of xenobiotic organic compounds: a case study of linear alkylbenzene sulfonates in the Lambro River, Italy. *Environ Toxicol Chem* 2004; 23(9):2267-78.

DeSantis TZ, Hugenholtz P, Larsen N, *et al***.** Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microb.* 2006; *72*(7), 5069–5072.

Dubois E, LeGuyader F, Haugarreau L, *et al*. Molecular epidemiological survey of rotaviruses in sewage by reverse transcriptase seminested PCR and restriction fragment length polymorphism assay. *Appl Environ Microbiol*. 1997; 63: 1794–1800

Economou V and Gousia P. Agriculture and food animals as a source of antimicrobial-resistant bacteria. *Infect. Drug Resist.* 2015; 8, 49–61

Emond E and Moineau S. Bacteriophages and food fermentations. In: McGrath S, van Sinderen D, eds. Bacteriophage: Genetics and Molecular Biology. Horizon Scientific Press/Caister Academic Press 2007; 93-124.

Fancello L, Trape S, Robert C, *et al*. Viruses in the desert: a metagenomic survey of viral communities in four perennial ponds of the Mauritanian Sahara. *ISME J*. 2013; 7(2):359-69

FAO Fisheries and Aquaculture Department. 2013. Global Aquaculture Production Statistics for the year 2011

Farkas A, Erratico C, Viganò L. Assessment of the environmental significance of heavy metal pollution in surficial sediments of the River Po. *Chemosphere*. 2007;68(4):761-8

Flórez AB and Mayo B. Microbial diversity and succession during the manufacture and ripening of traditional, Spanish, blue-veined Cabrales cheese, as determined by PCR-DGGE. *Int J Food Microbiol*. 2006 Jul 15;110(2):165-71.

Galli P, Crosa G, Occhipinti Ambrogi A. Heavy metals concentrations in acanthocephalans parasites compared to their fish host. *Chemosphere*. 1998;37(14-15):2983-8.

Giari L, Dezfuli BS, Lanzoni M, *et al*. The impact of an oil spill on organs of bream *Abramis brama* in the Po River. *Ecotox environ safe* 2012; 77:18-27.

Grave K, Lingaas E, Bangen M, *et al.* Surveillance of the overall consumption of antibacterial drugs in humans, domestic animals and farmed fish in Norway in 1992 and 1996. *J Antimicrob Chemother* 1999;43:243–52.

Guglielmetti S, Fracassetti D, Taverniti V, *et al*. Differential modulation of human intestinal Bifidobacterium populations after consumption of a wild blueberry (Vaccinium angustifolium) drink. *J. Agric. Food Chem.* 2013; 61, 8134−8140.

Hallen HK, Donato J, Wang HH, *et al.,* Call of the wild: antibiotic resistance genes in natural environments. *Nat. Rev. Microbiol*. 2010;8:251-259.

Hammami R, Zouhir A, Hamida JB, *et al*. BACTIBASE: a new web-accessible database for bacteriocin characterization. *BMC Microbiol.* 2007; 7:89.

Haque MM, Bose T, Dutta A, *et al*. CS-SCORE: Rapid identification and removal of human genome contaminants from metagenomic datasets. *Genomics* 2015; 106(2):116-21.

Herman KM, Hall AJ, Gould LH. Outbreaks attributed to fresh leafy vegetables, United States, 1973-2012. *Epidemiol Infect*. 2015 Oct;143(14):3011-21

Hyatt D, Chen GL, Locascio PF, *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 2010; *11*, 119.

Hu Y, Yang X, Qin J, *et al.* Metagenome-wide analysis of antibiotic resistance genes in a large cohort of human gut microbiota. *Nat Commun* 2013;4:2151.

Kang, YJ and Frank FJ. Characteristics of biological aerosols in dairy processing plants. *J. Dairy Sci*., 1990;73:621-626

Kallman O, Lundberg C, Wretlind B, *et al.* Gram negative bacteria from patients seeking medical advice in Stockholm after the tsunami catastrophe. *Scand J Infect Dis* 2006;38:448–50

Kenzaka T, Tani TK, Nasu M. High frequency phage mediated gene transfer in freshwater environments determined at single cell level. *ISME J*. 2010; 4, 648–659.

Kilic A, Senses Z, Kurekci AE, *et al.* Nosocomial outbreak of *Sphingomonas paucimobilis* bacteremia in a hemato/oncology unit. *Jpn J Infect Dis* 2007;60:394-96

Knapp CW, McCluskey SM, Singh BK, *et al.* Antibiotic resistance gene abundances correlate with metal and geochemical conditions in archived Scottish soils. *PLoS One*. 2011;6(11):e27300.

Levy SB and Marshall B. Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med*. 2004 Dec;10(12 Suppl):S122-9.

Lin JN, Lai CH, Chen YH, *et al. Sphingomonas paucimobilis* bacteremia in humans: 16 case reports and a literature review. *J Microbiol Immunol Infect* 2010;43: 35–42.

López-Bueno A, Tamames J, Velázquez D, *et al*. High Diversity of the Viral Community from an Antarctic Lake. *Science*. 2009; 326(5954), pp. 858-861

Macinga DR, Rather PN. The chromosomal 2′-N-acetyltransferase of *Providencia stuartii*: physiological functions and genetic regulation. *Front. Biosci. J. Virtual Libr*. 1999;4, D132–D140

Madera C, Monjardin C, Suarez JE. Milk contamination and resistance to processing conditions determine the fate of Lactococcus lactis bacteriophages in dairies. *Appl. Environ. Microbiol.* 2004; 70:7365-7371

Mann BA, Slauch JM. Transduction of low copy numbers of plasmids by bacteriophages P22. *Genetics*. 1997; 146, 447–456.

Marcó MB, Moineau S, Quiberoni A. Bacteriophages and dairy fermentations. *Bacteriophage*. 2012 Jul 1; 2(3): 149–158

Margulies M, Egholm M, Altman WE, *et al*. Genome sequencing in microfabricated high-density picolitre reactors. *Nature*. 2005; 437, 376–380

Marti E, Variatza E, Balcazar JL. Bacteriophages as are servoir of extended-spectrum b-lactamase and fluoroquinolone resistance genes in the environment. *Clin Microbiol Infect*.2014; 20, O456–O459.

Martinez JL. Antibiotics and antibiotic resistance genes in natural environments. *Science.* 2008; 321, 365–367

Martinez JL, Coque TM, Baquero F. Waht is a resistant gene? Ranking risk in resistome. *Nat Rev Microbiol.* 2015;13:116-123

Mazaheri Nezhad Fard R, Barton MD, Heuzenroeder MW. Bacteriophage-mediated transduction of antibiotic resistance inenterococci. *Lett.Appl.Microbiol*. 2011; 52,559–564.

McArthur AG, Waglechner N, Nizam F, *et al*. The Comprehensive Antibiotic Resistance Database. *Antimicrob Agents Chemother* 2013; 57, 3348-3357

Mc Grath S, Fitzgerald GF,Van Sinderen D. Bacteriophages in dairy products: Pros and cons. *Biotechnol J*. 2007 Apr;2(4):450-5

McIntyre K, Heap HA, Davey GP, and G. K. Y. Limsowtin. The distribution of lactococcal bacteriophage in the environment of a cheese manufacturing plant. *Int. Dairy J*. 1991; 1:183-197

Milani C, Hevia A, Foroni E, *et al*. Assessing the Fecal Microbiota: An Optimized Ion Torrent *16S rRNA* Gene-Based Analysis Protocol. *PLoS ONE.* 2013; 8:e68739.

Minot S, Sinha R, Chen J *et al*. The human gut viromes: inter-individual variation and dynamic response to diet. *Genome Res*. 2011. 21, 1616–1625

Modi SL, Lee HH, Spina CS, *et al.* Antibiotic tretment expands the resistance reservoir and ecological network of the phage metagenome. *Nature.* 2013;499:219-23.

Morar M, Bhullar K, Hughes DW, *et al.* Structure and mechanism of the lincosamide antibiotic adenylyltransferase LinB. *Structure* 2009, 17:1649-1659.

Muniesa M, Colomer-Lluch M, Jofre J. Could bacteriophages transfer antibiotic resistance genes from environmental bacteria to human-body associated bacterial populations? *Mob Genet Elements.* 2013a; 3(4): e25847.

Muniesa M, Colomer-Lluch M, Jofre J. Potential impact of environmental bacteriophages in spreading antibiotic resistance genes. Future Microbiol. 2013B; 8(6):739-51

Muscillo M, La Rosa G, Marianelli C, *et al*. A new RT-PCR method for the identification of reoviruses in seawater samples. *Water Res*. 2001; 35: 548–556.

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

Pal C, Bengtsson-Palme J, Rensing C, *et al*. BacMet: antibacterial biocide and metal resistance genes database. *Nucl Acids Res* 2014; 42 (D1): D737-D743.

Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. *Microbiol Mol Biol Rev* 1998;62:1–34

Parsley LC, Consuegra EJ, Kakirde KS *et al*. Identification of diverse antimicrobial resistance determinants carried on bacterial, plasmid, or viral metagenomes from an activated sludge microbial assemblage. *Appl. Environ. Microbiol*. 2010. 76, 3753–3757.

Penders J, and Stobberingh EE. Antibiotic resistance of motile aeromonads in indoor catfish and eel farms in the southern part of The Netherlands. *Int. J. Antimicrob. Agents*. 2008; 31, 261–265

Perry JA, Westman EL, Wright GD. The antibiotic resistome: what's new? *Curr Opin Microbiol*. 2014a Oct;21:45-50

Perry JA and Wright GD. Forces shaping the antibiotic resistome. *Bioessays*. 2014b Dec;36(12):1179-84

Perron GG, Whyte L, Turnbaugh PJ, *et al*. Functional characterization of bacteria isolated from ancient arctic soil exposes diverse resistance mechanisms to modern antibiotics. *PLoS ONE*. 2015;10:e0069533

Pettine M, Bianchi M, Martinetti W, *et al*. Contribution of the Lambro River to the total pollutant transport in the Po watershed (Italy). Sci Total Environ 1996; 192, 275–297

Ricard G, McEwan NR, Dutilh BE, *et al*. Horizontal gene transfer from Bacteria to rumen Ciliates indicates adaptation to their anaerobic, carbohydrates-rich environment. *BMC Genomics*. 2006 Feb 10;7:22.

Rodriguez-Brito B, Li L, Wegley L, *et al*. Viral and microbial community dynamics in four aquatic environments. *ISME J*. 2010;4:739–51

Rodríguez-Lázaro D, Cook N, Ruggeri FM, *et al*. Virus hazards from food, water and other contaminated environments. *FEMS Microbiol Rev*. 2012 Jul;36(4):786-814

Rosario K, Nilsson C, Lim YW, *et al*. Metagenomic analysis of viruses in reclaimed water. *Environ Microbiol*. 2009;11:2806–20.

Roux S, Enault F, Robin A, *et al*. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE.* 2012; e33641

Rzeżutka A & Cook N. Survival of human enteric viruses in the environment and food. *FEMS Microbiol Rev*. 2004; 28: 441–453

Salustiano VC, Andrade NJ, Brandao SCC *et al*. 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*. 2003; 34: 255-259

Sano E, Carlson S, Wegley L *et al*. Movement of viruses between biomes. *Appl. Environ. Microbiol*. 2004; 70, 5842–5846.

Scanlan PD, Buckling A. Co-evolution with lytic phages selected for the mucoid phenotype of Pseudomonas fluorescens SBW25. *ISME J*. 2012; 6, 1148–1158.

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

Smillie CS, Smith MB, Friedman J, *et al*. Ecology drives a global network of gene exchange connecting the human microbiome. *Nature*. 2011;480: 241–4.

Suttle CA. The significance of viruses to mortality in aquatic microbial communities. *Microb. Ecol*. 1994; 28, 237–243.

Taylor NG, Verner-Jeffreys DW, Baker-Austin C. Aquatic systems: maintaining, mixing and mobilising antimicrobial resistance? *Trends Ecol Evol*. 2011; 26(6):278-84.

Tchounwou PB, Yedjou CG, Patlolla AK *et al*. Heavy Metals Toxicity and the Environment. *EXS* 2012; 101: 133–164.

Teixeira P, Tacão M, Alves A. Antibiotic and metal resistance in a ST395 Pseudomonas aeruginosa environmental isolate: A genomics approach. *Mar Pollut Bull* 2016; S0025-326X(16)30490-8

Thingstad TF. Elements of a theory for the mechanisms controlling abundance, diversity, and biogeochemical role of lytic bacterial viruses in aquatic systems. *Limnol Oceanogr.* 2000; 45:1320-28.

Thomas CM and Nielsen KM. Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat.Rev.Microbiol*. 2005;3,711–721.

Thurber RV, Haynes M, Breitbart M, *et al*. Laboratory procedures to generate viral metagenomes. *Nat Protoc* 2009; 4:470–83.

Tseng CH, Chiang PW, Shiah FK, *et al*. Microbial and viral metagenomes of a subtropical freshwater reservoir subject to climatic disturbances. *ISME J*. 2013;7:2374–86.

Urbatzka R, van Cauwenberge A, Maggioni S, *et al*. Androgenic and antiandrogenic activities in water and sediment samples from the river Lambro, Italy, detected by yeast androgen screen and chemical analyses. *Chemosphere* 2007; 67(6):1080-7.

Vaz-Moreira I, Nunes OC, Manaia CM. Diversity and antibiotic resistance patterns of Sphingomonadaceae isolates from drinking water. *Appl Environ Microbiol* . 2011;77:5697–706

Viganò L, Barbiero G, Buffagni A, Mingazzini M, Pagnotta R. Assessment of the alterations of the aquatic environment downstream a polluted tributary of the river Po (Italy). *Aquat Ecosyst Health Manag*. 1999; 2, 55–69

Viganò L, Patrolecco L, Polesello S, *et al*. Interactions between trophic and toxic factors in a polluted urban river. *Ecotoxicol Environ Saf*. 2008; 69(1):49-57.

Viganò L, Arillo A, Bottero S, *et al*. First observation of intersex cyprinids in the Po River (Italy). *Sci Total Environ.* 2001; 269, 189–194

Viganò L, Farkas A, Guzzella L, *et al*. The accumulation levels of PAHs, PCBs and DDTs are related in an inverse way to the size of a benthic amphipod (Echinogammarus stammeri Karaman) in the River Po. *Sci Total Environ*. 2007; 373(1):131-45.

Van Boeckel TP, Brower C, Gilbert M, *et al*. Global trends in antimicrobial use in food animals. *Proc. Natl. Acad. Sci. U.S.A*. 2015;112, 5649–5654

Van Heijnsbergen E, Schalk JA, Euser SM *et al.,* Confirmed and Potential Sources of Legionella Reviewed. *Environ Sci Technol*. 2015 Apr 21;49(8):4797-815

Van Schaik W. The human gut resistome. *Philos Trans R Soc Lond B Biol Sci*. 2015 Jun 5; 370(1670): 20140087.

Von Wintersdorff CJ, Penders J, van Niekerk JM *et al*. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. *Front Microbiol*. 2016 Feb 19;7:173.

Wang S, Shi X. Molecular mechanisms of metal toxicity and carcinogenesis. *Mol Cell Biochem*. 2001;222(1-2):3-9

White RA, Callister SJ, Moore RJ, *et al*. The past, present and future of microbiome analyses. *Nat Prot*. 2016; 11,2049–2053

Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS Microbiol Rev*. 2011; 35: 957–76

Wright, G. D., and Poinar, H. Antibiotic resistance is ancient: implications for drug discovery. *Trends Microbiol*. 2012;20, 157–159

Zablocki O, Van Zyl L, Adriaenssens EM, *et al*. High-Level Diversity of Tailed Phages, Eukaryote-Associated Viruses, and Virophage-Like Elements in the Metaviromes of Antarctic Soils. Appl Environ Microb 2014;80:6888–897.