



UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA
Dipartimento di Scienze dell'Ambiente e del Territorio e di
Scienze della Terra
Dottorato in Biologia (XXVI Ciclo)



***Urban Airborne Microorganisms:
Biodiversity, Variability and
Community-Assembly Processes***

TUTOR *Prof. Paolo Tortora*
Settore Scientifico-Disciplinare BIO19

TESI DI DOTTORATO
Valentina Bertolini
Matricola n° 042350

To my granny,
to my parents,
to my husband and
to my little baby,
gratefully.

Index

| | |
|---|-----------|
| Index | II |
| 1. Introduction | 1 |
| 1.1. Ecology of microorganisms | 2 |
| 1.1.1 From macro-ecology to micro-ecology | 2 |
| 1.1.2 Microbial communities | 3 |
| 1.1.3 Community assembly processes | 3 |
| 1.1.4 Microbial biodiversity | 5 |
| 1.2. Methods in microbial ecology | 6 |
| 1.3. The Earth atmosphere | 7 |
| 1.4. Atmospheric Suspended Particulates | 9 |
| 1.4.1. Bioaerosols | 12 |
| 1.5. Aim of the project | 12 |
| 1.6. Summary of the thesis | 13 |
| 1.7. References | 17 |
| 2. Antibiotic resistance in bacteria associated with coarse atmospheric particulate matter in an urban area | 22 |
| 2.1. Introduction | 23 |
| 2.2. Materials and Methods | 25 |
| 2.2.1 Sampling and isolation of microbial strains | 25 |
| 2.2.2 Disc diffusion test | 26 |
| 2.2.3 Identification of <i>Staphylococcus aureus</i> strains | 27 |
| 2.2.4 Detection of antibiotic resistance genes | 27 |
| 2.3. Results | 28 |
| 2.3.1 Isolation of microbial strains | 28 |
| 2.3.2 Disc diffusion test | 30 |
| 2.3.3 Identification of <i>Staphylococcus aureus</i> strains | 33 |
| 2.3.4 Antibiotic resistance genes | 35 |
| 2.4. Discussion | 36 |
| 2.5. Acknowledgements | 38 |
| 2.6. References | 38 |
| 3. Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy | 43 |
| 3.1. Introduction | 44 |
| 3.2. Material and methods | 46 |
| 3.2.1 Sample collection and DNA extraction | 46 |
| 3.2.2 Quantitative PCR (qPCR) | 47 |

| | | |
|-----------|--|-----------|
| 3.2.3 | 16S rRNA fragment libraries by Illumina Genome Analyzer Iix | 48 |
| 3.2.4 | Sequence analyses | 48 |
| 3.2.5 | Environmental factors | 49 |
| | 3.2.5.1 Meteorological data and Particulate Matter (PM) concentration | 49 |
| 3.2.6 | Analysis of Inorganic Ions | 49 |
| 3.2.7 | Statistical Analysis | 50 |
| 3.2.8 | Sequence Data | 51 |
| 3.3. | Results | 52 |
| 3.3.1 | Quantification of airborne bacteria | 52 |
| 3.3.2 | Microbial community structure | 52 |
| 3.3.3 | Short-term temporal variation of community structure | 59 |
| 3.4. | Discussion | 59 |
| 3.4.1 | Seasonal variability of airborne bacterial abundance | 59 |
| 3.4.2 | Seasonal and short-term variability of community structure | 60 |
| 3.4.3 | Potential sources of bacteria and differences in the structure of bacterial communities associated with different PM size fractions | 62 |
| 3.5. | Acknowledgements | 63 |
| 3.6. | References | 64 |
| 4. | Spatio-temporal variability and community-assembly processes of PM10 associated bacterial communities across two cities of Northern Italy | 71 |
| 4.1. | Introduction | 72 |
| 4.2. | Material and methods | 74 |
| 4.2.1 | Sample collection and DNA extraction | 74 |
| 4.2.2 | Quantitative PCR (qPCR) | 75 |
| 4.2.3 | 16S rRNA fragment libraries by Illumina Hiseq 1000 | 76 |
| 4.2.4 | Sequence analyses | 77 |
| 4.2.5 | Environmental factors and chemical data | 77 |
| | 4.2.5.1 Meteorological data and Particulate Matter (PM) concentration | 77 |
| | 4.2.5.2 Analysis of Inorganic Ions | 77 |
| | 4.2.5.3 Analysis of elemental elements | 78 |
| 4.2.6 | Statistical analyses | 78 |
| | 4.2.6.1 Missing value imputation | 78 |
| | 4.2.6.2 Univariate analyses | 79 |
| | 4.2.6.3 Cluster analysis | 79 |
| | 4.2.6.4 Ordination techniques | 80 |
| 4.3 | Results | 81 |
| 4.3.1 | Quantification of airborne bacteria | 81 |
| 4.3.2 | Microbial community structures | 82 |
| | 4.3.2.1 Cluster Analysis | 83 |

| | | |
|----------|--|------------|
| 4.3.2.2 | Canonical Analysis and Principal Components Analysis | 84 |
| 4.3.2.3 | Constrained ordination | 85 |
| 4.3.2.4 | Variation partitioning analysis | 93 |
| 4.3.2.5 | Multivariate Regression Trees (MRT) | 93 |
| 4.3.2.6 | Indicator taxa | 94 |
| 4.4 | Discussion | 96 |
| 4.4.1 | Quantification of airborne bacteria | 96 |
| 4.4.2 | Variability of microbial community structure and community-assembly processes | 97 |
| 4.4.2.1 | Dispersal | 98 |
| 4.4.2.2 | Selection | 100 |
| 4.5 | Conclusions | 102 |
| 4.6 | References | 102 |
| 5 | Unraveling the bacterial diversity in the atmosphere | 110 |
| 5.1 | Introduction | 111 |
| 5.2 | Sampling methods | 112 |
| 5.3 | Description of airborne microorganisms | 114 |
| 5.3.1 | Quantification of airborne microorganisms | 114 |
| 5.3.2 | Cultivation based surveys | 116 |
| 5.3.3 | Molecular characterization of microbial communities | 118 |
| 5.3.3.1 | Evaluation of biodiversity of airborne microbial communities | 120 |
| 5.3.3.2 | Temporal and spatial variability in the structure of airborne microbial communities | 123 |
| 5.3.3.3 | Influence of environmental and meteorological factors on airborne bacterial communities | 124 |
| 5.3.3.4 | Identification of potential sources of airborne Bacteria | 125 |
| 5.4 | Future perspectives | 126 |
| 5.5 | Acknowledgements | 128 |
| 5.6 | References | 128 |
| 6 | Bacterial community structure on two alpine debris covered glaciers and biogeography of <i>Polaromonas</i> phylotypes | 138 |
| 6.1 | Introduction | 139 |
| 6.2 | Material and methods | 141 |
| 6.2.1 | Study sites and environmental data | 141 |
| 6.2.2 | Debris sampling and data extraction | 142 |
| 6.2.3 | 16S rRNA fragment sequencing | 143 |

| | | |
|----------|---|------------|
| 6.2.4 | Sequence processing and data analysis | 143 |
| 6.2.5 | Statistical methods | 144 |
| 6.2.6 | Data accessibility | 146 |
| 6.3 | Results | 146 |
| 6.3.1 | Structure of microbial communities | 147 |
| 6.3.2 | Biogeography of <i>Polaromonas</i> genus | 150 |
| 6.4 | Discussion | 152 |
| 6.5 | Acknowledgements | 157 |
| 6.6 | References | 157 |
| 7 | Perspectives | 167 |
| 7.1 | Future perspectives | 168 |
| 7.2 | References | 169 |
| 8 | Appendix | 171 |
| 8.1 | Supporting information of Chapter 2 | 172 |
| 8.2 | Supporting Information of Chapter 3 | 173 |
| 8.2.1 | Details on statistical methods | 173 |
| 8.2.1.1 | Missing PM2.5 data | 173 |
| 8.2.1.2 | ANOSIM | 174 |
| 8.2.1.3 | Short-term variation in bacterial community composition | 174 |
| 8.2.2 | Reference | 175 |
| 8.2.3 | Tables | 175 |
| 8.3 | Supporting Information of Chapter 4 | 180 |
| 8.4 | Supporting Information of Chapter 6 | 181 |
| 8.4.1 | Further information on glaciological and environmental features of the miage and belvedere glaciers | 181 |
| 8.4.1.1 | Miage glacier | 181 |
| 8.4.1.2 | Belvedere glacier | 183 |
| 8.4.1.3 | Comparison of ecological conditions at sampling sites on Miage and Belvedere glaciers | 184 |
| 8.4.2 | Further information on sampling methods | 184 |
| 8.4.3 | Further information on molecular analyses | 185 |
| 8.4.4 | Further details on statistical analyses | 186 |
| 8.4.5 | References | 187 |
| 8.4.6 | Tables | 189 |
| 8.4.7 | Figures | 196 |
| 9 | Ringraziamenti | 199 |

CHAPTER 1

Introduction

1.1 Ecology of microorganisms

1.1.1 From macro-ecology to micro-ecology

Although in theory, it is possible to suppose a continuity of life that encompasses all organisms, most biologists accept a fundamental divide between the small and simple organisms (microbes, predominantly prokaryotes and viruses) and the larger and more complex ones (macroorganisms, mainly eukaryotes). Although evolutionary theories provide a unifying framework for all biology, it is compulsory to include some major differences between the evolutionary processes of unicellular and pluricellular organisms.

Lateral gene transfer is one of these different processes, as is speciation and the real concept of what a species is in the kingdoms of prokaryotes. In the context of speciation, geographical isolation is believed to play fundamentally different roles for macrobes and microbes (O'Malley, 2008).

Biogeography is the descriptive and explanatory study of spatial patterns and processes in the distribution of biodiversity. However, the application of macrobiogeographical rules to microorganisms is not always possible and profitable. For example, the passive dispersal (e.g. via wind) of microbes is likely to be easier than for macrobes, simply because they are smaller. While active dispersal is rare for microbes. Consequently, the dispersal potential of microorganisms is probably larger than that of macroorganisms (Darcy et al, 2011).

Likewise, microorganisms can easily become dormant i.e. to enter in a reversible state of reduced activity in response to environmental stresses (Jones et al, 2010). It has been estimate that less than 10% of a typical microbial community may be active at the same time; thus the dormant component potentially represents a reservoir of genetic diversity (Nemergut et al, 2013).

In conclusion, one of the most important elements for future research will be understanding similarities and differences between microbial and microbial

biogeography and whether they can coexist within one general account (Martiny et al, 2006; Whitaker, 2006).

1.1.2 Microbial communities

In such a complex contest, the scientific community over years has introduced many different and sometimes ambiguous terminologies and definitions and there is still an animated debate on this theme. Hence, in the effort to clarify any ambiguity, at least in the frame of this thesis, here below there is the definition of microbial community that could represent a shared starting point.

We define a community as a group of potentially interacting organisms that co-occur in space and time. In the case of microbiology, communities are not discrete and their boundaries may vary both temporally and spatially, especially because of the inability to sample at a level that is both small enough to be relevant for microorganisms and large enough to be relevant for ecosystem processes. Thus, we refer as a microbial community a microbial assemblage of defined microorganisms that co-occur in space and time, without necessarily implying the existence of dynamics and interactions between populations (Nemergut et al, 2013).

1.1.3 Community assembly processes

An exhaustive study of a microbial community requires understanding the mechanisms of assembly and how they produce pattern in both space and time. On the basis of the Vellend's approach, four different processes have been identified, which drive the assembly of a community: selection, dispersal, diversification and drift (Vellend, 2010; Nemergut et al, 2013). Here below, the description of these four processes as described by Nemergut and colleagues (2013).

Selection

Selection is a large force shaping microbial community assembly and it is defined as "deterministic fitness differences between individuals". Different habitat (e.g. seawater, soils) harbor different suites of microorganisms and a variety of

Introduction

environmental factors have a critical role in determining bacterial assemblage structure and diversity.

Although selection should be similar in macrobial and microbial systems, the specifics of these processes will be different due to the vast diversity harbored within microbial communities and even within individual organisms. Thus, the complexities of the potential environmental and biological drivers of fitness are greatly magnified for these communities. Indeed, the metabolic breadth of microorganisms has been hypothesized to be a key factor in the generation and maintenance of microbial diversity. Finally, the prominent role of horizontal gene transfer (and recombination in general) in microbial diversification may affect microbial community assembly through selection.

Dispersal

The definition of dispersal is the “movement of organisms across space”. Because of the small size, high abundance and short generation time of microorganisms, dispersal processes have not been rigorously studied, much less quantified. Thus, the distributions of microbes are often used as proxies for dispersal. Given that it is difficult to understand the role of dispersal processes in community assembly. Dispersal is different from migration, in which a new organism is incorporated into a community from outside. Migration events are the result of dispersal as well as selection and possibly drift.

Microbial dispersal is typically a passive process (e.g. transport via air, water or mobile macroorganisms). Although it is often considered stochastic, it is not entirely true: taxa vary in dispersal ability, making dispersal probability not entirely random among species. For example, dormancy mechanisms may make organisms more resistant to the environmental stressors.

Diversification

The process of evolutionary diversification is fundamental to biogeography and community ecology, but unfortunately, we understand little about the spatial and temporal dynamics of how microbes evolve and these processes are difficult to study

empirically. The process of dormancy is likely to affect microbial evolutionary because of it can protect cells from death, at least until favorable conditions or stochastic factors lead to their growth. This could result in dramatic variation in evolutionary rates over time. The process of horizontal gene transfer may also change community assembly dynamics for microorganisms because of they can evolve through mutation rapidly.

Drift

Ecological drift, or stochastic changes in the relative abundance of organisms, may play an important role in microbial community assembly, especially when selection is weak, alpha diversity is low and the total number of community members is small. These conditions can be found in certain types of microbial communities, for example nutrient-rich systems such as wastewater treatment facilities. The vast majority of taxa in microbial communities are found in low relative abundances. However, low-abundance individuals may exist in dormant states, protecting them from extinction. A better understanding of the dynamics between dormancy and local extinctions is crucial to an appreciation of the role of drift in community assembly, particularly for microbial assemblages.

1.1.4 Microbial biodiversity

The term biodiversity indicates the biological diversity of an environment. Generally, there are two classes of biodiversity: inventory diversity or differentiation diversity. Inventory diversity metrics describe diversity within an environment (alpha-diversity), while differentiation diversity explains the turnover in diversity between environments (beta-diversity) (Lozupone and Knight, 2008). However, the terminology and metrics used in macro-ecology must be applied cautiously on micro-ecology because the methods available to characterize microbial diversity are inherently limited in their ability to detect the many numerically minor constituents of microbial communities (Bent and Forney, 2008). Exploring microbial diversity means to take in account of different peculiar characteristics typical of the sole

microorganisms (for example, the definition of the boundary of a single environment, or the capability to capture the actual biodiversity of a microbial community). Microorganisms constitute a special case of study, because of their intrinsic peculiarities (small dimensions, high rate of duplication, high dispersal capacity, etc.). The actual microbial diversity of most habitats seems to be extraordinarily big, with complex habitats containing an estimated 10^4 – 10^6 species in a single gram (Torsvik et al, 1998; Ovreas et al, 2003) and the Earth's biosphere containing more than 10^{30} individuals (Whitman et al, 1998) and an untold number of species. Hence, the first necessary step is to focus on the methods used to extensively collect samples and to generate the data for accurate biodiversity estimation.

1.2 Methods in microbial ecology

The methodologies of study of microbial communities represent not merely a technical factor for this research field, but mainly a critical issue for an exhaustive comprehension of this scientific subject. Moreover, the variety of methods used so far, represents an obstacle for data uniformity and comparability.

The small dimension, short generation time and metabolic versatility of microorganisms limit the use of observation-based surveys for the description of microbial populations and for coupling their taxonomy and ecological function (Polz et al 2006; Schmidt et al, 2007).

For a long time, cultivation based methods have been the gold standard to describe microorganisms. As indicated by the name, these methods imply point sampling, laboratory microorganism pure culture isolation and subsequent analysis of the isolates. Inevitably, this procedure permits the analysis of a limited number of isolates, not more than a few hundreds. Therefore, these methodologies heavily underestimate both the number and the diversity of microorganisms present in a given environment, because it is well known that less than 1% of total microorganism can be efficiently cultivated (see chapter 5).

Introduction

The development of cultivation independent methods PCR-based has expanded the possibility to detect and identify microorganisms unambiguously and avoids the cultivation biases linked to laboratory cultivability (Peccia et al, 2006). Using the 16S rRNA gene as a phylogenetic marker, it is possible to identify single microorganisms, make a community study or quantify the total number of bacteria. The common methodological framework implies PCR reaction with suitable primer sets, clone library construction and DNA fingerprinting analyses (for the community study). The advantages of these methods include large sensitivity and rapidity. However, the quantity of data retrievable with this approach reaches at maximum hundreds of clones. The representative of this analysis is still too limited to give a thorough description of a microbial community.

In this respect, the recent development of Next Generation Sequencing (NGS) technologies greatly improves the capability to describe microbial communities. In recent years, sequencing costs have rapidly decline and consequently the amount of available data has increased exponentially. Owing to their high throughput and the reducing cost per sequence, the NGS techniques have a great potential in describing the diversity and composition of microbial communities (Bowers et al, 2011a; Rinsoz et al, 2008). For example the Illumina technology can generate up to millions of amplicons in single run, thus providing high coverage both to amplicon-based and whole metagenomic studies of microbial communities and contributing to fill the gaps in the current knowledge of microbial community structure and dynamics (Claesson et al, 2010).

Computational advances, as well as new standards for contextualizing environmental microbial community composition data sets, will allow us to make the most of these data, facilitating cross-investigator and cross-system meta-analyses (Caporaso et al, 2010).

1.3 The Earth atmosphere

Introduction

The Earth planet is characterized by the presence of an atmosphere distinctive from other planetary atmospheres. This peculiar atmosphere makes life possible on the Earth. The term atmosphere derives from the Greek words *ατμός* (*atmos*) and *σφαίρα* (*sphaira*) that mean vapor and sphere. It represents the layer of gases surrounding the terrestrial crust and is held in place by the Earth gravity. The main functions of the atmosphere are the absorption of ultraviolet solar radiation, the warming of the surface through heat retention and the reduction of temperature extremes between day and night.

Although the atmosphere extends outward at least 10 000 kilometers (Km), most of its mass is close to Earth surface, concentrated at low altitude. More than half of the atmosphere mass lies below 6 Km and more than 98% of it lies within 26 Km of the sea level.

Observing its vertical extension, the atmosphere shows a vertical pattern of temperature that forms a series of layers in which temperature increase and decrease alternatively: troposphere, stratosphere, mesosphere, thermosphere and exosphere (starting from the Earth surface up to the planet boundary). The boundaries of the first three layers are called tropopause, stratopause and mesopause. The troposphere is characterized by an overturning of the air resulting in a vertical mixing and turbulence. It is extended about 8 Km above sea level over the poles and up to about 18 Km above the sea level at the equator.

The chemical composition of the atmosphere is relatively simple and uniform; the major components do not vary over time and space. The two most abundant gases are nitrogen (N₂) (78.09%) and oxygen (O₂) (20.95%) that constitute together about the 99% of atmosphere. The remaining 1% is composed of argon (0.93%) and small amounts of other gases often referred as trace gases (i.e. neon, helium, krypton and hydrogen). The atmosphere also contains a variable amount of other gases that can have a significant effect on weather and climate. Water vapor, for instance, highly varies in quantity in different locations (i.e. deserts, tropical oceans, etc.) and it is the source of all precipitations and clouds. Like water vapor, also Carbon Dioxide (CO₂)

has a considerable influence on climate because it is able to adsorb infrared radiation and thereby help warm the lower atmosphere. CO₂ is homogeneously distributed in the lower atmospheric layers, but its concentration has been rapidly increasing in this last years; this increase can cause a consistent warming of the planet. Another important gas is Ozone (O₃). It is mainly present between 15 and 48 Km from Earth's surface (a layer sometimes called the Ozonosphere) and it is an excellent absorber of ultraviolet solar radiation; this characteristic protects Earth life forms from deadly effects of radiation (Hess, 2010). Other minor gaseous components of anthropogenic origin are present in the atmosphere. Among these contaminants the most relevant are carbon, sulfur and nitrogen compounds. For instance, Carbon Monoxide (CO), a primary pollutant formed by the incomplete combustion of carbon-based fuels, which is extremely toxic for brain and organs; Sulfur Dioxide (SO₂), a corrosive and lung irritant *per se*, but also able to react in the atmosphere to produce secondary pollutants and finally Nitric oxide (NO), a gas that can react in the atmosphere to form Nitrogen Dioxide (NO₂), a corrosive gas.

1.4 Atmospheric Suspended Particulates

Atmospheric particulates, also known as aerosols or particulate matter (PM), are tiny pieces of solid or liquid particles suspended in the atmosphere as aerosol. Sources of particulate matter can be anthropogenic or natural. They can adversely affect human health and also have impacts on climate and precipitation. PM can be classified in many ways, but one of the most used is the classification on the basis of the particle size. First, it is possible to distinguish "coarse" and "fine" particles: respectively particles with aerodynamic diameter less than 10 μm (PM₁₀) and less than 2.5 μm (PM_{2.5}). Then, there are some "ultrafine" particles with aerodynamic diameter less than 1 μm (PM₁) and it is possible to find particles smaller referred as "nano-particles" (Figure 1).

Introduction

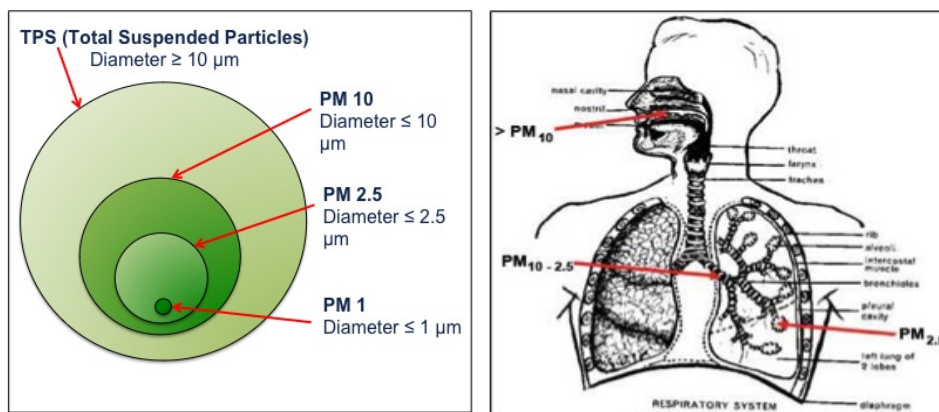


Figure 1 Size particles and relative penetration capability into the respiratory system.

This classification is particularly relevant because smaller particles are able to deeper penetrate into the respiratory system and to progressively provoke more serious damages. In particular, course PM are able to reach the pharynx, while fine PM can reach the bronchial tubes and the ultrafine the pulmonary alveoli.

The chemical composition, quantity and relative abundance of the principal chemical species of PM can vary depending on location, season and analyzed fraction. However, it is possible to identify a general trend of a peculiar area (Perrone et al, 2010). For example, in Milan urban area the main components are shown in figure 2 (Mantecca et al, 2012).

Combustion-related primary compounds, such as Polycyclic Aromatic Hydrocarbons (PAHs), were higher in fine winter fractions, while summer PMs were enriched in secondary organic aerosols and in crustal elements such as Fe, Al and Si. Similar results were reported for other European cities, such as Dresden, Germany (Brüggemann et al, 2009) and are explained by the presence of different sources, like the domestic heating in winter and the meteorological condition, like a lower atmospheric mixing layer in winter and a drier weather in summer, this last factor determining a higher resuspension of crustal elements. The concentrations of the PM in urban areas are normally higher than in areas with less anthropogenic pressure due

Introduction

to higher emission and low air dispersion. Meteorological factors such as the horizontal wind (speed and direction), the atmospheric stability, the altitude, the thermal inversion and the atmospheric vertical motions due to the orographic systems also affect the air pollution. In general the most significant pollution events occur under thermal inversion conditions. For instance, the Po Valley, in the North of Italy, represents one of the most PM polluted area in Italy, as well as in the European Union. Together with the high population density and the large number of industrial activities spread throughout the valley, the peculiar geographic position, enclosed between the Alps and the Appennines, determines the accumulation of gaseous and particulate pollutants especially during wintertime (Mantecca et al, 2012).

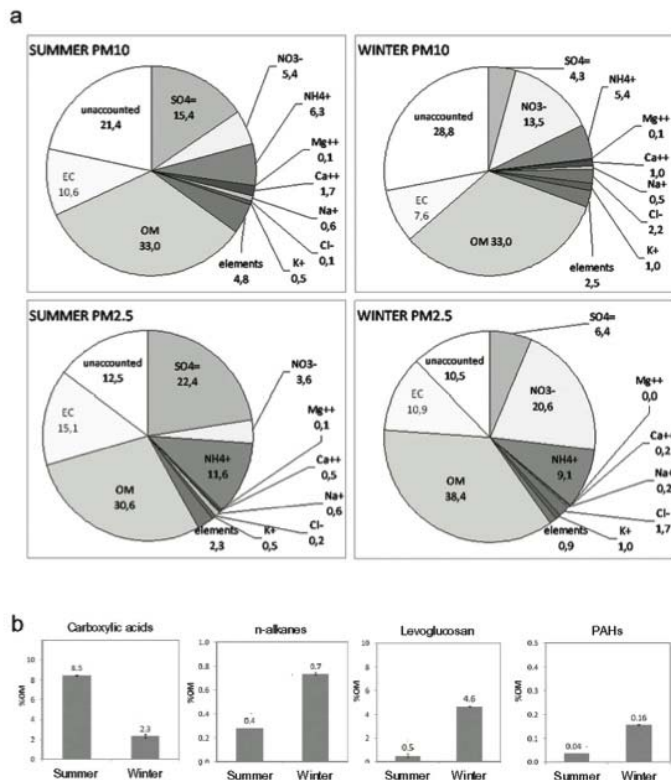


Figure 2 PM Characterization: a) chemical composition of summer and winter PM10 and PM2.5; b) histograms showing some specific primary (PAHs, levoglucosan and n-alkanes) and secondary (carboxylic acids) components of the OM in summer and

winter PM_{2.5} (reported as mass% of OM); OM - organic matter; EC - elemental carbon. Figure from Mantecca et al 2012.

1.4.1 Bioaerosols

Beside the extensively studied chemical fraction of PM, there is a likewise important fraction, which is less known and less investigated: the bioaerosol. The term bioaerosol indicates the presence in the atmosphere of suspended particles of biological origin such as bacterial cells, fungal spores, viruses, pollens and their by-products such as endotoxins (Stetzenbach et al, 2004).

In recent years a great interest has boosted on this fraction, because of the increasing evidences that it may have critical effects on biological systems (see chapter 5).

For example, bioaerosols could negatively affect human health and ecosystem functioning, by determining tissue inflammation (Camatini et al, 2010) and spreading of plant, livestock and human pathogens (Hirano et al, 1983; Pillai et al, 2002). Such effects are intensified in urban areas, where PM levels are particularly high. Although air quality has generally improved in the urban areas of developed countries in the second half of the last century, recent studies have provided evidence for a larger impact of urban air pollution on human health than previously hypothesized (Gouveia et al, 2006). Therefore, it has been proven that airborne microorganisms are able to influence atmospheric processes through the initialization of cloud condensation and ice nucleation phenomena, potentially altering precipitation patterns (Christner et al, 2008; Mohler et al, 2007). Finally, it is reasonable to hypothesize that some microorganisms present in the atmosphere could be capable to degrade different pollutants, although this is not already demonstrated.

1.5 Aim of the project

As described above, the atmosphere can be considered an extreme environment for microorganisms. High solar radiation, low moisture, low levels of nutrients and large dispersion occurrence could represent an obstacle to survival and reproduction of microorganisms. However, the presence and relevance of airborne bacteria have

been widely demonstrated (Womack et al, 2010). Due to this relevance, the biological fraction of PM raised a great interest in recent years. However, there is still a considerable lack of knowledge in this field (Polymenakou, 2012). The baseline variability in the types of microbes found in the atmosphere is just beginning to be assessed. Hence, it is currently difficult to establish linkages between airborne microbes and health outcomes. Long-term temporal analysis of microbial bioaerosol communities are needed to identify disturbances to this environment, which may be characteristic of microbial contamination events that are important to human and environmental health (Bowers et al, 2013).

Therefore, the aim of this thesis is to provide an exhaustive description of bacterial communities associated to PM of different size fractions in two Northern Italy urban sites, Milan and Venice and to put forward hypotheses about the processes underlying the assembly of the microbial community in urban atmosphere. A particular consideration will be given to address the following scientific questions: (i) which is the temporal and spatial variability of airborne microbial communities, (ii) which are, the environmental factors which drive community-assembly processes, (iii) which are the sources of airborne bacteria.

1.6 Summary of the thesis

This thesis is structured in different chapters addressing specific topics. The content of the chapters 2, 3, 5 and 6 have been already published as single articles in scientific journals while the paper containing the results presented in chapter 4 is in preparation. Here a brief summary of the content of each chapter is reported.

Chapter 2 reports the results of a cultivation-based study of the bacterial communities associated with PM₁₀ in Milan urban area. This study represented the first taxonomic snapshot of airborne bacteria in Milan through the cultivation of more than 400 isolates. Although the cultivation-dependent approach had severe biases for the definition of the microbial community structures, it allowed the determination of

Introduction

specific physiological traits of the isolates. In fact, this work specifically aimed at evaluating the presence of potentially harmful microorganisms belonging to the *Staphylococcus* genus. The antibiotic resistance and the presence of the corresponding genetic determinants were further investigated in the *Staphylococcus* isolates. Provided that the spreading of antibiotic resistant microorganisms is of particular concern in clinical settings, this work demonstrated that, although the evaluation of the actual risk posed by airborne microorganisms to human health is very challenging, the origin of antibiotic resistance genes can be traced in natural environments. The results reported in chapter 2 have been included in the following published paper:

Gandolfi, A. Franzetti, V. Bertolini, E. Gaspari, G. Bestetti (2011) "Antibiotic resistance in bacteria associated to coarse atmospheric particulate matter in an urban area" JOURNAL OF APPLIED MICROBIOLOGY 110: 1612-1620

Chapter 3 reports the first work in which quantitative PCR and Illumina technology were used to provide the description of airborne bacterial communities associated with Total Suspended Particulate (TSP) of the Milan urban area. Forty TSP samples were collected in ten-day sampling sessions, with one session per season. The mean bacterial abundance was about 10^4 ribosomal operons per m^3 of air and was lower in winter than in the other seasons. Communities were dominated by *Actinobacteridae*, *Clostridiales*, *Sphingobacteriales* and few proteobacterial orders (*Burkholderiales*, *Rhizobiales*, *Sphingomonadales*, *Pseudomonadales*). Chloroplasts were abundant in all samples. The variation in community composition observed within seasons was comparable to that observed between seasons, thus suggesting that airborne bacterial communities show large temporal variability, even between consecutive days. The structure of airborne bacterial communities therefore suggests that soil and plants are the sources, which contribute most to the airborne communities of Milan atmosphere, but the structure of bacterial community seems to depend mainly on the

source of bacteria that predominates in a given period of time. The results reported in chapter 2 have been included in the following published paper:

V. Bertolini, I. Gandolfi, R. Ambrosini, G. Bestetti, E. Innocente, G. Rampazzo, A. Franzetti (2013) “Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy” APPLIED MICROBIOLOGY AND BIOTECHNOLOGY – 97: 6561-6570

Chapter 4 reports a one-year survey on three urban different sampling locations in two cities of Northern Italy. The study concerned the assessment of the variability and the community-assembly processes that affected the airborne microbial communities associated with PM10. A total of 32 samples of PM10 for each sampling site were collected. The Illumina platform was used to sequence 16S rRNA gene libraries and to obtain the data necessary for a complete description of the airborne bacteria communities. Furthermore, the quantitative PCR was used to estimate the total number of bacteria present in the selected environments.

The main aims of this work were the assessment of the microbial community variability among seasons and sampling sites in order to gain insight into the processes that assemble the microbial community. The obtained data confirm that dispersion is an important community-assembly process. In particular, bacterial sources significantly contributed to change the community structure over time and space. Moreover, important clues that indicate possible selection processes (i.e. the effects of Ba, SO_4^{2-} and Mg^{2+}) were discovered.

During my PhD fellowship the increasing interest in the ecology of airborne microorganism led to the publication of several articles addressing this issue with the involvement of a few research groups around the World. Since the body of research became consistent the Editor-in-Chief of the journal “Applied Microbiology and Biotechnology” asked our research group to critically evaluate the results and the

methodologies used so far and provide a comprehensive picture of the state of the art. I was involved in this effort and, owing to the particular pertinence of the content, I report this published review in **Chapter 5**. The bibliographic details of the review are the following:

I. Gandolfi, V. Bertolini, R. Ambrosini, G. Bestetti, A. Franzetti (2013) "Unravelling the bacterial diversity in the atmosphere" APPLIED MICROBIOLOGY AND BIOTECHNOLOGY 97: 4727-4736

As reported above, long-range dispersal of microorganism occurred mainly through the atmosphere with significant implications for the global geographical distribution of these organisms. Moreover, high-elevation cold environments are considered ideal places to test hypotheses about basic ecological processes of microorganisms. In the frame of a research project dealing with the microbial ecology of Alpine environments I was involved in a study aiming at testing hypotheses about biogeographical distribution of bacterial populations. In particular, we investigated whether phylotypes of the genus *Polaromonas*, which is ubiquitous in cold environments, do present a biogeographical distribution. Although this peculiar topic is not the main issue of my thesis, it represents a not negligible aspect of the airborne microorganisms investigation. For this reason, it has been reported in **Chapter 6**. The paper reporting the obtained results is the following:

A. Franzetti, V. Tatangelo, I. Gandolfi, V. Bertolini, G. Bestetti, G. Diolaiuti, C. D'Agata, C. Mihalcea, C. Smiraglia and R. Ambrosini (2013) "Bacterial community structure on two alpine debris-covered glaciers and biogeography of Polaromonas phylotypes" ISME Journal – 7: 1483–1492.

1.7 References

Bent SJ, Forney LJ (2008) The tragedy of the uncommon: understanding limitations in the analysis of microbial diversity. *ISME J* 2: 689-695.

Bowers RM, Sullivan AP, Costello EK, Collett Jr JL, Knight R, Fierer N (2011b) Sources of bacteria in outdoor air across cities in the Midwestern United States. *Appl Environ Microbiol* 77, 6350-6356.

Bowers RM, Clements N, Emerson JB, Wiedinmyer C, Hannigan MP, Fierer N (2013) Seasonal variability in bacteria and fungal diversity of near-surface atmosphere. *Environ Sci Technol* DOI: 10.1021/es402970s.

Brüggemann E, Gerwig H, Gnauk TH, Müller K, Herrmann H. (2009) Influence of seasons, air mass origin and day of the week on size-segregated chemical composition of aerosol particles at a kerbside. *Atmos. Environ.* 43: 2456-2463.

Camatini M, Corvaja V, Pezzolato E, Mantecca P, Gualtieri M. (2010) PM10-biogenic fraction drives the seasonal variation of pro-inflammatory response in A549 cells. *Environ Toxicol* 27:63–73.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Tumbaugh PJ, Walters WA, Widmann J, Yatsunencko T, Zaneveld J, Knight R. (2010) "QIIME allows analysis of high-throughput community sequencing data". *Nat. Methods.* 7: 335–336.

Christner BC, Morris C, Foreman CM, Cai R, Sands DC (2008) Ubiquity of Biological Ice Nucleators in Snowfall, *Science* 319: 1214.

Introduction

Claesson MJ, Wang QO, O'Sullivan O, Greene-Diniz R, Cole R JR, Ross P, O'Toole PW (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Res* 38, e200.

Darcy JL, Lynch RC, King AJ, Robeson MS, Schmidt SK. (2011). Global distribution of *Polaromonas* phylotypes - evidence for a highly successful dispersal capacity. *PLoS One* 6:e23742. doi:10.1371/journal.pone .0023742.

Gouveia NC, Maisonet M. (2006). "Health effects of air pollution: an overview. In: *Air Quality Guidelines: Global Update 2005*". World Health Organization: Copenhagen, Denmark, pp 87–109.

Hess D. (2010). "Introduction to the atmosphere" in "McKnight's Physical Geography: A Landscape Appreciation", 10th ed., Pearson, London.

Hirano SS, Upper CD. (1983). "Ecology and epidemiology of foliar bacterial plant pathogens". *Annu Rev Phytopathol* 21:243–269.

Lozupone CA, Knight R (2008) Species divergence and the measurement of microbial diversity. *FEMS Microbiol. Rev.* 32: 557-558.

Mantecca P, Gualtieri M, Longhin E, Bestetti G, Palestini P, Bolzacchini E, Camatini M. (2012) Adverse biological effects of Milan urban PM looking for suitable molecular markers of exposure. *Chem Ind Chem Eng Q* 18: 635-641.

Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, Horner-Devine MC, Kane M, Krumins JA, Kuske CR, Morin PJ, Naeem S, Ovreas L, Reysenbach

Introduction

AL, Smith VH, Staley JT (2006) Microbial biogeography: putting microorganisms on the map. *Nat Rev Micro* 4,102-112

Mohler O, DeMott PJ, Vali G, Levin Z (2007) Microbiology and atmospheric processes: the role of biological particles in cloud physics. *Biogeosci Discuss* 4: 2559-2591.

Nemergut DR; Schmidt SK; Fukami T, O'Neill SP, Bilinska TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S (2013) Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* 77, 342-356.

O'Malley MA (2008) 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Stud Hist Phil Biol & Biomed Sci* 39: 314-325.

Ovreas L, Daae FL, Torsvik V, Rodriguez-Valera F. (2003). Characterization of microbial diversity in hypersaline environments by melting profiles and reassociation kinetics in combination with terminal restriction fragment length polymorphism (T-RFLP). *Microb Ecol* 46: 291–301

Peccia J, Hernandez M (2006) Incorporating polymerase chain reaction-based identification, population characterization, and quantification of microorganisms into aerosol science: A review. *Atmospheric Environment* 40, 3941-3961.

Perrone G, Eva E, Solarino S, Cadoppi P, Balestro G, Fioraso G, Tallone S (2010) Seismotectonic investigations in the inner Cottian Alps (Italian Western Alps): An integrated approach. *Tectonophysics* 496, 1-16.

Pillai SD, Ricke SC (2002) Bioaerosols from municipal and animal wastes: background and contemporary issues. *Can J Microbiol* 48, 681-696.

Introduction

Polymenakou PN (2012) Atmosphere: A Source of Pathogenic or Beneficial Microbes? Atmosphere 3: 87-102.

Polz MF, Hunt DE, Preheim SP, Weinreich DM (2006) Pattern and mechanisms of genetic and phenotypic differentiation in marine microbes. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 361: 2009-2021.

Rinsoz T, Duquenne P, Greff-Mirguet G, Oppliger A (2008) Application of real-time PCR for total airborne bacterial assessment: comparison with epifluorescence microscopy and culture dependent methods. Atmos Environ 42, 6767–6774.

Schmidt SK, Costello EK, Nemergut DR, Cleveland CC, Reed SC, Weintraub MN, Meyer AF, Martin AM (2007) Biogeochemical consequences of rapid microbial turnover and seasonal succession in soil. Ecology 88: 1379-1385.

Stetzenbach LD, Buttner MP, Cruz P (2004) Detection and enumeration of airborne biocontaminants. Curr Opin Microbiol 15: 170-174.

Torsvik V, Daae FL, Sandaa RA, Ovreas L (1998) Novel techniques for analysing microbial diversity in natural and perturbed environments. J Biotechnol 64: 53–62.

Vellend M (2010) Conceptual synthesis in community ecology. Quarterly review of biology 85: 183-206.

Whitaker RJ (2006) Allopatric origins of microbial species. Philosophical transactions of the royal society: series B 361: 1975-1984.

Whitman WB, Coleman DC, Wiebe WJ (1998) Prokaryotes: The unseen majority. Proc Natl Acad Sci USA 95:6578-6583.

Introduction

Womack AM, Bohannon BJM, Green JL (2010) Biodiversity and biogeography of the atmosphere. *Philos Trans R Soc B Biol Sci* 365, 3645–3653.

CHAPTER 2

**Antibiotic resistance in bacteria associated
with coarse atmospheric particulate matter in
an urban area**

2.1 Introduction

The common perception about the spreading of antibiotic-resistant bacteria and the genes conferring such resistance is generally connected with the overuse of antibiotics, especially in clinical settings. At present, the pathogens considered the most challenging to deal with are *methicillin-resistant Staphylococcus aureus* (MRSA), *vancomycin-resistant Enterococci* (VRE) and extended-spectrum β -lactamases (ESBL) producing *Enterobacteriaceae* (Livermore, 2009; Septimus and Kuper, 2009).

However, several environments other than hospitals have been identified as sources of both antibiotics and antibiotic-resistant bacteria. For example, the main anthropogenic inputs of them in the aquatic environment come from wastewater discharges, including those from urban areas, agricultural practices and animal feeding operations (Kümmerer, 2009). This is exacerbated by devices in wastewater plants and drinking water distribution systems often ineffective in removal of antibiotic traces and resistant bacteria. Moreover, also the use of subinhibitory concentrations, for example in aquaculture and livestock production, has been demonstrated to enhance the rate of transfer and dissemination of antibiotic resistance genes. For all these reasons, antibiotic resistance genes have started to be considered as environmental pollutants themselves (Kümmerer, 2009), also considering that their presence in several environments can increase the chance for pathogens to acquire and in turn transfer new resistance determinants (Martínez, 2009).

Although such anthropogenic-originated pressure may have a crucial importance in the spreading of antimicrobial resistance, it has to be also remembered that many antibiotics have a natural origin, being produced especially by soil micro-organisms (Martínez, 2008). Consequently, also antibiotic resistance genes have been probably originated from environmental strains, although it has been speculated that their original roles were different from the so-called 'weapon-shield' function shown in clinical contexts (Martínez, 2008). Accordingly with such observations, the

existence in environmental strains of a potential wide reservoir of antibiotic resistance genes, possibly still unknown, can be inferred. Nevertheless, as systematic studies on the antibiotic resistome in the environment have not been conducted so far, little is known about the actual behaviour of such genes in natural ecosystems (Aminov and Mackie, 2007; Allen et al, 2010). Owing to this lack of knowledge and to the potential chance for pathogens to acquire antibiotic resistance from environmental strains, it has been conservatively suggested to avoid as much as possible the contact between human pathogens and environmental microorganisms (Baquero et al, 2008; Martínez, 2009). From a practical point of view, however, such a goal is not so easy to achieve, because the microbial ecosystems are not completely isolated and the occurrence of gene exchange among different compartments has been demonstrated, although the actual transfer rate has not been precisely evaluated (Aminov and Mackie, 2007). This is especially true for the atmosphere, which is extremely mobile and therefore subject to extensive transport and contact with other environmental compartments. In addition, there is a significant lack of information about aerosolization of pathogens, their viability during atmospheric transport, their response to meteorological and climate change factors and their interactions with environmental bacteria in such an extreme compartment. This poses a serious obstacle to estimating the actual risk posed by airborne micro-organisms to human health (Lai et al, 2009).

The aim of this study was the assessment of antimicrobial resistance in airborne bacteria associated with coarse particulate matter (PM₁₀, i.e. the fraction of particulate matter whose aerodynamic diameter is $\leq 10 \mu\text{m}$) in an urban area, with specific considerations about the *Staphylococcus* genus, which is regarded as one of the most threatening micro-organisms in respiratory diseases, especially because of MRSA spreading. We chose to analyze PM₁₀ fraction because it is known that particles smaller than $10 \mu\text{m}$ are those, which can penetrate into the lungs (Griffin, 2007). Although genetic determinants carrying antibiotic resistance may

be potentially borne by culturable as well as non culturable microorganisms, we chose to limit the analysis to the former. We could therefore study also phenotypic resistances and associate them with a phylogenetic classification.

2.2 Material and Methods

2.2.1 Sampling and isolation of microbial strains

PM10 samples were collected with a 20–830 Eight Stage Cascade Impactor (Andersen Instruments, Waltham, MA, USA) on nitrocellulose filters (Orange Scientific, Waterloo, Belgium), at an air flux speed of 38 l min^{-1} . Characteristics of PM10 in the Milan urban area were previously described (Marcazzan et al. 2001). With the aim of detecting also the low-abundance populations, we decided to isolate at least half of the colonies in each sample. Because of the high number of obtained isolates, we chose to limit the analysis to a representative sample for each of the two considered seasons (summer and winter). Sampling was carried out for 24-h periods in late summer, on 15 September 2008, and in winter, on 11 and 12 February 2009. Winter sampling was carried out for a total of 48 h because it was assumed that the number of bacteria would be lower than in summer. The sampling site was located in Torre Sarca, Milan, (Italy, $45^{\circ}31'20''\text{N}$, $9^{\circ}12'45''\text{E}$), a typical urban station in the northern part of the city. The site is located next to a high traffic road, and it consists of a grass surrounded by small buildings and fenced by a wall approximately 1,50 m high. The sampler was placed near one of the small buildings and next to the grass, at about 20 m from the nearest roads and 50 m from the nearest traffic lights. PM sampling was performed at about 1 m from the ground. Photographs showing the sampling site and its surroundings are given as Supporting Information (Figs S1 and S2). Because of the site position, characteristics of collected particulate matter were mainly affected by car traffic. On the contrary, the presence of a small airport located at approximately 1 km from the site, with a very low air traffic counting only a few airplanes per week, had no influence on that. Also winds, which are generally

recognized to affect PM characteristics and distribution, are scarce and show no prevalent directions in Milan region. Moreover, our sampling was carried out in the complete absence of wind.

At the end of the sampling, each filter was directly transferred on agarized LB (Sambrook and Russell, 2001) added with 10 mg l^{-1} of actidione to inhibit eukaryotic growth; plates were incubated at 30°C until the formation of new visible colonies on nitrocellulose filters stopped. Filters from stage zero of the impactor were not considered because they collected $\text{PM} > 10$. To isolate pure cultures, colonies were picked and repeatedly streaked on fresh agarized LB (Biolife Italiana, Milano, Italy); from each filter roughly one-half of the colonies were randomly chosen for isolation. All obtained microbial strains were stored in 25% glycerol at -20°C .

A fragment of the 16S rRNA gene containing V4 and V5 hypervariable regions was amplified with Com primer set (Schwieger and Tebbe 1998). The taxonomic attribution of sequences was carried out by the Ribosomal Database Project (RDP) Bayesian Classifier (<http://rdp.cme.msu.edu/index.jsp>).

2.2.2 Disc diffusion test

Microbial isolates classified as belonging to taxa that also include common potential pathogenic species were tested for antibiotic susceptibility using the disc diffusion method on agarized Mueller-Hinton medium (MH), according to the Clinical Laboratory Standards Institute protocols (CLSI 2009). A total of 10 antibiotics were chosen representing the major classes of known antibacterial compounds: penicillins (piperacillin $100 \mu\text{g}$, PIP), third-generation cephalosporins (cefotaxime $30 \mu\text{g}$, CTX; ceftazidime $30 \mu\text{g}$, CAZ), carbapenems (imipenem $10 \mu\text{g}$, IPM), macrolides (oleandomycin $15 \mu\text{g}$, OL), tetracyclines (TE $30 \mu\text{g}$), aminoglycosides (gentamicin $10 \mu\text{g}$, GM), glycopeptides (vancomycin $30 \mu\text{g}$, VA), quinolones (ciprofloxacin $5 \mu\text{g}$, CIP) and rifamycins (rifampicin $30 \mu\text{g}$, RA). All antibiotics were provided by BioMerieux SA, Marcy l'Etoile, France. *Staphylococcus* strains were also screened for the resistance to

the more common macrolide erythromycin (E, 15 µg) and to cefoxitin (FOX, 30 µg) (Oxoid, Basingstoke, Hants, UK); the latter is used to identify MRSA strains, giving better results than oxacillin test (Broekema et al. 2009). *Escherichia coli* strain TOP 10 was used as control.

2.2.3 Identification of *Staphylococcus aureus* strains

Coagulase-positive *S. aureus* strains were identified using Baird Parker agar medium + RPF (rabbit plasma fibrinogen) supplement II (Baird-Parker 1962) (Biolife Italiana); plates were prepared according to the manufacturer's instructions and incubated at 37°C for 48 h after streaking of *Staphylococcus* strains. After incubation, coagulase positive staphylococci form black or grey or even white small colonies surrounded by an opaque halo.

2.2.4 Detection of antibiotic resistance genes

The *Staphylococcus* strains that resulted resistant to VA, OL, erythromycin, TE and / or FOX on the basis of the disc diffusion test were further investigated for the presence of genetic determinants conferring specific resistance to these antibiotics. Genomic DNA was extracted from fresh cultures using the phenol–chloroform method. (Sambrook and Russell 2001). The presence of the following genetic determinants was PCR-checked using specific primers: *vanA* for VA resistance (Dutka-Malen et al. 1995), *ermA* and *ermC* for macrolide resistance (Martineau et al. 2000), *tetK* and *tetM* for tetracycline resistance (Strommenger et al. 2003) and *mecA* for methicillin resistance (Zhang et al. 2004).

PCR conditions were set as follows: for *vanA* amplification, an initial denaturation at 95°C for 2 min, 35 cycles of 95°C for 60 s, 48°C for 30 s and 72°C for 45 s, and a final extension at 72°C for 5 min; for *ermA* and *ermC* amplification, an initial denaturation at 95°C for 2 min, 35 cycles of 95°C for 60 s, 51°C for 30 s and 72°C for 30 s, and a final extension at 72°C for 5 min; for *tetK* and *tetM* amplification, an initial denaturation at 95°C for 4 min, 35 cycles of 95°C for 30 s, 55°C for 45 s and 72°C for 45 s, and a final

extension at 72°C for 5 min; for *mecA* amplification, an initial denaturation at 95°C for 2 min, 35 cycles of 95°C for 60 s, 56°C for 30 s and 72°C for 30 s, and a final extension at 72°C for 5 min.

Each PCR was carried out in a total volume of 20 µl: 10 µl of GoTaq Green Master Mix 2x (Promega, Madison, WI, USA) containing 400 µmol l⁻¹ dNTPs and 3 µmol l⁻¹ MgCl₂, 1 µmol l⁻¹ of each primer and 1 µmol l⁻¹ of template DNA. The amplified PCR products were purified with Wizard® SV Gel and PCR Clean-Up System (Promega) and sequenced to confirm the presence of the antibiotic resistance genes.

2.3 Results

2.3.1 Isolation of microbial strains

A total number of 281 and 288 microbial strains were isolated from summer and winter PM10, respectively. The total load of culturable bacteria in particulate samples was 10³ and 5³ CFU m⁻³, respectively, for summer and winter PM10. The sequence of the hypervariable regions V5 and V6 of 16S rRNA was used as input for the RDP Classifier tool. The results of the RDP classification are shown in Table 1 and in Fig. S3. In summer PM10, there is a sharp prevalence of Gram-positive micro-organisms, especially belonging to the class of Bacilli, also including staphylococci, while in winter PM10 Gram-negative bacteria are more represented, especially the class of *Gammaproteobacteria*, which mainly includes *Pseudomonadaceae* and *Enterobacteriaceae*.

Among the isolated taxa in both seasons, families *Pseudomonadaceae* and *Enterobacteriaceae* and genera *Staphylococcus*, *Acinetobacter* and *Enterococcus* were selected for further analysis, because they contain, besides environmental strains, some species known as pathogens or opportunistic pathogens.

Chapter 2

Table 1 Ribosomal Database Project classification of isolates from summer and winter PM10. All strains that were further analysed are shown in bold.

| Summer PM10 | Winter PM10 |
|--|--|
| Root (281 sequences) | Root (288 sequences) |
| domain Bacteria (279) | domain Bacteria (284) |
| phylum "Bacteroidetes" (1) | phylum "Firmicutes" (88) |
| class Flavobacteria (1) | class "Bacilli" (87) |
| order "Flavobacteriales" (1) | order Bacillales (86) |
| family Flavobacteriaceae (1) | family "Listeriaceae" (1) |
| genus Chryseobacterium (1) | genus Brochothrix (1) |
| phylum "Actinobacteria" (17) | family Bacillaceae (12) |
| class Actinobacteria (17) | genus Bacillus (10) |
| subclass Actinobacteridae (17) | unclassified Bacillaceae (2) |
| order Actinomycetales (17) | family "Staphylococcaceae" (71) |
| suborder Micrococcineae (17) | genus Salinicoccus (2) |
| family Micrococcaceae (7) | genus Staphylococcus (68) |
| genus Micrococcus (3) | unclassified "Staphylococcaceae" (1) |
| genus Arthrobacter (4) | family Planococcaceae (1) |
| family Microbacteriaceae (10) | genus Sporosarcina (1) |
| genus Microbacterium (1) | unclassified Bacillales (1) |
| genus Curtobacterium (3) | unclassified "Bacilli" (1) |
| unclassified Microbacteriaceae (6) | unclassified "Firmicutes" (1) |
| phylum "Firmicutes" (210) | phylum "Proteobacteria" (168) |
| class "Bacilli" (210) | class Betaproteobacteria (1) |
| order "Lactobacillales" (3) | order Burkholderiales (1) |
| family "Enterococcaceae" (3) | family Burkholderiaceae (1) |
| genus Enterococcus (3) | genus Cupriavidus (1) |
| order Bacillales (200) | class Gammaproteobacteria (165) |
| family Planococcaceae (4) | order Xanthomonadales (4) |
| genus Kurthia (1) | family Xanthomonadaceae (4) |
| genus Sporosarcina (1) | genus Stenotrophomonas (4) |
| unclassified Planococcaceae (2) | order "Enterobacteriales" (57) |
| family "Staphylococcaceae" (26) | family Enterobacteriaceae (57) |
| genus Staphylococcus (23) | genus Pantoea (32) |
| unclassified "Staphylococcaceae" (3) | unclassified Enterobacteriaceae (25) |
| family "Paenibacillaceae" (1) | order Pseudomonadales (103) |
| genus Paenibacillus (1) | family Moraxellaceae (2) |
| family Bacillaceae (160) | genus Psychrobacter (2) |
| genus Lysinibacillus (2) | family Pseudomonadaceae (101) |
| genus Exiguobacterium (12) | genus Pseudomonas (98) |
| genus Bacillus (81) | unclassified Pseudomonadaceae (3) |
| unclassified Bacillaceae (65) | unclassified Gammaproteobacteria (1) |
| unclassified_Bacillales (9) | unclassified "Proteobacteria" (2) |
| unclassified "Bacilli" (7) | phylum "Actinobacteria" (17) |
| phylum "Proteobacteria" (46) | class Actinobacteria (17) |
| class Alphaproteobacteria (1) | subclass Actinobacteridae (17) |
| order Caulobacterales (1) | order Actinomycetales (17) |
| family Caulobacteraceae (1) | suborder Streptosporangineae (1) |

| | |
|--|------------------------------------|
| genus <i>Brevundimonas</i> (1) | family Nocardioptaceae (1) |
| class Gammaproteobacteria (45) | genus <i>Nocardopsis</i> (1) |
| order Pseudomonadales (27) | suborder Micrococccineae (13) |
| family Moraxellaceae (10) | family Sanguibacteraceae (1) |
| genus <i>Acinetobacter</i> (10) | genus <i>Sanguibacter</i> (1) |
| family Pseudomonadaceae (17) | family Microbacteriaceae (2) |
| genus <i>Pseudomonas</i> (11) | genus <i>Curtobacterium</i> (2) |
| unclassified Pseudomonadaceae (6) | family Micrococcaceae (10) |
| order "Enterobacteriales" (17) | genus <i>Kocuria</i> (2) |
| family Enterobacteriaceae (17) | genus <i>Micrococcus</i> (7) |
| genus <i>Citrobacter</i> (1) | unclassified Micrococcaceae (1) |
| genus <i>Buttiauxella</i> (3) | suborder Streptomycineae (3) |
| genus <i>Pantoea</i> (5) | family Streptomycetaceae (3) |
| unclassified Enterobacteriaceae (8) | genus <i>Streptomyces</i> (2) |
| unclassified Gammaproteobacteria (1) | unclassified Streptomycetaceae (1) |
| unclassified Bacteria (5) | unclassified Bacteria (11) |
| unclassified Root (2) | unclassified Root (4) |

2.3.2 Disc diffusion test

Disc diffusion test was performed on a total number of 72 summer and 171 winter isolates, divided as follows: 26 *Staphylococcus*, 16 *Pseudomonadaceae*, 3 *Enterococcus*, 17 *Enterobacteriaceae* and 10 *Acinetobacter* for summer PM10, and 67 *Staphylococcus*, 49 *Pseudomonadaceae* and 55 *Enterobacteriaceae* for winter PM10. In winter samples, no strains belonging to genera *Enterococcus* and *Acinetobacter* were isolated. Approximately 20% of the strains could not be regrown from glycerol stock and were therefore excluded from such analysis. Each isolate was classified as sensitive or resistant on the basis of the halo diameter on the plate. Strains showing an intermediate resistance were included in the 'sensitive' class in our analyses. The results of resistances to the 10 antibiotics are shown in Fig. 1a,b.

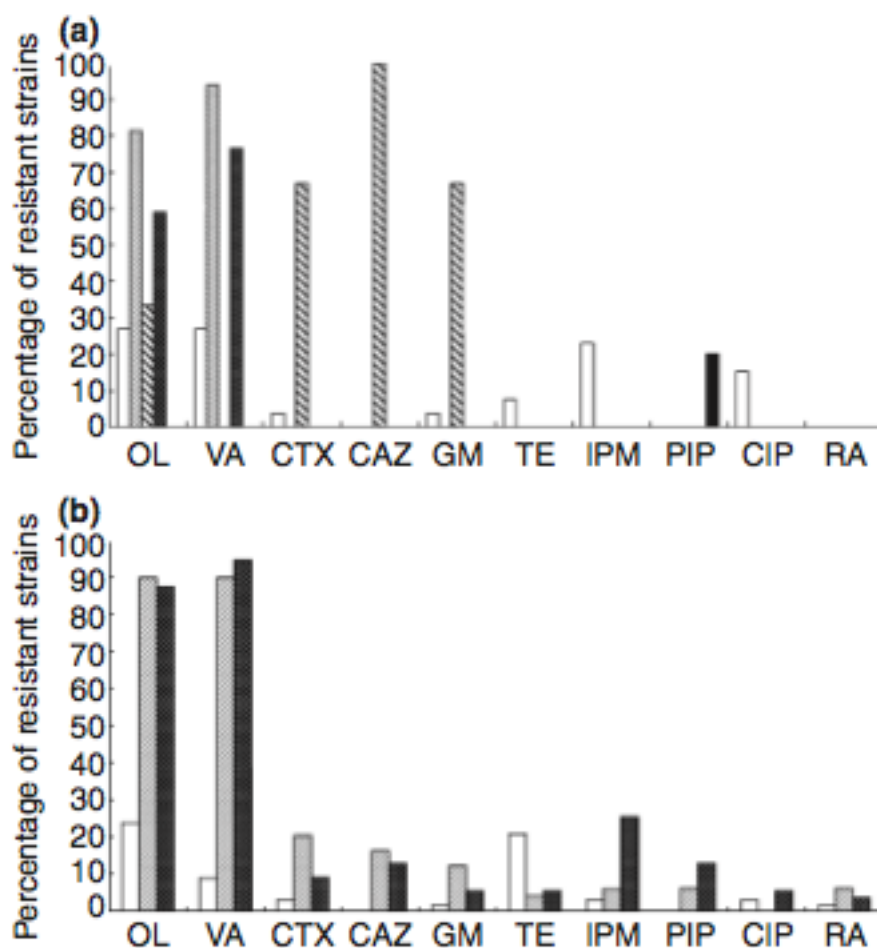


Figure 1 Percentage of resistant strains to the ten selected antibiotics: (a) summer isolates, (b) winter isolates. The absolute number of resistant strains for each taxon is shown in parentheses in the legend. OL, oleandomycin; VA, vancomycin; CTX, cefotaxime; CAZ, ceftazidime; GM, gentamicin; TE, tetracycline; IPM, imipenem; PIP, piperacillin; CIP, ciprofloxacin; RA, rifampicin. [a: (□) *Staphylococcus* (26); (▨) *Pseudomonadaceae* (16); (▩) *Enterococcus* (3); (■) *Enterobacteriaceae* (17); and (■) *Acinetobacter* (10) b: (□) *Staphylococcus* (67); (▨) *Pseudomonadaceae* (49); (■) *Enterobacteriaceae* (55)].

The main detected resistances, with percentages higher than 50% and up to 100%, are towards antibiotics to which some micro-organisms are known to have an intrinsic resistance, i.e., macrolides (OL) and VA for *Enterobacteriaceae* and *Pseudomonadaceae* and cephalosporins (CTX and CAZ) and aminoglycosides (GM) at low levels for enterococci. In addition to that, however, winter *Enterobacteriaceae* and *Pseudomonadaceae* isolates also showed resistances to the tested β -lactams (PIP, CTX, CAZ and IPM), with percentages ranging 6–25%. Summer *Acinetobacter* strains showed a certain degree (20%) of resistance only to PIP, while staphylococci resistances were more heterogeneous, being distributed among all tested antibiotics, with the exception of CAZ and PIP, although percentages of resistance were always not higher than 27% (OL and VA for summer staphylococci). In general, winter isolates showed higher degrees of resistance than summer isolates.

Staphylococcus strains were further tested for erythromycin (E) and FOX resistance. The results are shown in Table 2: a non-negligible percentage of them, around 10%, resulted resistant also to these two antibiotics, with a slightly higher percentage (17%) for erythromycin resistance in winter staphylococci. Moreover, strains showing resistance to both tested macrolides (E and OL) represented a higher percentage (15 and 19% for summer and winter isolates, respectively) than those obtained for resistances to only one of the two compounds. A classification of staphylococci divided in classes according to their resistance to one or more antibiotics is reported in Fig. 2, as calculated on the overall results about the 12 tested compounds. The percentage of isolates resistant to at least one antibiotic is very high, approximately 70% in both seasons, with a slightly higher value for summer strains. Furthermore, staphylococci harbouring a multiresistance, up to five different compounds, represent a large fraction, accounting for 38 and 31% of the isolates for summer and winter, respectively.

Table 2 – Antibiotic resistance of *Staphylococcus* isolates to cefoxitin (FOX) and macrolides (E, erythromycin; OL, oleandomycin).

| Antibiotic resistance | Summer <i>Staphylococci</i> | Winter <i>Staphylococci</i> |
|-----------------------|--------------------------------|--------------------------------|
| FOX | 12% (3) | 9% (6) |
| only E | 8% (2) | 17% (12) |
| only OL | 12% (3) | 4% (3) |
| both E and OL | 15% (4) | 19% (13) |

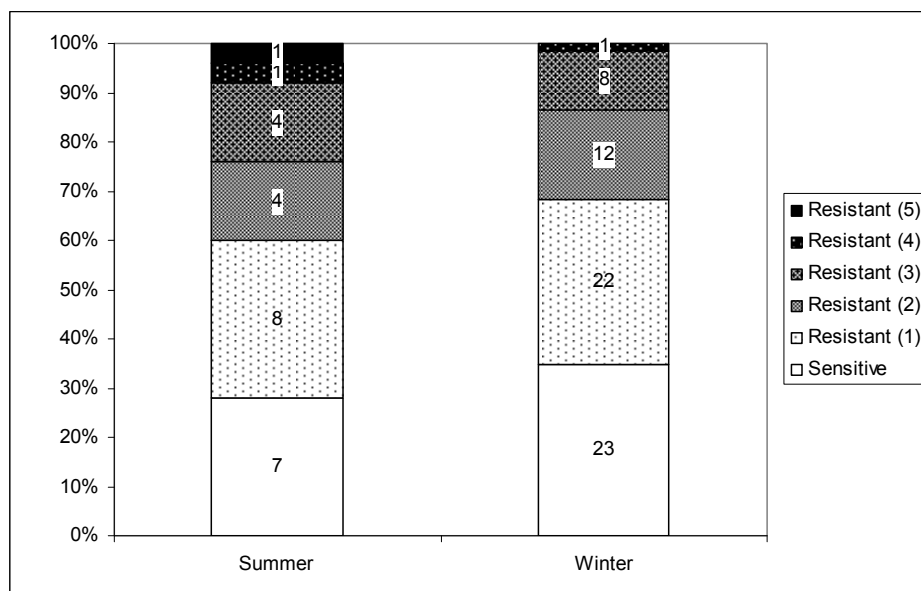


Figure 2 – Classification of *Staphylococcus* isolates according to the number of antibiotics against which they showed resistance, indicated by the number in parentheses in the legend. The absolute number of strains belonging to each class is shown on the histograms.

2.3.3 Identification of *Staphylococcus aureus* strains

None of the *Staphylococcus* strains produced a visible halo around the colonies on selective plates, thus resulting coagulase-negative. Therefore, it could be hypothesized that none of them belonged to *S. aureus* species. In addition, the type

strain showing the best match with the V5– V6 hypervariable regions of 16S rRNA of each *Staphylococcus* strain was identified through the RDP ‘Seqmatch’ tool; no best match corresponded with *S. aureus*. These results are shown in Fig. 3; a high percentage of both summer and winter staphylococci matched with *Staphylococcus succinus*, while the other most represented species were different between the two seasons, with a high prevalence of *Staphylococcus equorum* for winter isolates and a significant percentage of *Staphylococcus sciuri* in summer isolates. None of the strains showed the best match with the most common clinical or human-associated *Staphylococcus* strains, such as *Staphylococcus epidermidis* or *Staphylococcus saprophyticus*. Only one best match with *Staphylococcus lugdunensis*, which occasionally causes infections, was found.

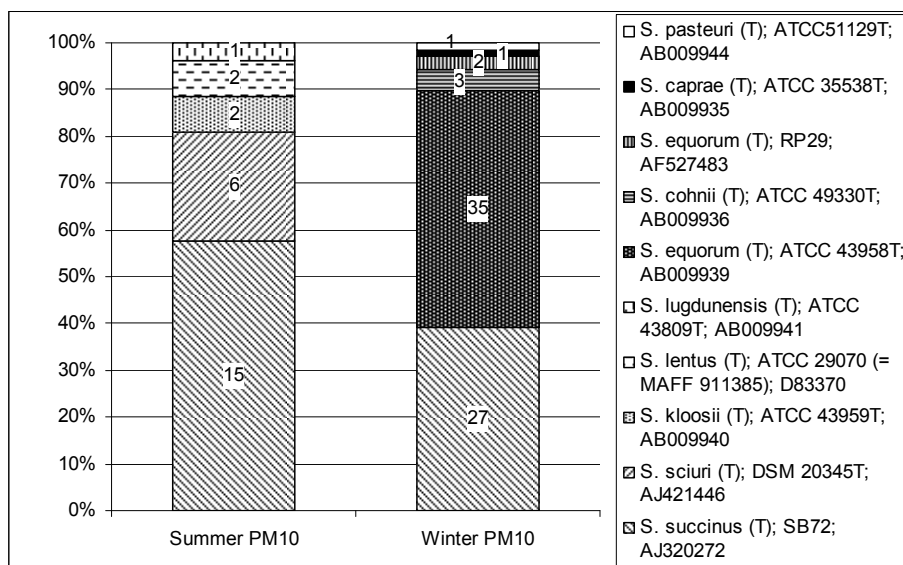


Figure 3 – Type strains showing the best match with the *Staphylococcus* isolates of this study. The absolute number of isolates matching a specific type strain is shown on the histograms. Deposit and accession numbers of the type strains are indicated in the legend.

2.3.4 Antibiotic resistance genes

As shown in Fig. 4a,b, 15 of 16 TE-resistant *Staphylococcus* strains were positive to tetK amplification, although the presence of tetM could not be detected. Sequencing of PCR products confirmed the high homology with other tetK genes present in databases and belonging to *S. aureus* strains. Three macrolide-resistant *Staphylococcus* strains gave a PCR product when ermC specific primers were used, while no products were obtained with primers targeting ermA. Sequencing of the amplicons showed a high homology with *S. aureus* ermC gene, as already found for tetK sequences.

PCR performed with primers for vanA gene gave only a specific PCR products, while mecA primers could not detect any methicillin-resistance gene.

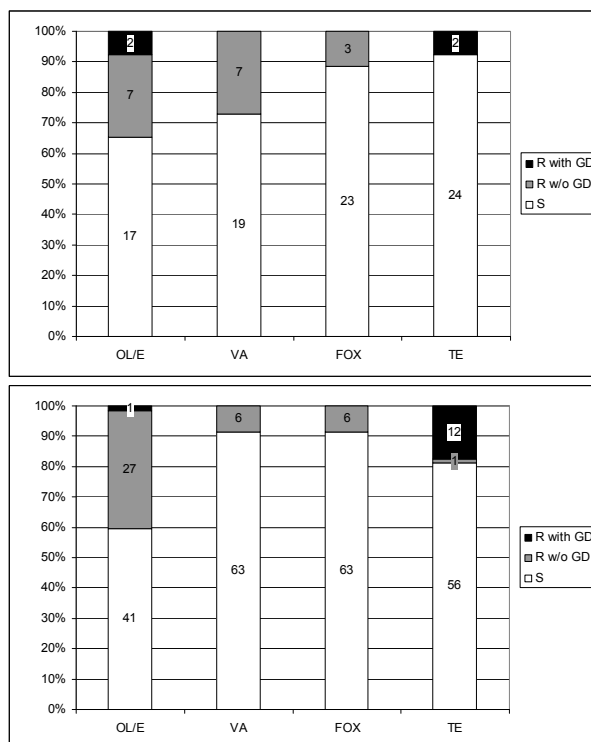


Figure 4 – Summary of resistances in *Staphylococcus* strains for antibiotics whose specific genetic determinant(s) was PCR-checked: (a) summer isolates, (b) winter isolates. S, sensitive; R w/o GD, resistant but the genetic determinant was not

detected; R with GD, resistant and the presence of the genetic determinant was confirmed. OL, oleandomycin, E, erythromycin, VA, vancomycin, FOX, ceftioxin, TE, tetracycline.

2.4 Discussion

The concentration of culturable bacteria in outdoor particulate is highly dependent on factors such as the particulate loads, the type of environment or the atmospheric and climatic conditions (Boreson et al. 2004; Harrison et al. 2005). For example, Fang et al. (2007) reported values between 71 and 22100 CFU m⁻³ in the city of Beijing, China, and values between 2 and 8500 CFU m⁻³ for other urban environments (see references therein). Our results are similar to the lowest obtained by other groups. Actually, the number of culturable isolates which can be obtained from air particulate is also strictly dependent on the intrinsic limitations of sampling conditions. In fact, it has been demonstrated that different sampling techniques, e.g., impaction on filters or agar plates, impingement or gravity deposition, can widely affect cultivability of prokaryotes (Buttner et al. 2002), especially because of the sampling stress. Therefore, comparison between microbial concentrations obtained under different operating conditions is not always reliable. Furthermore, sampling stress and the subsequent cultivation step are likely to cause a severe bias in the composition of the microbial community, so that the relative abundance of taxa found for isolated strains is not necessarily representative of the whole microbial community. In fact, if compared with cultivation-independent studies applying molecular methods (Maron et al. 2005; Brodie et al. 2007; Fierer et al. 2008; Polymenakou et al. 2008; Lee et al. 2010), our microbial community showed a lower degree of biodiversity. Nevertheless, the dominant taxa in our study coincided quite well with other cultivation-based surveys of airborne micro-organisms (Fang et al. 2007), where *Bacillus*, *Staphylococcus* and *Pseudomonas* were dominant.

The high resistance to macrolides (OL) and VA shown by *Pseudomonadaceae* and *Enterobacteriaceae* is of no particular concern, as the latter compound is a narrow-spectrum antibiotic targeting only Gram-positive bacteria (Arthur et al. 1996), and

also macrolides are essentially used in the treatment of such bacteria, in addition to a few Gram-negative cocci or intracellular micro-organisms, such as *Chlamydia* and *Rickettsia* (Leclercq 2002). Winter *Enterobacteriaceae* showed not-negligible percentages of resistance to β -lactams (PIP, CTX, CAZ and IPM). Nevertheless, as summer isolates were proven to be sensitive to such antibiotics, it cannot be definitively stated that ESBL-producing *Enterobacteriaceae*, which are a serious problem in clinical settings, are also highly widespread in the environment. Enterococci are known to have intrinsic resistances to cephalosporins (CTX and CAZ) and to low levels of aminoglycosides (GM) (Brooks et al. 2001), as confirmed by our data, although isolated strains were only three. On the contrary, our isolates were VA-sensitive, which means that VRE seem to be not particularly diffuse in the environment.

Staphylococci, conversely, showed a wide range of resistances in both seasons. In fact, summer isolates showed overall resistances to nine of 12 antibiotics, while winter isolates were resistant to 10 different compounds. Moreover, although the percentages of resistance to each antibiotic were generally low (always <30%), strains that showed a resistance to a single compound or a multiresistance accounted for the large majority of the isolates, thus demonstrating that antibiotic resistance is a widespread phenotype in airborne staphylococci. Such a diverse pattern of resistance, together with the high percentage of multiresistant strains, identifies staphylococci as a potential critical taxon present in the atmosphere, also considering the key role played by some species belonging to this genus in respiratory and lung diseases. Despite the non-negligible percentage of isolates resistant to macrolides, FOX, VA and TE, however, the corresponding genetic determinants could not generally be detected. In particular, only *tetK*, which mediates the resistance to TE through the action of an efflux protein, was found in almost all resistant isolates. It can be therefore hypothesized that the observed phenotypic resistances were attributed to aspecific mechanisms, such as cell wall reduced permeability or wide-spectrum efflux pumps. From this point of view, specific genetic determinants are less likely to be

transmitted among different *Staphylococcus* strains. This may reduce the theoretical risk associated with the inhalation of airborne micro-organisms.

Staphylococci isolated from PM10 generally showed no best matches with human pathogens, but rather with food or animal-associated species. Despite the absence of the former in our samples, however, the sequences of genetic determinants for TE and erythromycin resistance showed the highest similarity with those of *S. aureus* strains. Although this might be simply attributed to a bias in the submission of sequences to databases, as *S. aureus* is the most studied among Staphylococci, a substantial homology in genes harboured by environmental and clinical species can be inferred. For this reason, despite the relatively low abundance of antibiotic resistance determinants in airborne *Staphylococcus* strains, such homology remarks that the study of clinical and human-associated micro-organisms is not enough when the topic of concern is related to human health, but also environmental micro-organisms should be carefully considered. This is especially true if we consider that the overall genetic exchange among strains from different environments has not completely either elucidated or quantified so far.

2.5 Acknowledgements

This work was partially funded by Fondazione Cariplo (Milano, Italy) in the frame of the project TOSCA (Toxicity of particulate matter and molecular markers of risk). The authors thank Prof. Marina Camatini for the Andersen impactor and Giuseppe Merlino for helping in sampling surveys.

2.6 References

Allen HK, Donato J, Wang HH, Cloud-Hansen KA, Davies J and Handelsman J (2010) Call of the wild: antibiotic resistance genes in natural environments. *Nat Rev Microbiol* 8, 251–259.

Aminov RI and Mackie RI (2007) Evolution and ecology of antibiotic resistance genes. *FEMS Microbiol Lett* 271, 147–161.

Arthur M, Reynolds PE and Courvalin P (1996) Glycopeptide resistance in enterococci. *Trends Microbiol* 4, 401–407.

Baird-Parker AC (1962) An improved diagnostic and selective medium for isolating coagulase positive Staphylococci. *J Appl Bact* 25, 12–19.

Baquero F, Martínez JL and Canton R (2008) Antibiotics and antibiotic resistance in water environments. *Curr Opin Biotechnol* 19, 260–265.

Boreson J, Dillner AM and Peccia J (2004) Correlating bioaerosol load with PM_{2.5} and PM₁₀ concentrations: a comparison between natural desert and urban-fringe aerosols. *Atmos Environ* 38, 6029–6041.

Brodie EL, De Santis TD, Moberg Parker JP, Zubieta IX, Piceno YM and Andersen GL (2007) Urban aerosols harbor diverse and dynamic bacterial populations. *Proc Natl Acad Sci USA* 104, 299–304.

Broekema NM, Van TT, Monson TA, Marshall SA and Warshauer DM (2009) Comparison of cefoxitin and oxacillin Disk Diffusion Methods for detection of mecA mediated resistance in *Staphylococcus aureus* in a largescale study. *J Clin Microbiol* 47, 217–219.

Brooks GF, Butel JS and Morse SA (2001) Jawetz, Melnick, & Adelberg's Medical Microbiology, 22nd edn, p. 225 (Italian Edition). New York: McGraw-Hill.

Buttner MP, Willeke K and Grinshpun S (2002) Sampling and analysis of airborne microorganisms. In *Manual of Environmental Microbiology*, 2nd edn ed. Hurst, C.J., Crawford, R.L., Knudsen, G., McInerney, M. and Stetzenbach, L.D. pp. 814–820. Washington DC: ASM Press.

CLSI (2009) *Performance Standards for Antimicrobial Disk Susceptibility Tests; Approved Standard*, 10th edn. M02-A10, pp. 10–12; 27. Wayne: CLSI.

Dutka-Malen S, Evers S and Courvalin P (1995) Detection of glycopeptide resistance genotypes and identification to the species level of clinically relevant Enterococci by PCR. *J Clin Microbiol* 33, 24–27.

Fang Z, Ouyang Z, Zheng H, Wang X and Hu L (2007) Culturable airborne bacteria in outdoor environments in Beijing, China. *Microb Ecol* 54, 487–496.

Fierer N, Liu Z, Rodríguez-Hernandez M, Knight R, Henn M. and Hernandez, MT (2008) Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microbiol* 74, 200–207.

Griffin DW (2007) Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin Microbiol Rev* 20, 459–477.

Harrison RM, Jones AM, Biggins PDE, Pomeroy N, Cox CS, Kidd SP, Hobman JL, Brown NL et al. (2005) Climate factors influencing bacterial count in background air samples. *Int J Biometeorol* 49, 167–178.

Kummerer K (2009) Antibiotics in the aquatic environment – A review – Part II. *Chemosphere* 75, 435–441.

Lai K, Emberlin J and Colbeck I (2009) Outdoor environments and human pathogens in air. *Environ Health* 8(Suppl I), S15.

Leclercq R (2002) Mechanisms of resistance to macrolides and lincosamides: nature of the resistance elements and their clinical implications. *Clin Infect Dis* 34, 482–492.

Lee S, Lee H, Kim S, Lee HM, Kang H and Kim YP (2010) Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. *Sci Total Environ* 408, 1349–1357.

Livermore DM (2009) Has the era of untreatable infections arrived? *J Antimicrob Chemother* 64, i29–i36.

Marcazzan GM, Vaccaro S, Valli G and Vecchi R (2001) Characterisation of PM10 and PM2.5 particulate matter in the ambient air of Milan (Italy). *Atmos Environ* 35, 4639–4650.

Maron PA, Lejon DPH, Carvalho E, Bizet K, Lemanceau P, Ranjard L and Mougel C (2005) Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. *Atmos Environ* 39, 3687–3695.

Martineau F, Picard FJ, Lansac N, Menard C, Roy PH, Ouellette M and Bergeron MG (2000) Correlation between the resistance genotype determined by multiplex PCR assays and the antibiotic susceptibility patterns of *Staphylococcus aureus* and *Staphylococcus epidermidis*. *Antimicrob Agents Chemother* 44, 231–238.

Martínez JL (2008) Antibiotics and antibiotic resistance genes in natural environments. *Science* 321, 365–367.

Martínez JL (2009) Environmental pollution by antibiotics and by antibiotics resistance determinants. *Environ Pollut* 157, 2893–2902.

Polymenakou PN, Mandalakis M, Stephanou EG and Tselepidis A (2008) Particle size distribution of airborne microorganisms and pathogens during an intense African dust event in the Eastern Mediterranean. *Environ Health Perspect* 116, 292–296.

Sambrook J and Russell DW (2001) *Molecular Cloning – A Laboratory Manual*, 3rd edn, pp. A2.2; A8.9–A8.10. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Schwieger F and Tebbe CC (1998) A new approach to utilize PCR-single-strand-conformation polymorphism for 16S rRNA gene-based microbial community analysis. *Appl Environ Microbiol* 64, 4870–4876.

Septimus EJ and Kuper KM (2009) Clinical challenges in addressing resistance to antimicrobial drugs in the twenty-first century. *Clin Pharmacol Ther* 86, 336–339.

Strommenger B, Kettlitz C, Werner G and Witte W (2003) Multiplex PCR Assay for simultaneous detection of nine clinically relevant antibiotic resistance genes in *Staphylococcus aureus*. *J Clin Microbiol* 41, 4089–4094.

Zhang K, Sparling J, Chow BL, Elsayed S, Hussain Z, Church DL, Gregson DB, Louie T et al. (2004) New quadriplex PCR assay for detection of methicillin and mupirocin resistance and simultaneous discrimination of *Staphylococcus aureus* from coagulase-negative *Staphylococci*. *J Clin Microbiol* 42, 4947–4955.

CHAPTER 3

Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy

3.1 Introduction

Airborne particulate matter (PM) is a complex and dynamic mixture of components having different origins, both inorganic and biological, as well as different chemical and physical properties (Putaud et al, 2004). In the last decade a large number of studies have investigated the risks that PM poses both to human health and ecosystem functions (Samet et al, 2006), but most of these studies have typically focused on the non-biological fraction of suspended particles (Colbeck and Lazaridis, 2009). In recent years, interest has grown also in the biological fraction of PM (bioaerosols) (Jones and Harrison 2004), boosted by an increasing amount of evidence suggesting that this fraction may play a critical role in the effects of PM on biological systems (D'Amato, 2002; Finnerty et al., 2007; Jaenicke, 2005; Moorman et al, 2011). Bioaerosols have negative effects on human health and ecosystem functioning, by determining, for example, tissue inflammation (Camatini et al, 2010) and spreading of plant, livestock and human pathogens (Hirano and Upper, 1983; Pillai and Ricke, 2002). Such effects are exacerbated in urban areas, where the atmosphere is enriched in PM. Interest has therefore grown in the study of atmospheric pollution of urban areas, and recent studies have provided evidence for a larger impact of urban air pollution on human health than previously hypothesised (Gouveia and Maisonet, 2006).

Bioaerosols have a diverse composition. Fungal spores, pollens and dust mite allergens have been studied in detail (Beggs and Kerr, 2000), while the diversity of airborne bacteria and their potential sources remain poorly investigated (Bowers et al, 2009; 2011b; Maron et al, 2005; Maron et al, 2006; Womack et al, 2010). In addition, it has only recently been acknowledged that both the potential sources of PM and the possible dynamic of bacterial populations in the atmosphere may influence the spatial and temporal variation in airborne bacterial community structure (Franzetti et al, 2010). Therefore, there is a deep gap in the knowledge of the structure and dynamic of airborne bacterial communities, and in the ecological factors that drive their structure. It is also worth noting that it is even unclear

whether airborne bacteria are an actual ecological community, with the different populations growing and interacting in the atmosphere, or if they are simply a pool of organisms passively gathered together from different sources (Bowers et al, 2011b). For this reason, in this paper the term “bacterial community” is used to refer to the ensemble of bacterial groups present in the atmosphere, without necessarily implying the existence of dynamics and interactions between bacterial populations in this environment. Beside bacteria, PM of vegetable origin represents a considerable fraction of bioaerosols. *Chloroplasts* are revealed by the same analytical tools used for the analyses of airborne bacterial communities (e.g. massive parallel sequencing, see below), and their abundance provides information on the abundance of PM of vegetable origin. Plants, on the other hand, are one of the main sources of airborne bacteria, so that the relative abundance of *Chloroplast* sequences among bacterial ones may provide information on the potential sources of airborne bacteria (Brodie et al, 2007). For this reason, in the present work *Chloroplast* abundance was always included in the analyses, and the term “bacterial community” was used to refer in short to the ensemble of bacteria and *Chloroplasts* observed in a sample of PM, despite a portion of the community that was analysed was not bacteria.

Very little information is available regarding the total amount of bacteria present in the atmosphere at a given place and time, because this datum is technically difficult to determine, mainly due to limitations of microscopy techniques (Bowers et al, 2011a; 2011b). Culture-based microbiological methods have been the standard investigation procedure for a long time (Fang et al, 2007; Lighthart and Shaffer, 1995 a,b), despite the fact that it was well-recognised that they can capture only a small portion of the total microbial diversity (Pace, 1997; Peccia and Hernandez, 2006). As a result, the number of studies which have applied culture-independent techniques to the study of airborne bacteria is still limited. For example, such methods were used to assess the diversity of bacteria associated with small size particles during dust events (Polymenakou et al, 2008) and to characterise the spatial and temporal variations of bioaerosols in different geographical localities and/or different land uses (Angenent

et al, 2005; Bowers et al, 2009; 2011a; 2011b; Brodie et al, 2007; Franzetti et al, 2010; Frohlich-Nowoisky et al, 2009). However, the potential of culture-independent molecular methods, such as quantitative PCR (qPCR), rRNA gene sequencing and hybridisation (i.e. PhyloChip) and massive parallel sequencing technologies, for quantifying and describing the total diversity and composition of microbial communities has not yet been fully exploited (Bowers et al, 2011a; Rinsoz et al, 2008). Among these methods, the Illumina technology can reveal unprecedented diversity from even the most complex microbial environments (Claesson et al, 2010), and it is therefore a promising tool to investigate airborne bacterial communities and to contribute filling the gaps in the current knowledge of their structure and dynamics. In this work the Illumina technology and qPCR were used to provide a thorough description of airborne microbial communities of the Milan urban area (Northern Italy). Samples were taken in four ten-day sessions equally distributed across the different seasons in order to describe both short- and long-term variation in bacterial community structure and composition. The aims of the present study are therefore threefold: i) to deeply characterise the abundance and diversity of airborne microbial populations in this urban area; ii) to disentangle short-term (days) and seasonal variation of microbial communities; and iii) to identify the potential sources of the bacteria present in the atmosphere of this urban area.

3.2 Materials and methods

3.2.1 Sample collection and DNA extraction

Total suspended particulate matter (TSP) was sampled on quartz fibre filters (Whatman, Maidstone, England) with a high-volume sampler (ECHO HiVol, TCR TECORA, Milan, Italy) and a flux speed of 250 L min⁻¹. The sampling campaigns were carried out in 2010 in four ten-day periods in the four seasons as follows: 22 February–4 March; 24 May–3 June; 23 August–3 September; 24 November–4 December. Each collection period lasted 24 hours. The sampler was located in Milan, Italy (45°31'20''N, 9°12'45''E), in a typical urban area in the northern part of the city

near a high-traffic road. The characteristics of the sampling site have been previously described (chapter 2). The sampler was placed approximately 20 m from the nearest roads and 50 m from the nearest traffic lights. TSP sampling was performed at approximately 1.5 m from the ground.

Total bacterial DNA was extracted directly from a quarter of each filter using the FastDNA Spin for Soil kit (MP Biomedicals, Solon, OH, USA). Individual filter portions were cut into small pieces and loaded into the bead tube of the DNA extraction kit, after adding 1 M CaCO₃ in order to increase the pH, and then shaken at 200 rpm for 60 min. The remaining steps of the DNA extraction were performed according to the manufacturer's instructions.

3.2.2 Quantitative PCR (qPCR)

The abundance of airborne bacteria was estimated by the quantification of the number of copies of the gene 16S rRNA. This parameter is not directly related to cell number due to the presence of multiple ribosomal operons in the bacterial genomes; however it can be used to look at relative shifts in microbial biomass as the bias in copy number is likely constant across samples. A 466 bp fragment of the bacterial 16S rDNA (331-797 according to *E. coli* position, V3 and V4 hypervariable regions) was PCR-amplified with a universal primer set (Nadkarni et al, 2002). The PCR was performed in a total volume of 20 µL using the FluoCycleII Sybr reaction mix (Euroclone, Pero, Italy) with 0.3 µM (final concentration) forward and reverse primer. One µL out of the 85 µL of extracted from the filter was used as the template. The amplification was carried out under the following conditions: 95°C for 4 min, followed by 40 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 30 s, with acquisition of the fluorescence at the end of each 72°C elongation step. Genomic bacterial DNA used for standard concentration curves was extracted from pure cultures of *E. coli* K-12 substr. DH10B. The strain contains seven copies of the gene 16S rRNA (Accession number: NC_010473) (Durfee et al, 2008). The amount of extracted standard DNA was

measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). The standards and the samples were included in triplicate in each run.

3.2.3 16S rRNA fragment libraries by Illumina Genome Analyser Ix

The V5-V6 hypervariable regions of the 16S rRNA gene were PCR-amplified for Genome Analyser Ix sequencing. The PCR was performed in 3 x 75 µL volume reactions with @Taq® Hot start Taq polymerase (Euroclone, Pero, Italy), 4 mM MgCl₂, 0.2 mM dNTPs mix, and 1 µM of each primer. 783F and 1046R primers were used (Huber et al, 2007, Wang and Qian, 2009) and the cycling conditions were: initial denaturation at 94°C for 5 min; 29 cycles at 94°C for 50 s, 47°C for 30 s, and 72°C for 30 s and a final extension at 72°C for 5 min. At the 5' end of the 783F primer, one of eight 6 bp barcodes was also included to allow sample pooling and subsequent sequence sorting (see below and Table S1). The amplified products of 288 bp were purified with the Wizard SV PCR purification kit (Promega Corporation, Madison, WI, USA) and DNA quantity and purity was evaluated spectrophotometrically using a NanoDrop™ (Thermo Scientific, USA). Multiplexed samples were prepared with the Illumina Multiplexing Sample Preparation Oligonucleotide Kit, which provides 12 index oligos for pooling up to 12 samples per lane. Therefore, our customised tagging system with 8 barcodes at the 5' end of the 783F primer allowed pooling up to 96 samples per lane. Sequencing of all the pooled samples was performed on a single Illumina GA-Ix lane, using a paired-end 76 bp protocol and the 4.0 sequencing chemistry. Illumina Real Time Analysis software version 1.8.7 was used to perform the cluster extraction and base-calling processing analyses. The Illumina GA-Ix sequencing was carried out at the Department of Clinical and Preventive Medicine, University of Milano-Bicocca, Italy.

3.2.4 Sequence analyses

Each sequence was assigned to its original sample according to its index oligos and barcode. After sorting, the reverse read of each paired-end sequence was reverse-

complemented and merged with the corresponding forward read, inserting 20 Ns in between (Claesson et al, 2010). A quality cut-off was applied in order to remove sequences i) that did not contain the barcode, and ii) with an average base quality value (Q) lower than 30. The barcode was removed from sequences before further processing. The source code used for sequence processing is available upon request to the authors.

The taxonomic attribution of filtered sequences was carried out using the stand-alone version of RDP Bayesian Classifier, using a 50% confidence level as suggested for sequences shorter than 200 bp (Claesson et al, 2009; Wang et al, 2007). Operational Taxonomic Units (OTUs) were defined by the RDP classifier considering the fourth taxonomic level, which in most cases corresponded to Order. Recent papers demonstrated that short sequences of hypervariable regions of 16S rRNA gene provide enough variability to capture the differences in the structure of bacterial communities (Caporaso et al, 2012; Claesson et al, 2009; Jeraldo et al, 2011).

3.2.5 Environmental factors

3.2.5.1 Meteorological data and particulate matter (PM) concentration

Daily average temperature, wind speed at ground level, and relative humidity for Milan in the nearest meteorological station in the city, as well as the concentrations of PM smaller than 10 μm (PM10) and smaller than 2.5 μm (PM2.5) in three different urban stations during sampling periods were retrieved from the Agenzia Regionale per la Protezione dell'Ambiente (Regional Agency of Environmental Protection, ARPA-Lombardia) and are available at: <http://ita.arpalombardia.it/meteo/dati/riciesta.asp>.

3.2.6 Analysis of inorganic ions

A portion of the filter used to collect PM was dissolved in MilliQ® water in an ultrasonic bath and analysed by ion chromatography (IC, Dionex DX500 system) to determine the water soluble inorganic ion concentration (Na^+ , NH_4^+ , Mg^{2+} , Ca^{2+} , F^- , Cl^- , NO_3^- , SO_4^{2-}) as described by Squizzato et al. (2012).

3.2.7 Statistical analyses

The concentrations of PM₁₀ measured on the same day at three different weather stations in Milan were highly repeatable ($r = 0.953$, $F_{39,72} = 57.70$, $P < 0.001$) and were therefore averaged within each day. Missing data on the concentration of PM_{2.5} were imputed by regressing PM_{2.5} concentrations on PM₁₀ concentrations (see SI - § 1.1.1 for further details).

A hierarchical cluster analysis of the relative abundance of OTUs was performed with the HCLUST procedure in R 2.8.1 (R Development Core Team, 2008) with the complete linkage method on the Hellinger distance between sites. This distance depends on the difference in the proportion of OTUs between samples, decreases the importance of OTU abundance over occurrence and avoids the double-zero problem when comparing species composition between samples (De Cáceres et al, 2010; Legendre and Legendre, 1998).

Principal Component Analysis (PCA) on Hellinger-transformed community data (Legendre and Gallagher, 2001) was used to visualise the changes in airborne bacterial community among seasons. Similarly, changes in meteorological data, and ion and PM concentration in the atmosphere (hereafter “air and PM factors” for brevity) were visualised by means of PCA on normalised variables. PCA analyses were performed with the BIODIVERSITYR package (Kindt and Coe, 2005) in R.

The Analysis of Similarity (ANOSIM) was used to assess the difference in OTU composition between samples collected in different seasons. Pairwise ANOSIM was used to investigate differences in airborne microbial communities between seasons. P-values of pairwise ANOSIM were corrected according to the Bonferroni procedure to account for multiple statistical tests. Similarly, differences in air and PM factors between seasons were assessed with ANOSIM tests on the Euclidean distance between normalised variables. ANOSIM was performed with the VEGAN package (Oksanen et al, 2009) in R. See SI - § 1.1.2 for further information about ANOSIM.

Hellinger-transformed OTU relative abundances were related to air and PM factors by multivariate regression tree (MRT) analysis using the MVPART package (Therneau et al, 2007) in R. The best tree was chosen as the smallest tree whose cross-validated relative error (CVRE) was within one standard error of the CVRE of the best tree (Borchard et al, 2011). The difference in OTU relative abundance among groups identified by the MRT analysis was tested by ANOVA. P values were corrected for multiple testing according to the False Discovery Rate (FDR) procedure (Benjamini and Yekutieli, 2001) using the MULTTEST package in R.

Short-term variation in the structure of bacterial communities and in air and PM factors was investigated on the one hand by Mantel tests and Mantel correlograms between the matrix of Hellinger distances of bacterial communities or by the matrix of Euclidean distances between normalised air and PM factors, and on the other hand by the difference in days between samples taken during each ten-day period (see SI - § 1.1.3 for further details). Goodness-of-fit Mantel tests (*sensu* Legendre and Legendre, 1998) were used to compare variation in bacterial community structure among samples taken in the same season (within-season variation) and among samples taken in different seasons (between-season variation).

Mantel tests were performed using the ECODIST package (Goslee and Urban, 2007) in R, and P-values were corrected according to the FDR procedure.

3.2.8 Sequence data

De-multiplexed fastq-formatted DNA sequences are available on Sequence Read Archive (SRA) - Study accession number ERP001381 <http://www.ebi.ac.uk/ena/data/view/ERP001381>.

3.3 Results

3.3.1 Quantification of airborne bacteria

The mean number of copies of the 16S rRNA gene, which was used here as a measure of microbial biomass (see Methods), was $1.7 (0.3 \text{ SE}) \times 10^5$ ribosomal operons m^{-3} of sampled air. An ANOVA model disclosed significant differences in microbial biomass among seasons ($F_{3,26} = 5.11$, $P = 0.005$; Figure 1), and a post-hoc test (Tukey method) revealed that winter samples contained significantly less bacterial biomass than spring ($t_{36} = -2.91$, $P = 0.030$) and summer ones ($t_{36} = -3.70$, $P = 0.004$).

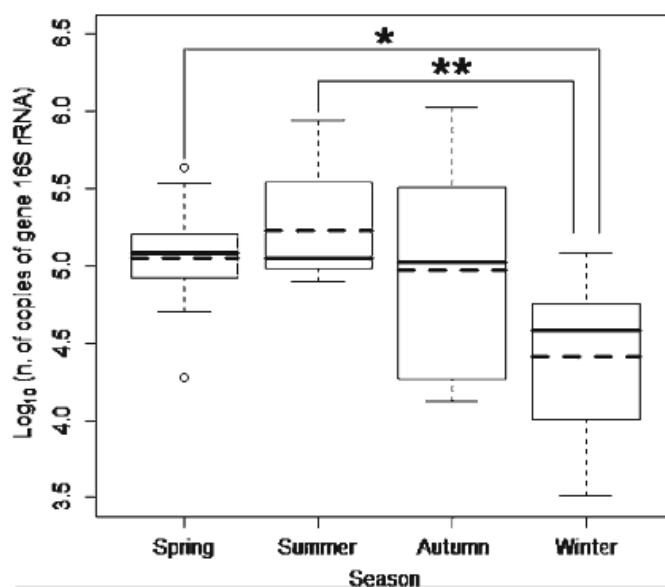


Figure 1. Box plot diagram of the TSP-associated number of copies of gene 16S rRNA, obtained by qPCR in the four seasons. Solid lines within boxes represent median values; dashed lines mean values. Lines with asterisks highlight significant difference between seasons (* = $P < 0.05$; ** = $P < 0.01$).

3.3.2 Microbial community structures

A total of 271,587 sequences that passed the quality filter were obtained from sequencing. The number of sequences of each sample ranged from 765 to 26,187.

A total of 107 OTUs were detected across the 40 collected samples. Among them, 16 OTUs represented more than 1% of the total number of the obtained sequences in at

least one sample (most abundant OTUs). In each sample, 26.7% to 39.0% of the sequences could not be classified and were discarded from subsequent analyses.

The taxonomic classification of bacteria based on the fourth taxonomic level is shown in Figure 2. Each community was dominated by few taxa throughout the year, with a marked prevalence of *Actinobacteridae*, *Clostridiales*, *Sphingobacteriales* and a few proteobacterial orders (*Burkholderiales*, *Rhizobiales*, *Sphingomonadales*, *Pseudomonadales*). A relevant fraction of *Chloroplasts* were also detected in all samples.

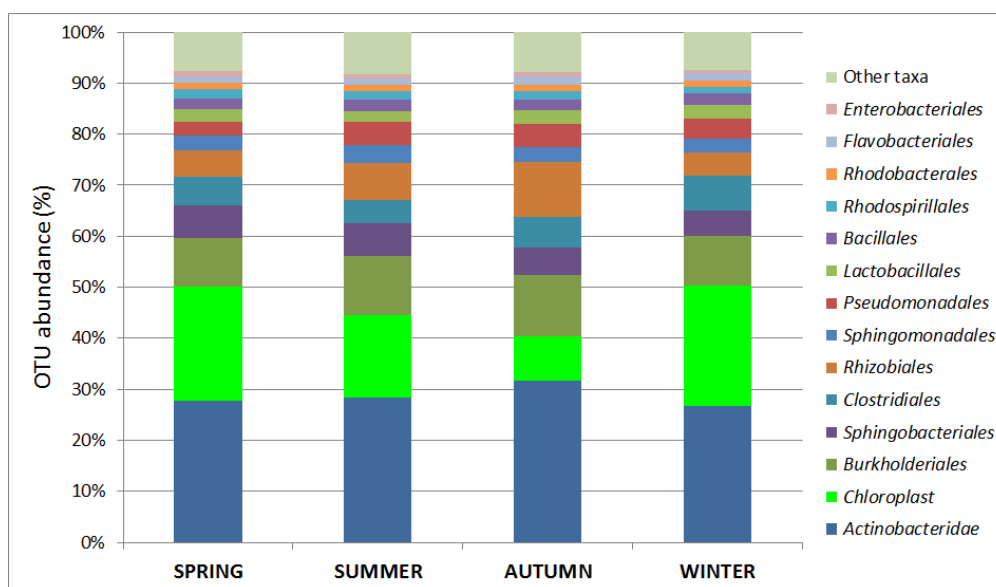


Figure 2. Taxonomic classification of the sequences using an RDP Bayesian classifier (50% confidence) at the fourth taxonomic level. Unclassified sequences were omitted from the figure.

On average, there was no significant difference either in the number of OTUs detected in different seasons ($F_{3,36} = 0.05$, $P = 0.99$) or in the Shannon–Weaver diversity index calculated from the OTU number detected at each sample ($F_{3,36} = 1.41$, $P = 0.26$) (Table 1).

Chapter 3

Table 1: Mean (SE) value of the quantitative environmental variables in the different seasons.

| Variable | Winter | Spring | Summer | Autumn |
|--|---------------|--------------|--------------|--------------|
| # OTUs | 52.0 (4.68) | 51.8 (4.26) | 50.0 (4.31) | 50.8 (4.10) |
| Shannon–Weaver diversity index | 2.30 (0.05) | 2.31 (0.06) | 2.40 (0.04) | 2.39 (0.03) |
| daily average temperature (°C) | 7.57 (0.56) | 20.21 (0.52) | 21.81 (0.87) | 2.33 (0.38) |
| wind speed (m s ⁻¹) | 1.51 (0.16) | 2.28 (0.28) | 2.22 (0.36) | 1.38 (0.11) |
| relative humidity (%) | 83.14 (4.16) | 56.15 (5.16) | 61.66 (5.4) | 88.85 (2.97) |
| PM10 concentration (µg m ⁻³) | 45.03 (6.87) | 25.58 (2.94) | 22.18 (2.29) | 38.48 (3.96) |
| PM2.5 concentration (µg m ⁻³) | 33.88 (6.53)* | 13.80 (2.48) | 9.60 (1.47) | 22.00 (2.73) |
| F ⁻ concentration (µg m ⁻³) | 0.24 (0.16) | 0.06 (0.02) | 0.03 (0.01) | 0.07 (0.02) |
| Cl ⁻ concentration (µg m ⁻³) | 0.63 (0.09) | 0.11 (0.05) | 0.24 (0.08) | 1.07 (0.31) |
| NO ₃ ⁻ concentration (µg m ⁻³) | 11.43 (2.37) | 9.03 (2.65) | 1.28 (0.2) | 10.02 (3.8) |
| SO ₄ ⁻ concentration (µg m ⁻³) | 1.46 (0.19) | 2.29 (0.46) | 2.21 (0.57) | 1.56 (0.21) |
| Na ⁺ concentration (µg m ⁻³) | 0.43 (0.03) | 0.49 (0.1) | 0.58 (0.09) | 0.92 (0.29) |

Chapter 3

| | | | | |
|--|-------------|-------------|-------------|-------------|
| NH ₄ ⁺ concentration (μg m ⁻³) | 3.85 (0.60) | 2.11 (0.69) | 1.00 (0.30) | 2.23 (0.48) |
| K ⁺ concentration (μg m ⁻³) | 0.30 (0.04) | 0.16 (0.02) | 0.13 (0.01) | 0.37 (0.06) |
| Mg ⁺⁺ concentration (μg m ⁻³) | 0.21 (0.04) | 0.15 (0.02) | 0.18 (0.02) | 0.12 (0.03) |
| Ca ⁺⁺ concentration (μg m ⁻³) | 2.66 (0.45) | 2.16 (0.26) | 2.43 (0.15) | 1.94 (0.49) |

* 2 missing data

The clustering analysis based on the relative abundance of the OTUs showed two main clusters (Figure S1). Qualitatively similar results could be obtained from an analysis restricted to the most abundant OTUs (Figure S2). The OTU clustering separated the spring (cluster A – Figure S2) and autumn samples (cluster B – Figure S2) into different clusters, while the winter and summer samples were scattered in the two different clusters.

On the PCA plot, the points representing the microbial communities of autumn and spring samples formed clearly separated clusters, whereas those of winter and summer samples largely overlapped with each other and with both the clusters of autumn and spring samples (Figure 3A). Almost identical results were obtained from the PCA restricted to the most abundant OTUs (Figure S3).

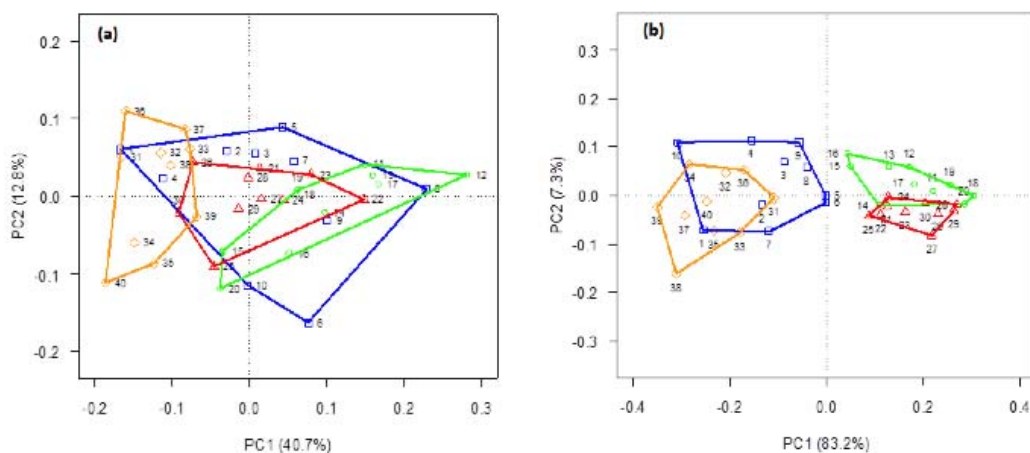


Figure 3. Two-dimensional PCA plot of (a) microbial communities and (b) air and PM factors at different seasons (squares and solid line: winter; circles and dashed line: spring; triangles and dotted line: summer; diamonds and dashed-dotted line: autumn). The percentage of variance explained by each axis is shown.

The ANOSIM showed overlapping, albeit significantly different, microbial communities between the four seasons, both in the analysis including all of the OTUs (global $R = 0.255$, $P < 0.001$) and in that restricted to the most abundant OTUs (global $R = 0.283$, $P < 0.001$). Pairwise tests showed that autumn communities were different, albeit with some overlap from spring communities ($R = 0.669$, $P_{\text{Bonf}} = 0.001$), and overlap but

were still separable, from summer and winter communities ($R \geq 0.361$, $P_{\text{Bonf}} \leq 0.002$). Conversely, spring, summer and winter communities were practically inseparable ($R \leq 0.116$, $P_{\text{Bonf}} \geq 0.284$). Identical results were obtained in post-hoc tests conducted on the analysis restricted to the most abundant OTUs (details not shown).

In general, all physical characteristics of air and PM differed between seasons (Table 1). The PCA plot based on standardised air and PM factors showed that the autumn and winter samples were well separated from the spring and summer samples. This interpretation was supported by ANOSIM ($R = 0.692$, $P = 0.001$). In fact, no significant difference in air and PM factors was found between the winter and autumn samples ($R = 0.252$, $P_{\text{Bonf}} = 0.056$), or between the spring and summer ones ($R = 0.153$; $P_{\text{Bonf}} = 0.256$). All other comparisons showed clearly separated groups ($R \geq 0.718$, $P_{\text{Bonf}} \leq 0.001$).

MRT analysis showed that airborne microbial communities were separated mainly by the average daily temperatures. Samples collected on days with an average temperature $<+6.32^{\circ}\text{C}$ formed a separate cluster with respect to those collected on days with temperature $\geq+6.32^{\circ}\text{C}$. This classification separated all of the autumn samples and three of the winter samples (1, 2 and 7; Table S2) from the others. The MRT analysis restricted to the most abundant OTUs consistently showed separation of samples according to a mean daily temperature higher or lower than $+6.32^{\circ}\text{C}$ (Figure 4 and Figure S4).

In particular, the abundance of two OTUs significantly differed between samples collected on days with temperature higher or lower than $+6.32^{\circ}\text{C}$ ($P_{\text{FDR}} = 0.008$, both in the analysis conducted on all of the OTUs and in that restricted to the most abundant ones): *Actinobacteridae* were significantly more abundant in samples collected on cold days, while *Chloroplasts* were significantly more abundant in samples collected on warm days (Figure 5).

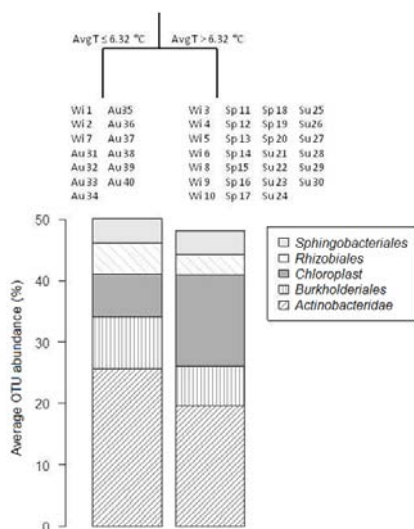


Figure 4. MRT analysis of OTU abundance and air and PM factors. Bar-plots show the average abundance of the five most common OTUs in samples, which accounts for more than 3% of sequences, collected on days with average temperature lower (left bar) or higher (right bar) than +6.32°C. Bar-plots showing the average abundance of all the common OTUs are shown in Fig. S3.

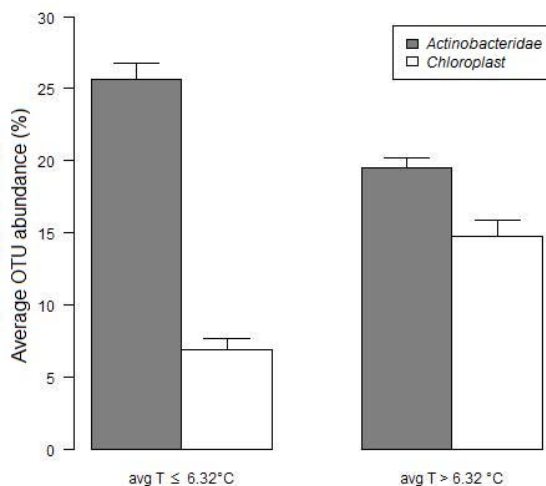


Figure 5. Average relative abundance of the OTUs that significantly differed between the clusters of samples identified by MRT analysis. Bars represent standard errors.

3.3.3 Short-term temporal variation of community structure

Mantel tests revealed no significant correlation of Hellinger distances between communities and the number of days between samplings within the same season ($r_M \leq 0.215$, $P_{FDR} \geq 0.199$). Similarly, Mantel correlograms disclosed no significant temporal autocorrelation between samples taken at different time-lags within the same season ($|r_M| \leq 0.441$, $P_{FDR} \geq 0.067$). The same analyses conducted for Euclidean distances between normalised air and PM factors revealed no significant correlation with the number of days between samplings ($r_M \leq 0.197$, $P_{FDR} \geq 0.223$), nor significant temporal autocorrelation at any time-lag ($|r_M| \leq 0.367$, $P_{FDR} \geq 0.275$).

A goodness-of-fit Mantel test indicated that bacterial communities within a single season were significantly more similar (less distant) than those sampled in different seasons ($r_M = -0.206$, $P < 0.001$). In particular, goodness-of-fit Mantel tests comparing the distance between communities sampled within each season and those sampled in all the others, indicated that significance of the overall test arose because summer communities were significantly more similar to each other than to those sampled in the other seasons (summer: $r_M = -0.225$, $P_{FDR} < 0.001$). Conversely, communities sampled during the other seasons were equally dissimilar to those collected during the same or in other seasons ($|r_M| \leq 0.123$, $P_{FDR} \geq 0.119$).

3.4 Discussion

3.4.1 Seasonal variability of airborne bacterial abundance

In this study we determined the abundance of airborne bacteria in an urban area of Northern Italy across the four seasons by a culture-independent technique, namely qPCR. The observed values ranged from 3.3×10^3 to 1.0×10^6 ribosomal operons per m^3 (Figure 1) with a significantly lower mean abundance in winter than in spring and summer. These values are much larger than those reported in the only other study that used qPCR technique on the same PM dimensional fraction by Lee et al. (2010) in Seoul (South Korea), and in previous culture-based reports from urban areas

(Bovallius et al, 1978; Fang et al, 2007). However, these latter values may have been underestimated due to cultivation biases (Peccia and Hernandez, 2006). In fact, higher values of bacterial concentration were obtained in more recent studies that applied culture-independent methods (Bowers et al, 2009, 2011a, 2011b; Lee et al. 2010, Maron et al, 2005). However, the limited number of these studies, the use of different techniques (qPCR or direct microscopic count), and the different PM fractions investigated, hampered direct comparisons with the results from the present study. In addition, in those studies, no relationship between bacterial abundance, season and general features of investigated sites was observed. Only recently, Bowers et al. (2011b) observed a significantly higher abundance of bacteria in summer than in winter for PM_{2.5} in Detroit and Cleveland (USA), where the annual variation in temperature is generally similar to that in Milan. It could be speculated that the lower abundance of bacteria in winter might therefore be due to a reduced release from plant, soil and water sources, and to climatic conditions that are less favourable for bacterial growth during the winter season both in Milan and in the U.S. Midwestern cities.

3.4.2 Seasonal and short-term variability of community structure

Differences were found in the structure of bacterial communities sampled across different seasons. In particular, microbial communities were similar during the winter, spring and summer, while the autumn communities were different from those observed in the other seasons. In fact, the ANOSIM clearly separated the communities of samples collected in the coldest period (Autumn, 24th November–4th December) from those of other seasons. In addition, MRT analysis disclosed a significant difference in the structure of microbial communities sampled on days when the mean temperature was above or below +6.32°C, with more *Actinobacteridae* and less Chloroplasts in samples collected on days colder than +6.32°C. It is worth noting that all of the autumn days in this dataset had a lower mean temperature than +6.32°C. These findings are in agreement with the results of

most of the previous studies conducted in different parts of the world, which consistently showed seasonal variability in bacterial communities (Bowers et al, 2011a; Bowers et al, 2011b; Brodie et al, 2007; Fierer et al, 2008; Polymenakou et al, 2008). However, very few studies were able to identify a single meteorological factor that affects the structure of microbial communities (see e.g. Brodie et al, 2007; Maron et al, 2006). This suggests that the seasonal variability in bacterial communities is probably due to the combined effects of differences in several meteorological factors, the chemical composition of particulate matter and the relative importance of the main sources of the air particles, rather than to temperature *per se*.

Over a shorter time-scale, no indication of temporal autocorrelation was found in the structure of bacterial communities. In addition, only summer communities differed from each other less than those collected in other seasons, while communities sampled during the other seasons were equally dissimilar to those collected during the same or in other seasons. The similar structure of summer communities might be due to the stability of the air and PM factors in this season, which also appears to be true from the PCA plot. Indeed, in this plot, which explained more than 90% of the total variance, summer samples were closer to each other than those of the other seasons (Figure 3B). The large variation in bacterial community composition even within a few days observed in the other season is, however, not surprising, as previous studies reported an even larger variability both intra-day and intra-week (Fierer et al, 2008; Maron et al, 2006). Airborne bacterial communities seem therefore to be influenced mainly by the predominant source of bacteria rather than by an intrinsic ecological dynamic, as they do not seem to show an ecological succession with time. Alternatively, temporal variations in community structure, if any, may occur over longer time periods than the 10-days investigated in the present study.

3.4.3 Potential sources of bacteria and differences in the structure of bacterial communities associated with different PM size fractions

The observed seasonal differences in bacterial communities and a general knowledge of the ecology of the dominant bacterial taxa allowed speculation about the relative importance of different potential sources of airborne bacteria in different seasons (Fierer et al. 2008; Franzetti et al. 2010). Plants are probably the main source of bacteria found in the atmosphere in warm seasons, as suggested by the abundance of plant-associated bacteria, such as *Sphingomonadales* and *Rhizobiales*, and the high number of *Chloroplast* sequences retrieved in samples collected in warm periods. Conversely, soil-inhabiting bacteria, such as *Actinobacteriales*, *Pseudomonadales* and *Burkholderiales*, prevailed in the atmosphere during cold seasons.

The observed variability in the composition of bacterial communities was lower than that reported in a previous study on the finer PM fractions (Franzetti et al, 2010). In fact, a strong seasonal shift in bacterial community composition had been observed on PM₁₀ and PM_{2.5} in the same sampling location in Milan (Franzetti et al, 2010). In the present study, which in contrast is focused on TSP, the seasonal differences were less exacerbated, thus suggesting that airborne bacterial communities associated with different dimensional fractions of PM may differ. In fact, in the present study, which was focused on TSP, *Actinobacteridae* were the most abundant taxon and *Burkholderiales* were among the dominant taxa throughout the whole year (Figure 2). On the contrary, in the previous study, which focused on PM₁₀ and PM_{2.5}, *Actinobacteridae* were the dominant taxon only in the winter and were observed at a very low abundance in the summer. In addition, only a few *Burkholderiales* were found in both seasons while *Sphingomonadales*, *Sphingobacteriales* and *Rhizobiales* were the most abundant taxa in summer. Hence, analyses of TSP and finer PM fractions (<10 µm) showed different seasonal patterns of variation in the most abundant bacterial taxa. This suggests that the bacteria associated with particles larger than 10 µm, which are usually excluded from the analyses in all the works that consider PM₁₀ only, significantly contribute to the structure of the entire airborne

bacterial community. *Sphingomonadales* and *Rhizobiales* are typical plant-associated bacteria, while *Actinobacteridae* and *Burkholderiales* are typically found in soil. It can therefore be supposed that the relative contribution of soil and plant sources differs for bacteria associated with particles larger than 10 μm with respect to those associated with PM₁₀. Admittedly, these speculations are only qualitative and a quantitative assessment of source contribution to the airborne communities as well as a precise knowledge of the quantitative distribution of bacteria between fine (<10 μm) and coarse (>10 μm) particles would have allowed a more detailed interpretation of these findings. Unfortunately, this information was unavailable in the present study.

In conclusion, in recent years, encouraging progress in next-generation sequencing techniques has led to an increase in our comprehension of the structure and diversity of airborne bacterial communities. The results of the present study confirmed that airborne bacterial communities showed a large variability even between consecutive days, thus suggesting that their structure mainly depends on the source of bacteria that predominates in a given period of time. In addition, a comparison of the results of the present and of a previous study conducted in the same area indicates that different bacterial communities could be found in different PM fractions, thus suggesting that a detailed investigation of the structure, temporal variability, and potential sources of bacterial communities associated to PM of different size may disclose further insights into the ecology of bioaerosols.

3.5 Acknowledgements

This work was partially supported by CARIPLO FOUNDATION (Milano - ITALY) in the frame of the project TOSCA (Toxicity of particulate matter and molecular markers of risk). The authors thank Rocco Piazza and Alessandra Pirola for DNA sequencing and bioinformatics support.

3.6 References

Angenent LT, Kelley ST, St Amand A, Pace NR, Hernandez MT (2005) Molecular identification of potential pathogens in water and air of a hospital therapy pool. *Proc Natl Acad Sci USA* 102, 4860–4865.

Beggs CB and Kerr KG (2000) The threat posed by airborne micro-organisms. *Indoor Built Environ* 9, 241–245.

Benjamini Y and Yekutieli (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29, 1165–1188.

Borchard D, Gillet F, Legendre F (2011) *Numerical Ecology with R*. Springer: New York, doi: 10.1007/978-1-4419-7976-6.

Bovallius A, Bucht B, Roffey R, Anas P (1978) Three-year investigation of the natural airborne bacterial flora at four localities in Sweden. *Appl Environ Microb* 35, 847–852.

Bowers RM, Lauber CL, Wiedinmyer C, Hamady M, Haller AG, Fall R, Knight R, Fierer N (2009) Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. *Appl Environ Microb* 38, 6029–6041.

Bowers RM, McLetchie S, Knight R, Fierer N (2011a) Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *ISME J* 5, 601–612.

Bowers RM, Sullivan AP, Costello EK, Collet JL, Knight R, Fierer N (2011b) Source of bacteria in outdoor air across cities in the midwestern United States. *Appl Environ Microbiol* 77, 6350–6356.

Brodie EL, DeSantis TZ, Parker JP M, Zubietta IX, Piceno YM, Andersen GL (2007) Urban aerosols harbour diverse and dynamic bacterial populations. *Proc Natl Acad Sci USA* 104, 299–304.

Camatini M, Corvaja V, Pezzolato E, Mantecca P, Gualtieri M (2010) PM10-biogenic fraction drives the seasonal variation of pro-inflammatory response in A549 cells. *Environ Toxicol* 27, 63–73.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621-1624.

Claesson MJ, O'Sullivan O, Wang Q, Nikkila J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4, e6669.

Claesson MJ, Wang QO, O'Sullivan O, Greene-Diniz R, Cole JR, Ross P, O'Toole PW (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Res* 38, e200.

Colbeck I, Lazaridis M (2009) Aerosols and environmental pollution. *Naturwissenschaften* 97, 117–131.

D'Amato G (2002) Environmental urban factors (air pollution and allergens) and the rising trends in allergic respiratory diseases. *Allergy* 57, 30–33.

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119, 1674–1684.

Chapter 3

Durfee T, Nelson R, Baldwin S, Plunkett III G, Burland V, Mau B, Petrosino JF, Qin X, Muzny DM, Ayele M, Gibbs RA, Csörgo B, Pósfai G, Weinstock GM, Blattner FR (2008) The complete genome sequence of *Escherichia coli* DH10B: Insights into the biology of a laboratory workhorse. *J Bacteriol* 190, 2597-2606.

Fang ZG, Ouyang ZY, Zheng H, Wang XK, Hu LF (2007) Culturable airborne bacteria in outdoor environments in Beijing, China. *Microb Ecol* 54, 487–496.

Fierer N, Liu Z, Rodriguez-Hernandez M, Knight R, Henn M, Hernandez MT (2008) Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microb* 74, 200–207.

Finnerty K, Choi JE, Lau A, Davis-Gorman G, Diven C, Seaver N, Linak WP, Witten M, McDonagh PF (2007) Instillation of coarse ash particulate matter and lipopolysaccharide produces a systemic inflammatory response in mice. *J Toxicol Environ Health A* 70, 1957–1966.

Franzetti A, Gandolfi I, Gaspari E, Ambrosini R, Bestetti G (2010) Seasonal variability of bacteria in fine and coarse urban air particulate matter. *Appl Microbiol Biotechnol* 90, 745–753.

Frohlich-Nowoisky J, Pickersgill DA, Despres VR, Poschl U (2009) High diversity of fungi in air particulate matter. *Proc Natl Acad Sci USA* 106, 12814–12819.

Gandolfi I, Franzetti A, Bertolini V, Gaspari E, Bestetti G (2011) Antibiotic resistance in bacteria associated to coarse atmospheric particulate matter in an urban area. *J Appl Microbiol* 110, 1612–1620.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *J Stat Soft* 22, 1–19.

Gouveia NC, Maisonet M (2006) Health effects of air pollution: an overview. In: Air Quality Guidelines: Global Update 2005. World Health Organization: Copenhagen, Denmark, pp 87–109.

Hirano SS, Upper CD (1983) Ecology and epidemiology of foliar bacterial plant pathogens. *Annu Rev Phytopathol* 21, 243–269.

Huber J, Mark Welch D, Morrison H, Huse S, Neal P, Butterfield D, Sogin M (2007) Microbial population structures in the deep marine biosphere. *Science* 318, 97-100.

Jaenicke R (2005) Abundance of cellular material and proteins in the atmosphere. *Science* 308, 73.

Jones AM, Harrison RM (2004) The effects of meteorological factors on atmospheric bioaerosol concentrations—a review. *Sci Total Environ* 326, 151–180.

Kindt R, Coe R (2005) Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre (ICRAF): Nairobi, Kenya.

Lee SH, Lee HJ, Kim SJ, Lee HM, Kang H, Kim YP (2010) Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. *Sci Total Environ* 408, 1349–1357.

Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia* 129, 271–280.

Legendre P, Legendre L (1998) Numerical ecology (2nd ed). Elsevier: Amsterdam, The Netherlands.

Lighthart B, Shaffer BT (1995a) Viable bacterial aerosol particle size distributions in the midsummer atmosphere at an isolated location in the high desert chaparral. *Aerobiologia* 11, 19–25.

Lighthart B, Shaffer BT (1995b) Airborne bacteria in the atmosphere surface layer: temporal distribution above a grass seed field. *Appl Environ Microbiol* 61, 1492–1496.

Maron PA, Lejon DPH, Carvalho E, Bizet K, Lemanceau P, Ranjard L, Mougel C (2005) Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. *Atmos Environ* 39, 3687–3695.

Maron PA, Mougel C, Lejon DPH, Carvalho E, Bizet K, Marck G, Cubito N, Lemanceau P, Ranjard L (2006) Temporal variability of airborne bacterial community structure in an urban area. *Atmos Environ* 40, 8074–8080.

Moorman JE, Zahran H, Truman BI, Molla MT (2011) Current asthma prevalence—United States, 2006–2008. *MMWR Surveill Summ* 60(Suppl.) 84–86.

Nadkarni MA, Martin FE, Jacques NA, Hunter N (2002) Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* 148, 257–266.

Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson GL, Solymos P, Stevens MHH, Wagner H (2009) *Vegan: Community Ecology Package*. R package version 1.15-3. <http://cran.r-project.org/>, <http://vegan.r-forge.r-project.org/>

Pace NR (1997). A molecular view of microbial diversity and the biosphere. *Science* 276, 734–740.

Peccia J, Hernandez M (2006) Incorporating polymerase chain reaction-based identification, population characterization and quantification of microorganisms into aerosol science: a review. *Atmos Environ* 40, 3941–3961.

Pillai SD, Ricke SC (2002) Bioaerosols from municipal and animal wastes: background and contemporary issues. *Can J Microbiol* 48, 681–696.

Polymenakou PN, Mandalakis M, Stephanou EG, Tselepides A (2008) Particle size distribution of airborne microorganisms and pathogens during an intense African dust event in the Eastern Mediterranean. *Environ Health Perspect* 116, 292–296.

Putaud J, Raes F, Van Dingenen R, Brüggemann E, Facchini M, Decesari S, Fuzzi S, Gehrig R, Hüglin C, Laj P, Lorbeer G, Maenhaut W, Mihalopoulos N, Müller K, Querol X, Rodriguez S, Schneider J, Spindler G, Ten Brink H, Tørseth K & Wiedensohler A (2004) A European aerosol phenomenology - 2: Chemical characteristics of particulate matter at kerbside, urban, rural and background sites in Europe. *Atmos Environ* 38, 2579-2595.

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria.

Rinsoz T, Duquenne P, Greff-Mirguet G, Oppliger A (2008) Application of real-time PCR for total airborne bacterial assessment: comparison with epifluorescence microscopy and culture-dependent methods. *Atmos Environ* 42, 6767–6774.

Samet JM, Brauer M, Schlesinger R (2006) Particulate Matter. In: *Air Quality Guidelines: Global Update 2005*. World Health Organization: Copenhagen, Denmark, pp 217–305.

Squizzato S, Masiol M, Innocente E, Pecorari E, Rampazzo G, Pavoni B (2012) A procedure to assess local and long-range transport contributions to PM_{2.5} and secondary inorganic aerosol. *Aerosol Science* 46, 64–76.

Therneau TM, Atkinson B, Ripley B, Oksanen J, De'ath G (2007) mvpart: Multivariate partitioning. R package version 1.2-6. <http://cran.r-project.org/web/packages/mvpart/index.html>.

Chapter 3

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microb* 73, 5261–5267.

Wang Y and Qian P (2009). Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. *PLoS ONE* 4, e7401.

Womack AM, Bohannon BJM, Green JL (2010) Biodiversity and biogeography of the atmosphere. *Philos Trans R Soc B Biol Sci* 365, 3645–3653.

CHAPTER 4

**Spatio-temporal variability and community-
assembly processes of PM10 associated
bacterial communities across two cities of
Northern Italy**

4.1 Introduction

Microorganisms are ubiquitous on the Earth, they can potentially be found in every environment (O'Malley, 2008). Although the atmosphere could be considered an extreme environment for microbes due to its low nutrient concentration, high solar radiation, etc., bacteria and fungi have been detected in various atmospheric layers (Griffin, 2004; Wainwright et al, 2003).

In recent years, the study of airborne bacteria had gained increasing interest due to the emerging evidences that microorganisms act as drivers of meteorological changes and diseases (Deguillaume et al, 2008; Jie et al, 2011).

The works on this topic gave a significant contribution to the comprehension of microbial dynamics and variability. For example, the changes in airborne microbial community structure across time and space were investigated (Bowers et al, 2009; 2012; Fahlgren et al 2010; Fierer et al, 2008b Franzetti et al, 2011). Different environments and conditions were taken in account: urban (Bowers et al, 2011b; Brodie et al, 2007; chapter 3; Maron et al, 2006) rural and different land-use types (Bowers et al, 2011a), upper troposphere (DeLeon-Rodriguez et al, 2013). Therefore, also sources of bacteria were investigated. Soil, leaf surfaces and in some cases animal feces were retrieved as the most relevant (Bowers et al, 2011b; Bowers et al, 2013; chapter 3). Finally, the bacterial abundance was assessed using different methods such as direct plate count (Lighthart, 1997; Peccia and Hernandez, 2006), epifluorescence microscopy (Maron et al, 2005) and quantitative PCR (qPCR) (Lee et al, 2010; chapter 3). Despite these evidences, the assessment of diversity and spatio-temporal variability of airborne microorganisms remains still under-investigated and the main attention went to potentially harmful bacteria deriving from punctual sources of contamination (i.e. hospitals, livestock, farms, waste treatment plants, etc.) (See chapter 5).

Cultivation-dependent methods were the first to be applied (Fang et al, 2007) while, subsequently, cultivation independent techniques and, more recently, next-

generation sequencing-based (NGS) techniques were used (Peccia and Hernandez, 2006; Franzetti et al, 2011). Owing to their high throughput and the reducing cost per sequence, the NGS techniques have a great potential in describing the diversity and composition of microbial communities (Bowers et al, 2010; Rinsoz et al, 2008). For example the Illumina technology can generate up to 600 Gbases in single run, thus providing high coverage both to amplicon-based and whole metagenomic studies of microbial communities and contributing to fill the gaps in the current knowledge of microbial community structure and dynamics (Claesson et al, 2010).

However, at present there is still a lack of information about community functions and to the best of our knowledge there are no studies on airborne bacteria that includes the chemical composition of PM in the community structure analysis.

In this work we focused the attention on Milan and Venice urban areas, these are two big and industrialized cities located in the Po valley (Italy). Indeed, the atmosphere of urban areas represents an interesting environment for microbiological surveys due to the considerable amount of air pollution and the high population density. In fact, it is well known that living in industrialized cities increases the health risk of inflammation and lung irritation because of the more intense exposition to particulate matter (PM), especially fine particles (Schwartz et al, 2002).

In particular, the Po valley that is a large plane in the middle of Northern Italy, results as one of the most polluted areas in Italy, as well as in Europe. This is due to the large number of industrial activities present in the area and to its peculiar geographical location and climatic conditions. The valley is enclosed between the Alps at North and the Appennines at South and this position allows the accumulation of pollution. Moreover, the occurrence of thermal inversion conditions, very frequent in the Po valley, represents one of the most significant pollution events (Mantecca et al, 2012). We collected 32 samples of PM₁₀ for each of the three sampling sites during an annual campaign and used the Illumina platform to sequence 16S rRNA gene libraries and to obtain the data necessary for a complete description of the airborne bacteria

communities. Furthermore, the quantitative PCR was used to estimate the total number of bacteria present in the selected environments.

The aim of this work was the assessment of the microbial community variability among seasons and sampling sites in order to gain insight into the processes that assemble the microbial community. In particular, dispersal and selection processes were assessed by the evaluation of (i) the potential relevance of airborne bacterial sources and (ii) the role of environmental and meteorological conditions on shaping the microbial community.

4.2 Material and Methods

4.2.1 Sample collection and DNA extraction

PM₁₀ was sampled in three different sites. In Milan (Lombardy, Italy) PM₁₀ was sampled in a typical urban station in the northern part of the city (45°30'35, 430''N, 9°12'38, 5239''E). A high-volume sampler (ECHO HiVol, TCR TECORA, Milan, Italy) was used with a flux speed of 200 l min⁻¹. The sampler was placed at about 20 m from the nearest roads. PM₁₀ sampling was performed at about 1 m from the ground on quartz fibre filters (Whatman, Maidstone, England).

In Venice 1 (Mestre, Veneto, Italy), PM₁₀ was sampled in an urban-traffic station located in the city centre near a crowded road (45°29'12, 426''N, 12°13'20, 327''E).

While in Venice 2 (Porto Marghera, Veneto, Italy), it was sampled in a sub-urban industrial area in the North-Western part of the city (45°26'18, 623''N, 12°12'13, 388''E). The high-volume samplers used for Venice 1 and 2, were provided by the Regional Agency of Environmental Protection (ARPA-Veneto, detailed information are available on the website www.arpa.veneto.it) and the sampling was performed on quartz fibre filters (Whatman, Maidstone, England) at flux speed of 500 l min⁻¹. The sampling campaigns of Venice 1 and 2 were executed by the Prof. G. Rampazzo research group of the University of Ca' Foscari, Venice.

Eight samples for each season and site were collected (Table 1). Each collection period lasted 24 hours. Unfortunately, the sample relative to Venice 2, summer, 6 July 2011, got lost and hence further analyses were impossible.

Table 1 – PM10 Sampling Campaign

| SEASON | PM10 SAMPLES | | | | | | | |
|--------|--------------|-----------|------------|------------|------------|------------|------------|------------|
| SUMMER | 6/7/2011 | 7/7/2011 | 13/7/2011 | 14/7/2011 | 20/7/2011 | 21/7/2011 | 27/7/2011 | 28/7/2011 |
| AUTUMN | 5/10/2011 | 6/10/2011 | 12/10/2011 | 13/10/2011 | 19/10/2011 | 20/10/2011 | 26/10/2011 | 27/10/2011 |
| WINTER | 11/1/2012 | 12/1/2012 | 18/1/2012 | 19/1/2012 | 25/1/2012 | 26/1/2012 | 1/2/2012 | 2/2/2012 |
| SPRING | 28/3/2012 | 29/3/2012 | 4/4/2012 | 5/4/2012 | 11/4/2012 | 12/4/2012 | 18/4/2012 | 19/4/2012 |

Total bacterial DNA was extracted directly from a quarter of each filter aseptically cut using the FastDNA Spin for Soil kit (MP Biomedicals, Solon, OH, USA). Individual filter portions were cut into small pieces and loaded into the bead tube of the DNA extraction kit, after the adding of 1M CaCO₃ in order to increase the pH, and shaken at 200 rpm for 60 min. The remaining steps of the DNA extraction were performed according to the manufacturer's instructions.

4.2.2 Quantitative PCR (qPCR)

The plasmid pCR2.1 (Life technologies Italia, Monza, Italy) containing a fragment of the bacterial 16S rDNA was used for standard concentration curves. The 466-bp fragment of the bacterial 16S rDNA (331-797 according to *E. coli* position) was obtained by PCR amplification with the universal primer-set 331 Forward (5'-TCCTACGGGAGGCAGCAGT -3') and 797 Reverse (5'-GGACTACCAGGTATCTAATCCTGTT-3') (Nadkarni et al, 2002).

The concentration of plasmidic DNA was measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies). Serial dilutions of the plasmidic DNA were used as the standard for qPCR.

The qPCR reaction was carried out in a total volume of 10 µL using the FluoCycleII Sybr reaction mix (Euroclone, Pero, Italy) with 0.3 µM forward and reverse primer

final concentration. The amplification was carried out with the Eco Real-Time PCR system (Illumina, CA, USA) under the following conditions: 95°C for 4 min, and 40 cycles of 95°C for 15 s, 60°C for 30 s and 72°C for 30 s, with acquisition of the fluorescence on the FAM canal at the end of each 72°C elongation step. The standards and the samples were included in triplicate in each run.

Due to their low abundance, it was impossible to quantify some DNA samples with the standard protocol. For this reason they were previously concentrated ten-fold and subsequently processed by qPCR. The samples processed in this manner were all the samples collected in Venice 1 and seven samples of Venice 2 (7, 20, 28 July 2011, 5, 19, 26 October 2011 and 18 January 2011). Despite this fact, several samples resulted even non quantifiable. Number of ribosomal operons was available for 68 samples.

4.2.3 16S rRNA fragment libraries by Illumina Hiseq 1000

For Illumina Hiseq sequencing the V5-V6 hypervariable regions of the 16S rRNA gene were amplified in 3 x 75 µL volume reactions with @Taq® Hot start Taq polymerase (Euroclone, Pero, Italy), 4 mM MgCl₂, 0.2 mM dNTPs mix, 1 µM of each primer. 783F and 1027R primers were used (Huber et al, 2007, Wang and Qian, 2009) and the cycling conditions were: initial denaturation at 94°C for 5 min; 29 cycles of 94°C for 50 s, 47°C for 30 s, and 72°C for 30 s and final extension at 72°C for 5 min. At the 5' end of 783F primer, one of 15 6-bp barcodes was also included to allow sample pooling and subsequent sequence sorting. The amplified products were purified with the Wizard SV PCR purification kit (Promega Corporation, Madison, WI, USA) and DNA quantity and purity were spectrophotometrically evaluated by NanoDrop™ (Thermo Scientific, USA).

15 purified amplicons with different barcodes were pooled in 100 µl samples with a DNA concentration of 40 ng/µl. Multiplexed sequencing of all the pooled samples were performed on a single Illumina Hiseq 1000 lane, using a paired-end 2x100 base-pair protocol and the 4.0 sequencing chemistry. The cluster extraction and base-calling processing analyses were performed by using the Illumina CASAVA Analysis

software, version 1.8. Illumina Hiseq 1000 sequencing was carried out at BMR Genomics, Padua, Italy.

4.2.4 Sequence analyses

Each sequence was assigned to its original sample according to its barcode. After sorting, the reverse read of each paired-end sequence was reverse-complemented and merged with the corresponding forward read, inserting 10 Ns in between (Claesson et al, 2010). A quality cut-off was then applied in order to remove sequences i) that did not contain the barcode, ii) with an average base quality value (Q) lower than 30. The barcode was removed from sequences before further processing.

The taxonomic attribution of filtered sequences was carried out using the stand-alone version of RDP Bayesian Classifier, using 50% of confidence as suggested for sequences shorter than 200 bp (Wang et al, 2007; Claesson et al, 2009).

4.2.5 Environmental factors and chemical data

4.2.5.1 Meteorological data and Particulate Matter (PM) concentration

Meteorological data (daily average temperature, daily rainfall, wind prevalent direction and speed, global solar radiation, relative humidity) and PM10 concentrations relative to the sampling periods were retrieved from Regional Agency of Environmental Protection (respectively ARPA-Lombardia for Milan site and ARPA-Veneto for Venice 1 and 2 sites) and are available at: www.arpalombardia.it and at: www.arpa.veneto.it.

4.2.5.2 Analysis of inorganic ions

The filter portion was dissolved in MilliQ® water in ultrasonic bath for 60 minutes and then analyzed by ion chromatography (IEC) to determine water soluble inorganic ions (Na^+ , NH_4^+ , Mg^{2+} , Ca^{2+} , K^+ , F^- , Cl^- , NO_3^- , NO_2^- , SO_4^{2-}) as described in Squizzato et al, 2012. These analyses were performed at the University of Ca' Foscari, Venice.

4.2.5.3 Analysis of elemental elements

The filter portion was disaggregated by the MDR microwave digestion system (Microwave Digestor Rotor Ethos 1600, Milestone) (Jarvis et al, 1992) and then analyzed by ICP-MS (Inductively Coupled Plasma Mass Spectrometry) and ICP-OES (Inductively Coupled Plasma Optical Emission Spectrometry) to determine the elemental element concentration. The acid mix used to mineralize the samples was composed by HNO₃, H₂O₂ and HF (Karthikeyan et al, 2006). The elements analyzed were Al, S, Fe, Ca, Mg, K, Ti, Cr, Mn, Zn, Cu, Ba, As, Ag, Cd, Ni, Pb, Sb, V, Co. These analyses were performed at the University of Ca' Foscari, Venice.

4.2.6 **Statistical analyses**

The complete dataset is composed by five variables accounting for meteorological conditions at sampling locations, one variable accounting for amount of PM in the atmosphere, two variables indicating the sampling location and season and 30 variables accounting for the chemical composition of PM. An Operational Taxonomic Unit (OTU) is defined on the basis of the classification at fourth taxonomic ranks of the RDP classifier.

OTUs found in one sample only (singletons) were removed because they may inflate variance explained by models (Legendre and Legendre 1998). A plot of OTU's mean relative abundance shows a discontinuity after *Sphingomonadales* (Fig. S1), which suggest that OTUs with mean relative abundance larger than 100 can be considered abundant OTUs. Therefore, abundant OTUs are 15 (including Chloroplasts) and excluding unclassified sequences.

4.2.6.1 Missing value imputation

Four variables among those describing chemical composition of PM were excluded from any subsequent analysis because they contained > 10% of missing values (variable = number of missing values: F = 24, NO₂⁻ = 24, Ag = 12, Na = 12). Missing

values in other variables were imputed. Data imputation is strongly recommended as it can avoid biased estimates from statistical analyses (Nakagawa and Frekleton, 2011) and is practically needed in multivariate analyses because a single missing datum should determine a large loss of information. We used the mice function (with default settings) in the MICE package (van Buuren and Groothuis-Oudshoorn, 2011) in R 3.0.2 (R Core Team 2013) to perform Markov Chain Monte Carlo (MCMC) multiple imputation by Gibbs sampling. MCMC methods for multiple imputation have the advantage that distributional assumption of data containing missing values (most commonly multivariate normality) are not necessary (Nakagawa and Frekleton, 2011).

4.2.6.2 Univariate analyses

Number of ribosomal operons was log-transformed to achieve normality (Shapiro-Wilk test of normality after log-transformation: $W = 0.987$, $P = 0.727$). Variation in the log-number of ribosomal OTUs according to season, city and their interaction was analyzed by generalized least squares ANOVA assuming a Gaussian error distribution and correcting for heteroscedasticity. Variation in the number of OTUS at each sample according to the same predictors was analyzed in a GLM assuming a Poisson error distribution.

4.2.6.3 Cluster analysis

A hierarchical cluster analysis of the relative abundance of OTUs was performed with the complete linkage method on the chi-square distance between samples. The chi-square distance was used here because Correspondence Analysis and Canonical Correspondence Analysis used in the following steps of the analyses (see below) approximates the chi-square distance among objects (Legendre and Legendre, 1998). In addition, this distance avoids the double-zero problem when comparing species composition between samples (Legendre and Legendre, 1998). Cluster analysis was performed with the HCLUST procedure in R 3.0.2 (R Core Team 2013) on both all OTUs and the most abundant ones.

4.2.6.4 Ordination techniques

Correspondence Analysis (CA) is an unconstrained ordination method used to explore the variability in bacterial community composition (Legendre and Legendre, 1998; Córdova-Kreylos et al, 2006). CA was performed on OTU abundances to visualize changes in airborne bacterial communities. This ordination technique was chosen because it is suitable to analyze species assemblages that are not monotonically related to one another (Ter Braak, 1985). Changes in meteorological data, including PM concentration in the atmosphere (“meteorological conditions” hereafter) and in chemical composition of the PM (“PM composition” hereafter) were visualized by means of Principal Components Analyses (PCA) on normalized variables. Meteorological data from Venice were identical for the two sampling sites, and were considered only once in these analyses.

We then aimed at relating bacterial community composition to environmental conditions and PM composition, and to assess the relative importance of these groups of variables in shaping bacterial community. “Environmental conditions” here refers meteorological conditions, plus a four-levels factor defining the season, a three-level factor defining the sampling site, and the two-way interactions among them. Factors were converted to dichotomous indicator variables whenever necessary. The procedure of analysis was as follows. First, we explored the structure of the correlation among both environmental variables and chemical variables and removed variables whose squared correlation coefficient was larger than 0.6 ($r^2 \geq 0.6$). This step avoided entering collinear variables in the following analyses. Second, we used Canonical Correspondence Analysis (CCA), a constrained ordination method similar to CA, to relate the bacterial community structure to PM composition and Environmental variables (Legendre and Legendre, 1998, Córdova-Kreylos et al, 2006). Models were simplified by backward removing non-significant predictors. When a categorical variable was significant, we conducted post-hoc pairwise comparisons among levels. In this case significance was adjusted according to the False Discovery

Rate (FDR) procedure (Benjamini and Yekutieli, 2001). Third, we used the variation partitioning (VarPart) in Canonical Analysis to assess the separate and joint contribution of environmental variables and PM composition in shaping bacterial communities. Fourth, we used Multivariate Regression Trees (MRT) to individualize thresholds in the values of environmental and PM variables that determined changes in the structure of bacterial communities. Fifth, we used the Indicator Species analysis (Dufrêne and Legendre, 1997) to identify species typical of different seasons, sampling locations, and groups identified by the MRT analysis (Borchard et al, 2011). CA, PCA and CCA analyses were performed with the BIODIVERSITYR package (Kindt and Coe, 2005) in R. VarPart analysis was performed with the VarCan software (Peres-Neto et al, 2005). MRT analysis was performed using the MVPART package (Therneau et al, 2007) in R. The best tree was chosen as the smallest tree whose cross-validated relative error (CVRE) was within 1 standard error of the CVRE of the best tree (Borchard et al, 2011). FDR procedure was performed with the MULTTEST package and Indicator Species analysis with the INDICSPECIES package (De Càceres and Legendre, 2009) in R.

4.3 Results

4.3.1 Quantification of airborne bacteria

The number of ribosomal operons per sample, which was used here as a measure of bacterial abundance (see chapter3 for details), ranged between 2.376×10^2 and 4.876×10^5 , with a mean value of $3.026 \times 10^4 \pm 8.675 \times 10^3$ SE. Log-transformed number of ribosomal operons varied significantly according to season ($F_{3,56} = 3.298$, $P = 0.027$) and sampling site ($F_{2,56} = 10.382$, $P < 0.001$), but not according to their interaction ($F_{6,56} = 1.872$, $P = 0.102$). Post hoc tests indicated that summer samples contained a lower number of ribosomal operons than both spring and winter samples (Figure 1A) and that Milan samples contained a larger number of ribosomal operons than Venice 2 samples (Figure 1B).

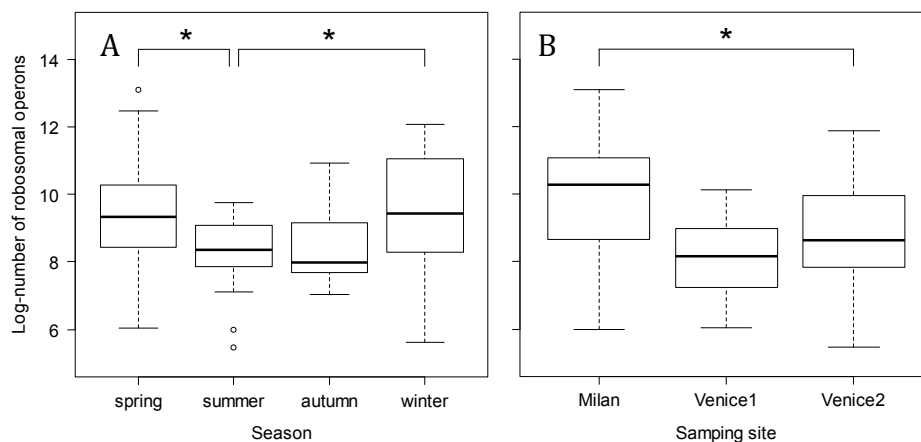


Figure 1 Number of ribosomal operons in A) different seasons and B) different sampling sites. The solid line represent the median value, the top and the bottom of the boxes represent the first and the third quartile while whiskers include 95% of data. Segments with asterisks denote seasons or sampling sites that differ significantly at Tuckey post-hoc test ($|z| \geq 2.414$, $P \leq 0.042$).

4.3.2 Microbial community structures

A total of 82,294,235 paired reads that passed the quality filter were obtained from sequencing. The number of paired reads of each sample ranged from 18,089 to 2,089,122 with a mean value of 587,755 for sample. In each sample, 22.2% to 42.3% of the sequences (mean value: 29,1%) could not be classified and were discarded from subsequent analyses. To avoid differences in coverage among the samples 10000 randomly-selected paired reads were chosen for each sample. This quantity still allows a detailed description of the community structures and a proper evaluation of alpha- and beta-diversity (Caporaso et al, 2012).

Overall 108 OTUs were identified in the PM of the three sampling sites, 102 of which were identified in more than one samples (non-singletons), while the number of OTUs identified at each sample ranged between 45 and 78, with a mean value of 57.0 ± 0.5 SE.

Number of OTUs per sample did not differ significantly according to season ($\chi^2_3 = 3.793$, $P = 0.285$), sampling site ($\chi^2_2 = 1.776$, $P = 0.411$), or their interaction ($\chi^2_6 = 3.251$, $P = 0.777$).

The taxonomic classification of bacteria based on the fourth taxonomic level showed an annual spatial variability not particularly marked (data not shown). The three analyzed sites were dominated by few orders, particularly *Burkholderiales*, *Actinomycetales* and *Chloroplast* that constituted, taken together, about 50% of the total communities. This pattern was particularly conserved in autumn and summer when the three communities are more similar with the exception of *Enterobacteriales* that were discretely abundant (about 8-10%) in Venice 2 during the summer. In winter, in all sites *Burkholderiales* were particularly abundant. In spring the presence of *Rhodobacterales* (about 30%) in Venice 2 was particularly relevant.

In Milan and Venice 1, the two urban sites, temporal variability followed a well-known trend: a progressive decrease of *Actinomycetales* and *Burkholderiales* with a parallel increase of Chloroplast associated with the temperature increase. Cold seasons were dominated by microorganisms generally retrieved in soil while in warm seasons there were mainly taxa of vegetable origin.

4.3.2.1 Cluster Analysis

Cluster analysis showed five main clusters both when applied to all OTUs and on the most abundant ones (Figure 2). Sample classification was generally consistent between the analyses, with 81 over 95 (i.e. 85.3%) of sample classified in the same cluster in both analyses. In addition, two clusters were consistently identified in these two analyses (Figure 1). One of these clusters, entirely composed by spring samples from Venice was characterized by a high abundance of *Rhodobacterales*, while the other, composed by spring and some summer samples from both Venice and Milan was characterized by a high abundance of Chloroplasts (Figures 2 and 3).

4.3.2.2 Canonical Analysis and Principal Components Analysis

The CA on OTU abundance suggested the existence of two ecological gradients. The first one goes from autumn and winter samples from both Venice and Milan on the one side, to summer and some spring samples from Milan on the other. The second gradient differentiated spring samples from Milan, from those from Venice (Figure 4A and B). This pattern of variation indicates that larger difference in bacterial community composition among cities could be observed in spring than in other seasons, thus suggesting that a season by city interaction should be included in the following analyses. Interestingly, the most extreme samples along this second gradient are those with high abundance of *Rhodobacterales* coherently classified in a separate cluster by previous analyses (see above and Figure 3).

Chemical composition of PM seems also to vary more in winter than in summer and more in Milan than in Venice, as indicated by the larger areas of the polygons including samples collected in different seasons and cities (Figure 4C). Variable scores indicate that the first axis of this PCA (PCA1) is a vector of decreasing concentration of almost all substances, and it is therefore a vector of decreasing abundance of PM₁₀, as confirmed by its strong negative correlation (Pearson $r = -0.755$) with amount of PM₁₀ in the atmosphere (note that amount of PM₁₀ was not among variables describing PM composition). The second axis (PC2) is instead a vector of increasing S, Cl and NH₄⁺ concentration and decreasing Al, Ca and Ca²⁺ concentration. Its positive relationship with S concentration probably explains why winter samples from Milan have the larger scores on this axis. Milan atmosphere is indeed highly enriched in S during winter due to by discharges of domestic heating systems (Brüggemann et al, 2008; Mantecca et al, 2012).

The first axis of the analysis of meteorological conditions is a vector of increasing temperature (score = 2.049) and wind speed (2.214) and decreasing humidity (-2.013) and PM₁₀ concentration (-2.824), while the second is a vector of increasing radiation and decreasing precipitation. It is therefore not surprising that winter and summer days occur at opposite sides along the first axis, and that spring, the rainiest season in

northern Italy shows large elongation along the second axis (Figure 4D). Meteorological conditions seem to be more variable in Milan than in Venice, as indicated by the very large polygon including sampling days in Milan, which also includes all the most extreme points.

4.3.2.3 Constrained ordination

Variables Fe, Cr, Cu, Ti, Ca, Ni, K, Zn, Co were removed from the dataset because they were strongly correlated ($r^2 \geq 0.652$) with other chemical variables. Square correlation among the remaining variables was ≤ 0.564 .

CCA analyses conducted on all OTUs or only on the most abundant ones gave consistent results. Indeed both analyses indicated that bacterial community structure varied according to the concentration of the same chemical constituents of PM, namely Ba, SO_4^{2-} and Mg^{2+} , and according to the same environmental variables, namely city, season, their interaction, wind speed, and humidity (Table 2).

Figure 2A Hierarchical cluster analysis on A) all OTUs. Numbers represent sample ID. Color bars at the bottom evidence the city (brown = Milan, violet = Venice 1, purple = Venice 2) or the season (green = spring, red = summer, orange = autumn, blue = winter) in which a given sample was collected. Symbols denote the clusters that are consistently identified in both analyses.

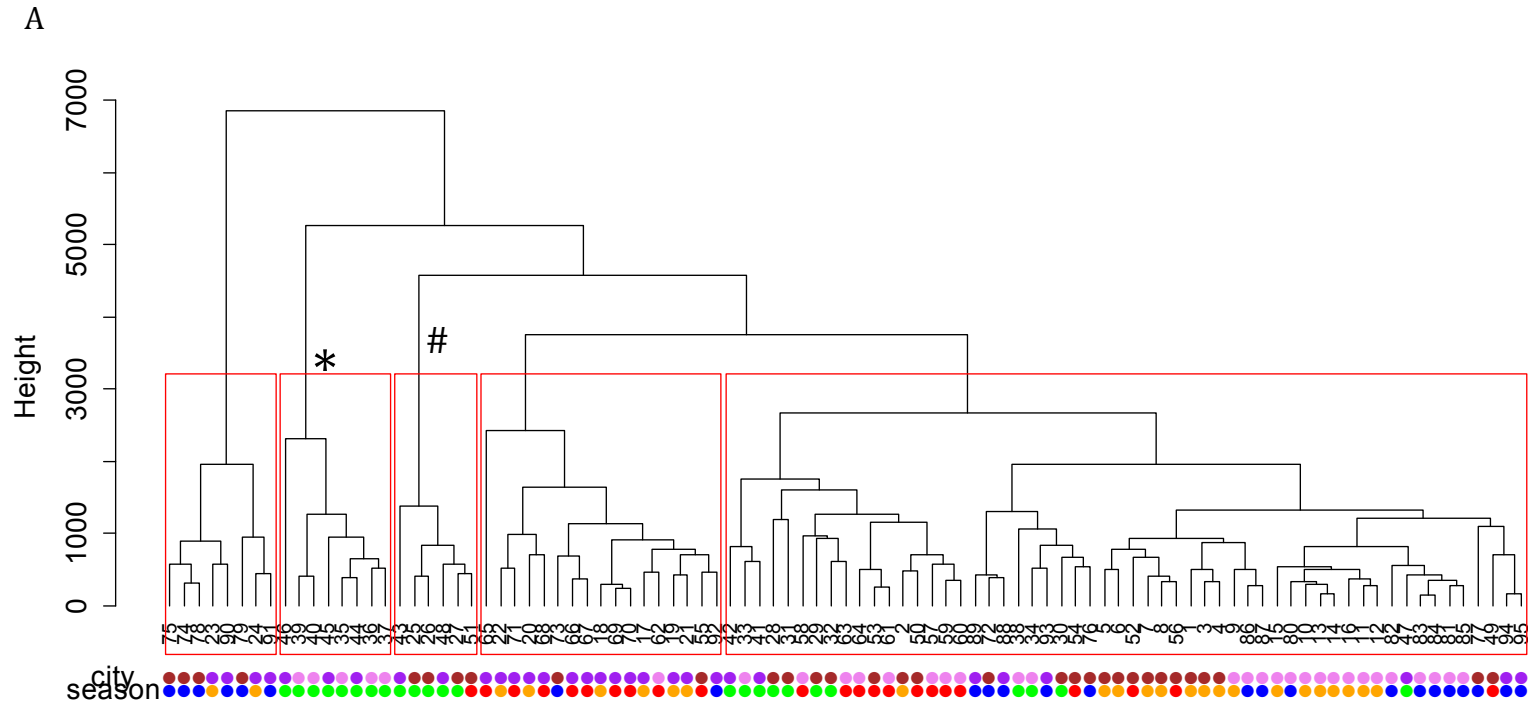
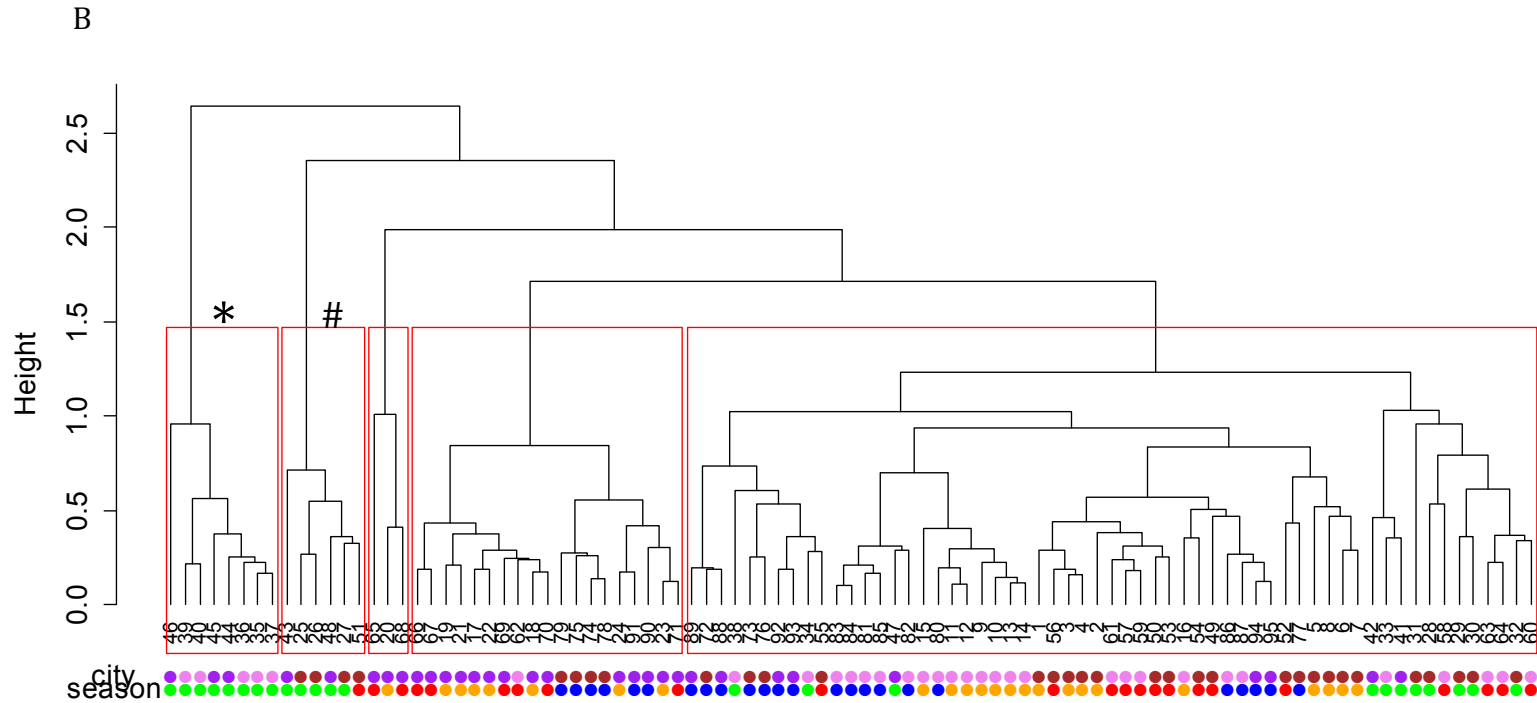


Figure 2B Hierarchical cluster analysis on B) abundant OTUs. Numbers represent sample ID. Color bars at the bottom evidence the city (brown = Milan, violet = Venice 1, purple = Venice 2) or the season (green = spring, red = summer, orange = autumn, blue = winter) in which a given sample was collected. Symbols denote the clusters that are consistently identified in both analyses.



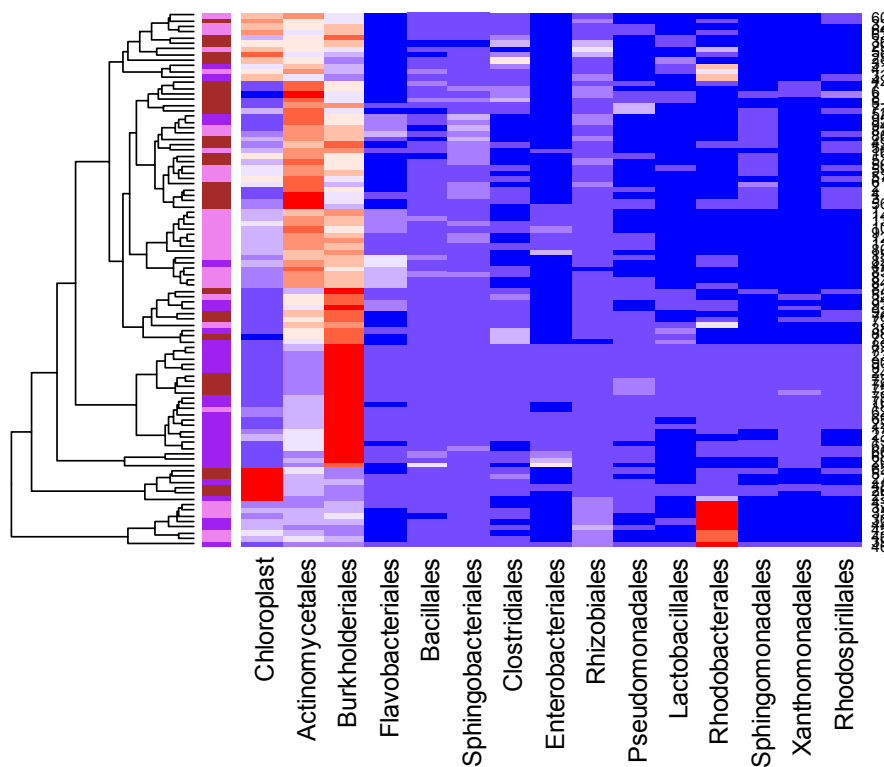


Figure 3 Heat-plot of the abundances (blue = low abundance, red = high abundance) of the most abundant OTUs detected in the 95 sample (numbers on the right). The dendrogram represents the hierarchical clustering analysis. Symbols denote the cluster consistently identified by both the analysis run on all the OTUs and on the most abundant ones. Colors at the base of the dendrogram represent sampling sites (brown = Milan, violet = Venice 1, purple = Venice 2)

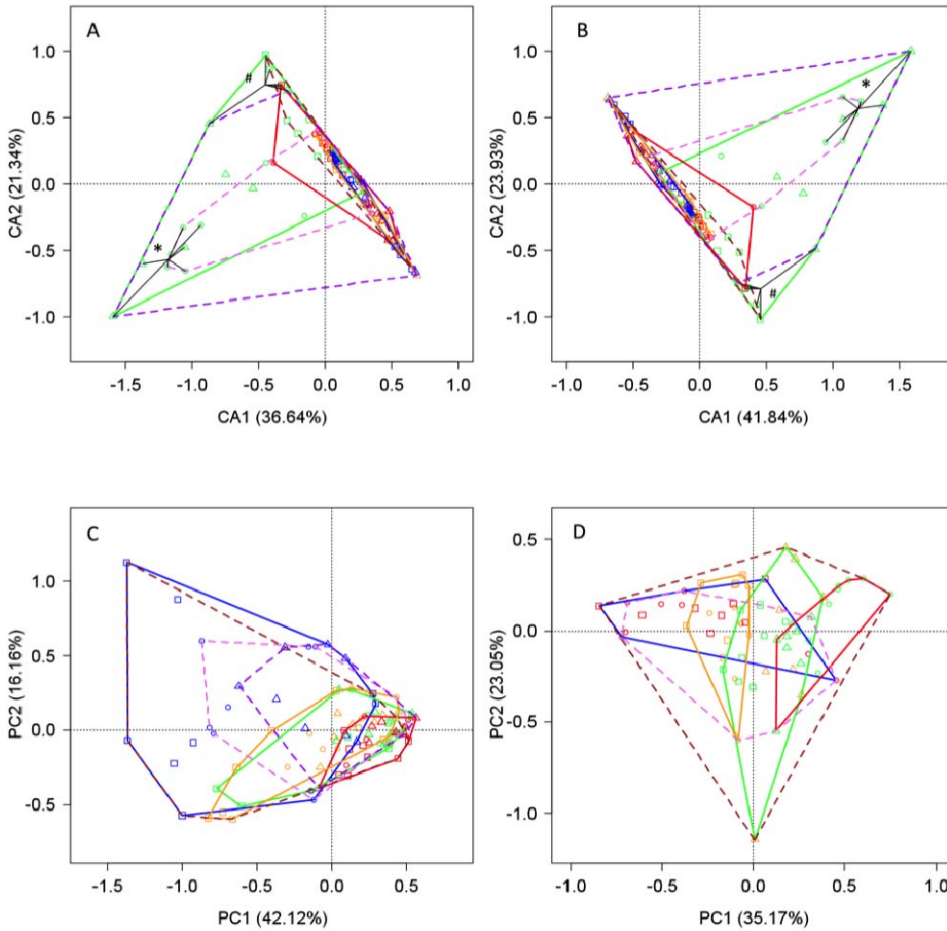


Figure 4 Canonical Analysis on the relative abundance of A) all the OTUs and B) the most abundant OTUs, and Principal Components Analysis on the standardized C) concentration of chemical substances in the PM and D) standardized meteorological variables, including PM abundance in the atmosphere. Symbols denote data collected at different cities (squares = Milan, circles = Venice 1, triangles = Venice 2), symbol colors denote the season (green = spring, red = summer, orange = autumn, blue = winter). Solid polygons include data collected at the same season (same colors as before) while dashed polygons include data collected at the same city (brown = Milan, violet = Venice 1, purple = Venice 2). The tiny black lines connect samples classified in the clusters indicated by the corresponding symbols in Figure 2A and 2B.

Table 2 CCA analyses of bacterial community structure based on All OTUs or only abundant OTUs in relation to variables accounting for PM composition and environmental conditions. The main effects of city and season were calculated by including in the model appropriate Helmert contrasts.

| | Effect | All OTUs | | | | | Abundant OTUs | | | | |
|-----------------------------|-------------------------------|----------|----------|--------|------------|-------|---------------------|----------|--------|------------|-------|
| | | Df | χ^2 | F | N. Per. | P | Df | χ^2 | F | N. Per. | P |
| PM composition | Ba | 1 | 0.016 | 4.958 | 199 | 0.005 | 1 | 0.015 | 5.887 | 199 | 0.005 |
| | SO ₄ ²⁻ | 1 | 0.010 | 3.220 | 199 | 0.015 | 1 | 0.009 | 3.656 | 199 | 0.010 |
| | Mg ²⁺ | 1 | 0.010 | 3.190 | 399 | 0.022 | 1 | 0.009 | 3.409 | 1299 | 0.035 |
| Environmental conditions | Season | 3 | 0.135 | 14.085 | 199 | 0.005 | 3 | 0.138 | 18.001 | 199 | 0.005 |
| | City | 2 | 0.046 | 7.209 | 199 | 0.005 | 2 | 0.047 | 9.185 | 99 | 0.005 |
| | Wind speed | 1 | 0.008 | 2.502 | 399 | 0.025 | 1 | 0.008 | 2.990 | 999 | 0.031 |
| | Humidity | 1 | 0.013 | 3.888 | 199 | 0.005 | 1 | 0.011 | 4.454 | 199 | 0.005 |
| | Season x City | 6 | 0.124 | 6.437 | 199 | 0.005 | 6 | 0.121 | 7.881 | 199 | 0.005 |
| C | <i>Overall test</i> | | | | | | <i>Overall test</i> | | | | |
| | Model | 16 | 0.391 | 7.634 | 199 | 0.005 | 16 | 0.391 | 9.585 | 199 | 0.005 |
| | Residual | 78 | 0.250 | | | | 78 | 0.200 | | | |

Post-hoc pairwise comparisons on both all and the most abundant OTUs indicated that spring samples significantly differed from those of all other seasons ($F_{1,39} \geq 12.112$, $P_{FDR} \leq 0.001$), and that also autumn samples differed from winter ones ($F_{1,39} \geq 3.387$, $P_{FDR} \leq 0.026$). Bacterial communities at all sampling sites were also generally different from one another, irrespective from season ($F_{1,53} \geq 6.052$, $P_{FDR} \leq 0.001$).

The significant interaction effect was investigated by pairwise comparisons among seasons within city, which indicated that seasonal variation in bacterial community structure differed among sampling locations. Indeed, in Venice 1, spring samples differed from those of all other seasons ($F_{1,12} \geq 7.215$, $P_{FDR} \leq 0.027$), in Venice 2 spring samples differed only from autumn ones ($F_{1,12} \geq 8.656$, $P_{FDR} \leq 0.014$), while in Milan spring samples differed from winter ones ($F_{1,12} \geq 4.528$, $P_{FDR} \leq 0.048$). These differences were significantly observed both in analyses run on all OTUs and on the abundant ones. In addition, analyses run on the abundant OTUs indicated that in Milan autumn samples also differed from both spring and winter ones ($F_{1,12} \geq 4.853$, $P_{FDR} \leq 0.043$) (Figure 5).

The significant interaction was also investigated by comparing sampling locations within seasons (Figure 5), which indicated that variation in bacterial communities among sampling locations differed in different seasons. Indeed in spring samples from Milan differed from those from Venice 1 ($F_{1,12} \geq 0.864$, $P_{FDR} \leq 0.031$), in summer no significant difference was observed among sampling locations ($F_{1,12} \leq 2.348$, $P_{FDR} \geq 0.416$), in autumn samples from Venice 2 were significantly different from both those of Milan and Venice 1 ($F_{1,12} \geq 4.176$, $P_{FDR} \leq 0.026$), and in winter bacterial community of Venice 1 differed from both those of Milan and Venice 2 ($F_{1,12} \geq 7.721$, $P_{FDR} \leq 0.028$). In addition, analyses run on the abundant OTUs indicated that in autumn samples from Venice 1 also differed from those from Milan ($F_{1,12} = 4.102$, $P_{FDR} = 0.028$), so that in this season bacterial communities from all sampling locations differed from one another. All other comparison were non-significant (details not shown).

Legend:

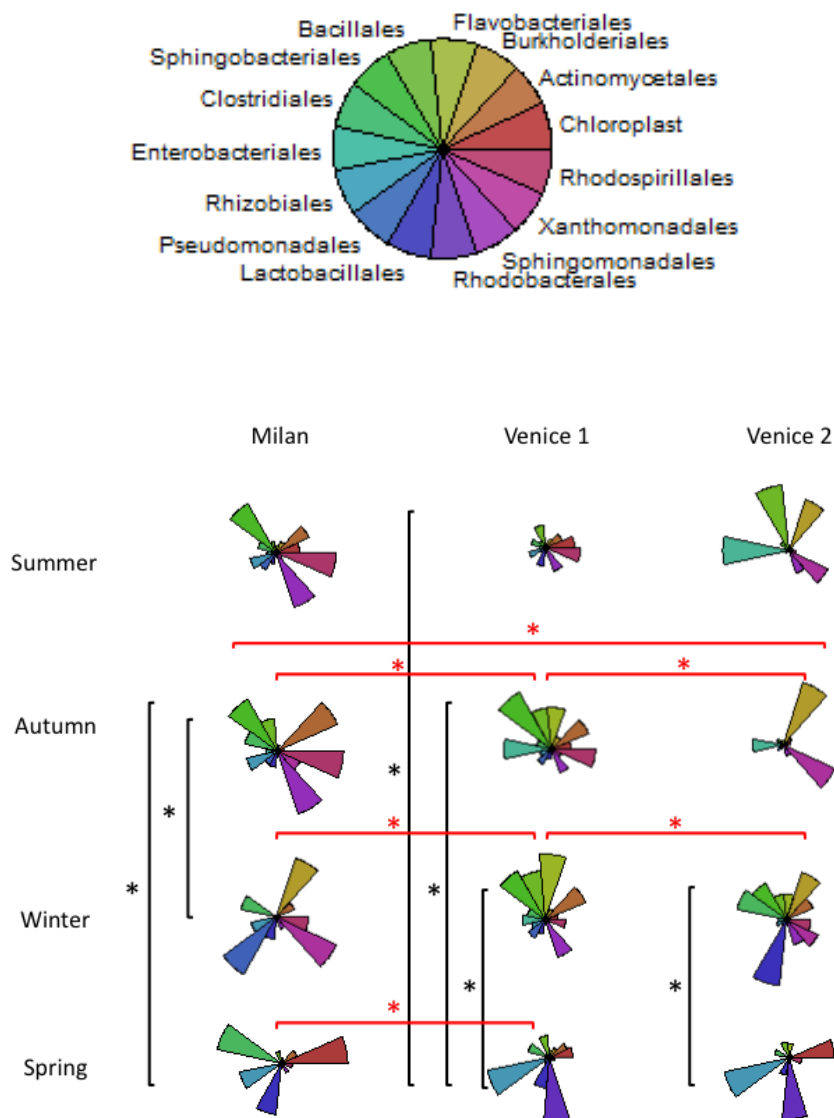


Figure 5 Significant interactions in bacterial communities by pairwise comparisons among sampling locations within seasons and viceversa. The legend shows the bacterial community structure. The red segments show the significant comparisons, while the black ones the significant comparisons among seasons within sampling locations.

4.3.2.4 Variation partitioning analysis

The variation partitioning analysis indicated that bacterial communities seem mostly related to environmental conditions, which *per se* explained a large part of the variation in bacterial community structure, equal to 47.3% in all OTUs and 53.0% on the abundant ones. Conversely, the PM composition *per se* only explained 4.2% of variance both on all and the most abundant OTUs. However, contribution of PM composition in shaping bacterial community structure, albeit minor, seems largely independent from that of environmental conditions, as the percentage of variation explained by their common contribution is only 1% in the analysis of all OTUs and 1.5% in that of the most abundant ones (Figure 6).

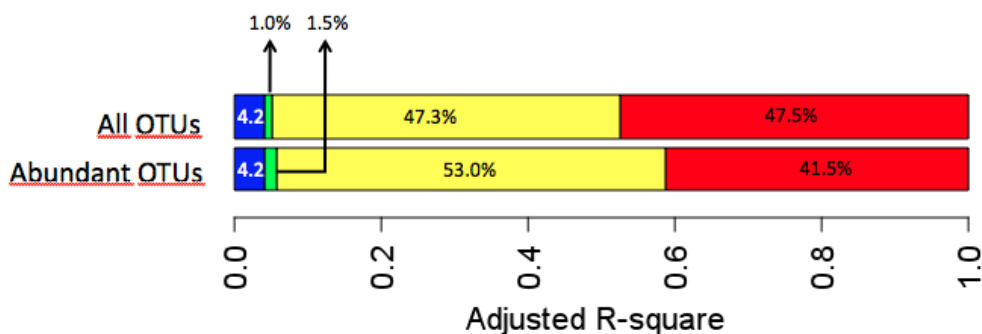


Figure 6 Variation partitioning analysis on all OTUs and most abundant OTUs. In Blue is shown the percentage of variance explained by PM composition *per se*. In yellow the percentage of variance explained by environmental conditions *per se*. In green the variance explained by both PM composition and environmental conditions. In red the variance not explained.

4.3.2.5 Multivariate Regression Trees (MRT)

MRT analysis indicated that bacterial communities can be divided in 7 groups according to PM composition and environmental variables selected in the previous analysis (Figure 7). Results of the analysis on all OTUs and the abundant ones were strictly consistent. The regression tree first separated spring samples from those of other seasons. Spring samples were then separated based on the sampling location, with samples from Milan on the one side (Group 1) and samples from Venice 1 and

Venice 2 on the other (Group 2). Samples from summer, autumn and winter were also separated according to sampling site, but with samples from Venice 2 from the one side and samples from Milan and Venice 1 on the other. This structure of the classification is consistent with the significant season by city interaction. Samples from Venice 2 in summer, autumn and winter were then further separated according to humidity being lower (Group 3) or higher than 89% (Group 4). Samples of the same seasons but from Milan or Venice 1 were instead separated according to Ba concentration being lower (Group 5) or higher than 1021 ng/m³, and this latter group was further separated according to SO₄²⁻ concentration being higher (Group 6) or lower than 3556 ng/m³ (Group 7).

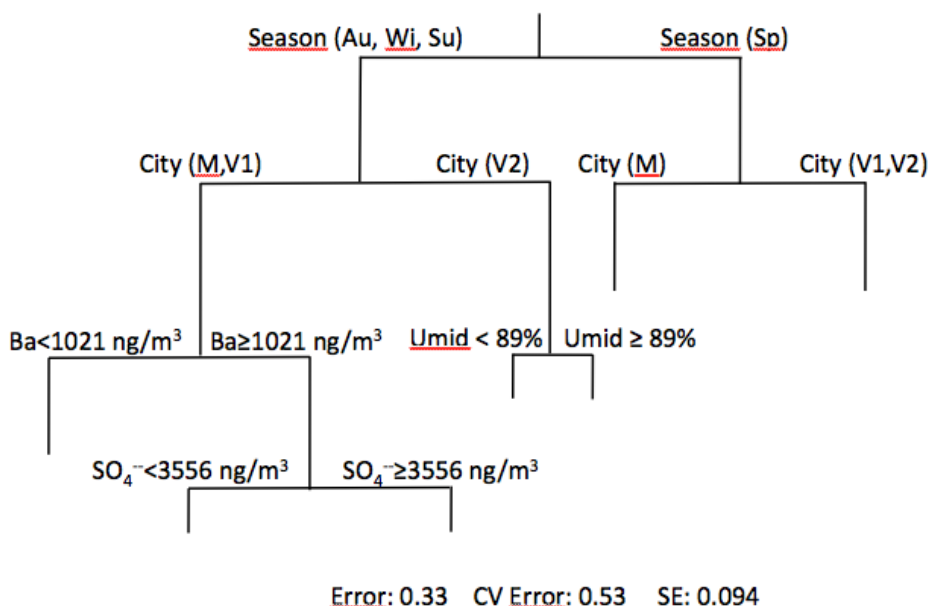


Figure 7 Multivariate Regression Tree (MRT).

4.3.2.6 Indicator taxa

The Indicator Taxa analysis suggested that no OTU were significantly associated with summer samples, *Spingobacteriales* and *Actinomycetales* with autumn samples, *Flavobacteriales*, *Pseudomonadales*, *Burkholderiales* and *Xanthomonadales* with

winter samples and *Rhodobacterales*, Chloroplasts and *Rhizobiales* with spring samples. The same analysis also indicated that *Pseudomonadales*, *Clostridiales*, *Rhodospirillales*, *Sphingomonadales* and *Actinomycetales* were typical of samples from Milan, *Flavobacteriales* of samples from Venice 1 and *Burkholderiales* and *Xanthomonadales* of samples from Venice 2.

When we used the Indicator Taxa analysis to individualize OTUs typical of the seven groups indicated by the MRT we found that Chloroplasts were typical of group 1, *Rhodobacterales* and *Rhizobiales* of group 2, *Burkholderiales* of group 3, *Sphingobacteriales* of group 5, *Pseudomonadales* and *Xanthomonadales* of group 6, *Actinomycetales* of group 7, while no OTU was typically associated with group 4 (Table 3).

Table 3 Results from separate Indicator Taxa analyses aiming at identifying OTUs typical of different seasons, cities, and groups identified by the MRT analyses.

| Group | Environmental condition | Taxon | IndVal | P |
|---------------|-------------------------|---------------------------|--------|-------|
| SEASON | | | | |
| 1 | Summer | NO OTU | | |
| 2 | Autumn | <i>Sphingobacteriales</i> | 0.538 | 0.050 |
| | | <i>Actinomycetales</i> | 0.534 | 0.015 |
| 3 | Winter | <i>Flavobacteriales</i> | 0.635 | 0.005 |
| | | <i>Pseudomonadales</i> | 0.609 | 0.005 |
| | | <i>Burkholderiales</i> | 0.592 | 0.005 |
| | | <i>Xanthomonadales</i> | 0.570 | 0.035 |
| 4 | Spring | <i>Rhodobacterales</i> | 0.861 | 0.005 |
| | | Chloroplast | 0.689 | 0.005 |
| | | <i>Rhizobiales</i> | 0.601 | 0.005 |
| CITY | | | | |
| 1 | Milan | <i>Pseudomonadales</i> | 0.673 | 0.005 |
| | | <i>Clostridiales</i> | 0.658 | 0.020 |
| | | <i>Rhodospirillales</i> | 0.637 | 0.005 |
| | | <i>Sphingomonadales</i> | 0.632 | 0.015 |
| | | <i>Actinomycetales</i> | 0.605 | 0.035 |
| 2 | Venice 1 | <i>Flavobacteriales</i> | 0.71 | 0.005 |
| 3 | Venice 2 | <i>Burkholderiales</i> | 0.686 | 0.005 |
| | | <i>Xanthomonadales</i> | 0.650 | 0.005 |

| MRT Group | | | | |
|-----------|--|---------------------------|-------|-------|
| 1 | Spring in Milan | Chloroplast | 0.628 | 0.005 |
| 2 | Spring in Venice 1 and Venice 2 | <i>Rhodobacterales</i> | 0.838 | 0.005 |
| | | <i>Rhizobiales</i> | 0.474 | 0.050 |
| 3 | Other seasons in Venice 2 with humidity < 89% | <i>Burkholderiales</i> | 0.549 | 0.005 |
| 4 | Other seasons in Venice 2 with humidity ≥ 89% | NO OTU | | |
| 5 | Other seasons in Milan or Venice 2 with Ba concentration < 1021 ng/m ³ | <i>Sphingobacteriales</i> | 0.439 | 0.045 |
| 6 | Other seasons in Milan or Venice 2 with Ba concentration ≥ 1021 ng/m ³ and SO ₄ ²⁻ concentration ≥ 3556 ng/m ³ | <i>Pseudomonadales</i> | 0.560 | 0.005 |
| | | <i>Xanthomonadales</i> | 0.552 | 0.005 |
| 7 | Other seasons in Milan or Venice 2 with Ba concentration ≥ 1021 ng/m ³ and SO ₄ ²⁻ concentration < 3556 ng/m ³ | <i>Actinomycetales</i> | 0.450 | 0.005 |

4.4 Discussion

4.4.1 Quantification of airborne bacteria

The evaluation of the total number of airborne bacteria present in the urban atmosphere of Milan and Venice, Northern Italy, was estimated by the measure of the ribosomal operon number per sample. It was determined by a culture-independent technique, specifically the qPCR. The observed values ranged between 2.4×10^2 and 4.9×10^5 ribosomal operons per m³. Summer samples contained a lower number of ribosomal operons than both spring and winter samples and Milan samples contained a larger number of ribosomal operons than Venice 2 samples (Figure 1). In general, these values are much larger than those reported in the only other study that used qPCR technique in Seoul (South Korea) (Lee et al. 2010), and in previous culture-based reports from urban areas (Bovallius et al. 1978; Fang et al. 2007). The lower abundance of these latter values may be ascribed to the underestimation of cultivation-dependent methods (Peccia and Hernandez 2006). In fact, higher values of bacterial concentration were obtained in studies that applied culture-independent

techniques (Bowers et al. 2009, 2011a, 2011b; Maron et al. 2005). Particularly, Bowers research group, measuring the total microbial abundance by epifluorescence microscopy, found values that ranged from about 10^5 to about 10^6 cells/ m^3 in all the three works with a relative spatial and temporal stability, while in Maron et al, 2005 the mean value found was 10^4 cells/ m^3 . Our results were located in the lower part of this range, however the mean bacterial abundance found so far, even in different parts of the world, seemed to be relatively homogeneous.

The results of this work were directly comparable with those obtained in our previous study in which the TSP fraction in Milan was assessed (see chapter 3), because of the use of the same technique and the study of the same geographical areas. The abundance retrieved in the present study (PM10) was slightly lower than in the TSP (Total Suspended Particulates) both referred to the air volume and the particulate weight. In particular, considering the number of ribosomal operons per g of particulate, in winter the mean obtained values had the same order of magnitude, respectively 2.9×10^8 n° rib op/g for TSP and 7.3×10^8 for PM10 in winter. The same occurred in spring, even with a higher order of magnitude, respectively 2.7×10^9 for TSP and 2.4×10^9 for PM10. Instead, summer and autumn showed a lower abundance in PM10 rather than in TSP. The mean values for summer were 3.7×10^9 (TSP) and 1.9×10^8 (PM10) in summer and 1.6×10^9 in autumn. While the mean values in autumn were 1.6×10^9 for TSP and 4.5×10^8 for PM10. The PM10/TSP ratio was estimated with the conversion factor proposed by Erbes, 1996.

4.4.2 Variability of microbial community structure and community-assembly processes

As reported in the introduction chapter the processes that drive microbial community assembly are dispersal, selection, diversification and drift. We considered the processes of diversification and drift negligible in the context of our study. Indeed, a single year time frame does not allow a significant genetic diversification while the ecological drift plays an important role in environments in which the selection is weak

and the alpha diversity is low (Nemergut et al. 2013). For these reason we focus the discussion of the results to gain insight into the contribution of dispersal and selection in assembling the microbial communities of the atmosphere.

4.4.2.1 Dispersal

Briefly, the dispersal process includes all the factors, which drive the “movement of microorganisms across space”, for instance, transport via air, water or mobile macroorganisms (Nemergut et al, 2013). Hence, the study of dispersal implied the research of those variables, which indicated the potential sources of airborne bacteria.

Previous works hypothesize soil, leaves, sea and animal feces as potential sources of airborne bacteria and recorded a temporal variability in the relative contribution of sources (Bowers et al, 2011; 2013; chapter 3). For instance, Bowers and colleagues (2011) investigating airborne communities of four US Mid Western cities, surprisingly discovered that, as well as the more predictable source (soil and leaves), the dog feces represented a source of bacteria, especially in winter. Moreover, in Bowers et al, 2013 was identified cow feces as the dominant source in early autumn and spring in the rural environment of Greeley (CO, USA), this suggested also a spatial variability of the sources.

In this work, thorough the indicator taxa approach; we found *Sphingobacteriales* (IndVal= 0.538, $p=0.05$) and *Actinomycetales* (IndVal= 0.534, $p=0.015$) as indicators of autumn samples. Bacteria belonging to these orders are typically found in soil and they could be considered as an indicator of the soil source. Similar results were also found in our previous study on Milan urban area (see chapter 3).

Another clear indication of the seasonal pattern of sources was the relevance of spring samples and the presence of Chloroplast (IndVal= 0.689, $p=0.05$) and *Rhizobiales* (IndVal= 0.601, $p=0.005$). Clearly, this result suggested the relevance of the vegetal source in shaping our microbial communities. In addition, the cluster analysis highlighted two relevant clusters among the five individuated. One of these

two clusters was composed by spring and some summer samples from both Venice and Milan and was characterized by a high abundance of Chloroplasts.

Moreover, the other cluster was entirely composed by spring samples from Venice and was characterized by a high abundance of *Rhodobacterales*. This is the first work on the atmosphere in which the *Rhodobacterales* abundance was particularly high, especially in this geographic area (Franzetti et al, 2011; chapter 3). The presence of this taxon in a lagoon environment, like the one surrounding Venice, is not surprising. Bacteria belonging to this order are purple non-sulfur bacteria and include species, which possess an extensive range of metabolisms. For instance *Rhodobacter* is a photosynthetic organism, it can grow in both the light and the dark and it can be in freshwater or marine environments (Yunyun et al, 2013). Moreover, their abundance in spring could be ascribed to the temperature increase and to the possible algae growth increment, conditions that could promote the growth of this particular bacterial taxon (Gilbert et al, 2012). Therefore this taxon could be considered typical of Venice 2 sampling location in spring. This result is in agreement with previous findings reporting Bowers and coworkers (2011a) that the composition of the airborne bacterial communities was significantly related to land use type (agricultural, suburban or forest) and that these differences were imputable to shifts in bacterial sources, rather than local meteorological conditions (Bowers et al., 2011a). Moreover they also demonstrated that the composition of the microbial community in the atmosphere is influenced not only by changes in the relative contribution of the different sources in different seasons as extensively reported (Bowers et al, 2011a; 2011b; 2012; 2013), but also by the seasonal dynamics of the microbial community in the sources themselves.

Finally, the significant contribution of the wind speed, revealed by CCA, confirmed importance of the dispersal process in shaping the microbial communities in the atmosphere. Indeed, particles from different sizes and locations could be transported to the sampling points at different wind speeds. A back trajectories analysis will be applied to elucidate the origin of the samples. This is a valid tool to evaluate and

compare the origin of air masses that transport bioaerosols, particularly on areas located in the middle of two very different environments like Venice (i.e. marine vs. continental) (Maki et al, 2013; Polymenakou, 2012; Woo et al, 2013).

4.4.2.2 Selection

Different habitat harbor different suites of microorganisms and a variety of environmental factors have a critical role in determining bacterial community structure and diversity. The complexity of the potential environmental drivers of fitness made extremely difficult to isolate a single selection agent. In fact, only few studies were able to identify single meteorological factors that affect the structure of microbial communities in the atmosphere. In particular, Maron et al. (2006) individuated temperature and relative humidity as the most contributors to the differences microbial communities of a French urban area. While in Bowers et al, 2012 the only selector variable found was temperature.

The variation partitioning analysis indicated that bacterial communities seem mostly related to environmental conditions, which *per se* explained a large part of the variation (47.3% for all OTUs and 53.0% for the abundant ones) in bacterial community structure. Most of variables included in the category “environmental conditions” are categorical variables (i.e. season and city) and meteorological factors that exhibit strong seasonal correlation (i.e. temperature and humidity), thus describing both different sources and different environments. Hence, the variance explained by the environmental conditions accounts for processes related to both dispersion and selection and these two effects were difficult to disentangle. Nevertheless, variables with a selective effect included in the ‘environmental conditions’ category could exist, however with our data we were not able to identify them.

Conversely, the category “PM composition” included ions (Na^+ , NH_4^+ , Mg^{2+} , Ca^{2+} , K^+ , F^- , Cl^- , NO_3^- , NO_2^- , SO_4^{2-}) and elemental elements (Al, S, Fe, Ca, Mg, K, Ti, Cr, Mn, Zn, Cu, Ba, As, Ag, Cd, Ni, Pb, Sb, V, Co). Although some of these elements and compounds

showed different concentrations in different seasons (i.e. nitrates, ammonium and Cl^- (even with a lower mean abundance) presented a peak in winter while they were less abundant in summer), most of them did not display seasonality or strong differences in Milan and Venice (i.e. Mg^{2+} and Na^+ remained constant and, Ca^{2+} and SO_4^{2-} slightly varied). For these reasons we considered the category “PM composition” suitable to account for selection processes.

Variation partitioning analysis showed that the PM composition *per se* only explained 4.2% of variance both on all and the most abundant OTUs. However, contribution of PM composition in shaping bacterial community structure, albeit minor, seems largely independent from that of environmental conditions, as the percentage of variation explained by their common contribution is only 1% in the analysis of all OTUs and 1.5% in that of the most abundant ones (Figure 6). As a matter of fact, CCA disclosed a significant effect of three PM components, namely Ba, SO_4^{2-} and Mg^{2+} , on microbial community structure. The concentration of SO_4^{2-} and Mg^{2+} , essential for bacterial growth, may affect the survival and growth of the microbial populations differentially, thus affecting their relative abundance (Scherer and Sahm, 1981; Savai et al, 2000). MRT analysis consistently grouped the samples of Milan and Venice 1 collected in summer, autumn and winter according to Ba concentration being lower (Group 5) or higher than 1021 ng/m³, and this latter group was further separated according to SO_4^{2-} concentration being higher (Group 6) or lower than 3556 ng/m³ (Group 7). On the one hand, the finding that *Actynomicetales* were retrieved as indicator taxon at lower concentration of sulphate is consistent with the ability of this taxon of producing spore to survive in nutrient-limiting conditions. On the other hand the correlation between high concentrations of Ba (>1021 ng/m³) and the *Pseudomonadales* was quite surprising since *Pseudomonas* genus is known to be particularly sensitive to BaCl_2 (Sivolodskii, 2012). Other analysed elements and compounds could potentially have a role in bacterial metabolism. Some of the analyzed ions (Na^+ , Mg^{2+} , Ca^{2+} , K^+ and Cl^-) are involved in the basic cellular activities such as membrane ion channels exchanges (Tisa et al, 2000). NH_4^+ , Mg^{2+} , NO_3^- , NO_2^- ,

SO₄²⁻ can act as electron donor, electron acceptors or reduced compounds from anoxic respirations. Other elements may have toxic properties on bacteria, such as Ti, Cr, Zn, Cu, As, Ag, Cd, Ni, Pb, Sb, V and Co. However, our data did not disclose any significant effect of these elements and compounds on the microbial community structure.

4.5 Conclusions

In conclusion, our data confirm that dispersion is an important community-assembly process as also previously described in literature (see references above). In particular, bacterial sources significantly contributed to change the community structure over time. The changes over time are due both to the different relative contribution of the sources in the different season and to the intrinsic changes in the structure of sources themselves across seasons. Moreover, we found important clues that indicate possible selection processes like the significant effect of Ba, SO₄²⁻ and Mg²⁺ that allow the selection of some microbial populations rather than others. These clues support the hypothesis that the presence of some nutrients in the aerosol lets to presume that the atmosphere could harbor the sufficient conditions to hypothesize a living bacterial community (Polymenakou, 2012).

4.6 References

Benjamini Y, Yekutieli (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29:1165–1188.

Borchard D, Gillet F, Legendre F (2011) *Numerical Ecology with R*. Springer: New York, doi: 10.1007/978-1-4419-7976-6.

Bovallius A, Bucht B, Roffey R, Anas P (1978) Three-year investigation of the natural airborne bacterial flora at four localities in Sweden. *Appl Environ Microb* 35, 847–852.

Bowers RM, Lauber CL, Wiedinmyer C, Hamady M, Haller AG, Fall R, Knight R, Fierer N (2009) Characterization of airborne microbial communities at a high-elevation site and their potential to act as atmospheric ice nuclei. *Appl Environ Microb* 38: 6029–6041.

Bowers RM, McLetchie S, Knight R, Fierer N (2011a) Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *ISME J* 5:601–612.

Bowers RM, Sullivan AP, Costello EK, Collet JL, Knight R, Fierer N (2011b) Source of bacteria in outdoor air across cities in the midwestern United States. *Appl Environ Microbiol* 77: 6350–6356.

Bowers RM, McCubbin IB, Hallar AG, Fierer N (2012) Seasonal variability in airborne bacterial communities at a high-elevation site. *Atmos Environ* 50, 41-49.

Bowers RM, Clements N, Emerson JB, Wiedinmyer C, Hannigan MP, Fierer N (2013) Seasonal variability in bacteria and fungal diversity of near-surface atmosphere. *Environ Sci Technol* DOI: 10.1021/es402970s.

Brodie EL, De Santis TZ, Moberg Parker JP, Zubietta IX, Piceno YM, Andersen GL (2007) Urban aerosols harbor diverse and dynamic bacterial populations. *P Natl Acad Sci USA* 104,299-304.

Brüggemann E, Gerwig H, Gnauk TH, Müller K, Herrmann H. (2009) Influence of seasons, air mass origin and day of the week on size-segregated chemical composition of aerosol particles at a kerbside. *Atmos. Environ.* 43: 2456-2463.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-

throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621-1624.

Claesson MJ, O'Sullivan O, Wang Q, Nikkila J, Marchesi JR, Smidt H, de Vos WM, Ross RP, O'Toole PW (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4, e6669.

Claesson MJ, Wang QO, O'Sullivan O, Greene-Diniz R, Cole R JR, Ross P, O'Toole PW (2010) Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Res* 38, e200.

Córdova-Kreylos AL, Cao Y, Green PG, Hwang HM, Kuivila KM, Lamontagne MG, Van De Werfhorst LC, Holden PA, Scow KM (2006) Diversity, composition and geographical distribution of microbial communities in California salt marsh sediments. *Appl. Environ. Microbiol.* 72: 3357-3366.

De Cáceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119:1674–1684.

Deguillaume L, Leriche M, Amato P, Ariya PA, Delort AM, Pöschl U, Chaumerliac N, Bauer H, Flossmann AI, Morris CE. (2008). "Microbiology and atmospheric processes: Chemical interactions of primary biological aerosols." *Biogeosciences* 5:1073–1084.

DeLeon-Rodriguez N, Latham TL, Rodriguez-R LM, Barazesh JM, Anderson BE, Beyersdorf AJ, Ziemba LD, Bergin M, Nenes A, Konstantinidis T (2013) Microbiome of the upper troposphere: species composition and prevalence, effects of tropical storms, and atmospheric implications. *P Natl Acad Sci USA* online first.

Dufrêne M, Legendre P (1997) Species assemblages and indicator species: The need for a flexible asymmetrical approach. *Ecological Monographs* 67, 345-366.

Erbes RE (1996) A practical guide to air quality compliance. 2nd ED. Chapter 4 p. 41 'Ambient air quality standards and atmospheric dispersion'. John Wiley and Sons, Inc. NY USA.

Fahlgren C, Hagström A, Nilsson D, Zweifel UL. (2010) "Annual variations in the diversity, viability, and origin of airborne bacteria." *Appl. Environ. Microbiol.* 76, 3015–3025

Fang ZG, Ouyang ZY, Zheng H, Wang XK, Hu LF (2007) Culturable airborne bacteria in outdoor environments in Beijing, China. *Microb Ecol* 54:487–496.

Fierer N, Liu Z, Rodriguez-Hernandez M, Knight R, Henn M, Hernandez MT (2008) Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microb* 74:200–207.

Franzetti A, Gandolfi I, Gaspari E, Ambrosini R, Bestetti G (2010) Seasonal variability of bacteria in fine and coarse urban air particulate matter. *Appl Microbiol Biotechnol* 90:745–753.

Gilbert JA, Steele JA, Caporaso JG, Steinbruck L, Reeder J, Temperton B, Huse S, McHardy AC, Knight R, Joint I, Somerfield P, Fuhrman JA, Field D. (2012) Defining seasonal marine microbial community dynamics. *ISME J* 6:298-308.

Griffin, DW. (2004) "Terrestrial microorganisms at an altitude of 20,000 m in Earth's atmosphere." *Aerobiologia* 20:135–140

Huber J, Mark Welch D, Morrison H, Huse S, Neal P, Butterfield D, Sogin M (2007) Microbial population structures in the deep marine biosphere. *Science*: 97-100.

Jarvis KE, Gray AL, Houk RS (1992) Handbook of inductively coupled plasma mass spectrometry. Blackie Academic and professional.

Jie Y, Houjin H, Feng J, Jie X. (2011) The role of airborne microbes in school and its impact on asthma, allergy, and respiratory symptoms among school children Rev. Med. Microbiol. 22: 84–89.

Karthikeyan S, Joshi UM, Balasubramanian R (2006) Microwave assisted sample preparation for determining water-soluble fraction of trace elements in urban airborne particulate matter: evaluation of bioavailability. Analytica Chimica Acta. 576: 23-30.

Kindt R, Coe R (2005) Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. World Agroforestry Centre (ICRAF): Nairobi, Kenya.

Lee SH, Lee HJ, Kim SJ, Lee HM, Kang H, Kim YP (2010) Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. Sci Total Environ 408, 1349–1357.

Legendre P, Legendre L (1998) Numerical ecology (2nd ed). Elsevier: Amsterdam, The Netherlands.

Lighthart B (1997) The ecology of bacteria in the alfresco atmosphere. FEMS Microbiol Ecol 23, 263-274.

Maki T, Kakigawa M, Kobayashi F, Yamada M, Matsuki A, Hasegawa H, Iwasaka Y. (2013) Assessment of composition and origin of airborne bacteria in the free-troposphere over Japan. Atmos. Environ. 74: 73-82.

Mantecca P, Gualtieri M, Longhin E, Bestetti G, Palestini P, Bolzacchini E, Camatini M. (2012) Adverse biological effects of Milan urban PM looking for suitable molecular markers of exposure. *Chem Ind Chem Eng Q* 18: 635-641.

Maron PA, Lejon DPH, Carvalho E, Bizet K, Lemanceau P, Ranjard L, Mougél C (2005) Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. *Atmos Environ* 39, 3687–3695.

Maron PA, Mougél C, Lejon DPH, Carvalho E, Bizet K, Marck G, Cubito N, Lemanceau P, Ranjard L (2006) Temporal variability of airborne bacterial community structure in an urban area. *Atmos Environ* 40, 8074–8080.

Nadkarni MA, Martin FE, Jacques NA, Hunter N (2002) Determination of bacterial load by real-time PCR using a broad-range (universal) probe and primers set. *Microbiology* 148:257–266.

Nakagawa S, Frekleton RP (2011) Model averaging, missing data and multiple imputation: a case study for behavioural ecology. *Behav Ecol Sociobiol* 65: 103-116.

Nemergut DR; Schmidt SK; Fukami T, O'Neill SP, Bilinska TM, Stanish LF, Knelman JE, Darcy JL, Lynch RC, Wickey P, Ferrenberg S (2013) Patterns and Processes of Microbial Community Assembly. *Microbiol Mol Biol Rev* 77: 342-356.

O'Malley MA (2008) 'Everything is everywhere: but the environment selects': ubiquitous distribution and ecological determinism in microbial biogeography. *Stud Hist Phil Biol & Biomed Sci* 39: 314-325.

Peccia J, Hernandez M (2006) Incorporating polymerase chain reaction-based identification, population characterization and quantification of microorganisms into aerosol science: a review. *Atmos Environ* 40:3941–3961.

Peres-Neto PR, Jackson DA, Somers KM (2005) How many principal components? stopping rules for determining the number of non-trivial axes revisited. *Comput Stat Data An* 49: 974-997.

Polymenakou PN (2012) Atmosphere: a source of pathogenic or beneficial microbes? *Atmosphere* 3, 87-102.

Rinsoz T, Duquenne P, Greff-Mirguet G, Oppliger A (2008) Application of real-time PCR for total airborne bacterial assessment: comparison with epifluorescence microscopy and culture-dependent methods. *Atmos Environ* 42: 6767–6774.

Savai J, Kojima H, hashimoto A, Shoji S, Sawaki T, Hacoda A, Kawada E, Kokugan T, Shimizu M (2000) Antibacterial characteristics of magnesium oxide powder. *World J Microb biot* 16: 187-194

Schwartz J, Laden F, Zanobetti A (2002) The concentration-response relation between PM_{2.5} and daily deaths. *Environ Health Perspect.* 110: 1025-1029.

Scherer P, Sahm H (1981) Influence of sulphur-containing compounds on the growth of *Methanosarcina bakeri* in a defined medium. *Eur J Appl Microbiol.* 12:28-35.

Sivolodskii EP (2012) Determination of the Sensitivity of Bacteria to Barium Ions, a Taxonomic Marker of the Genus *Pseudomonas*. *Microbiol.* 81: 112-117.

Squizzato S, Masiol M, Innocente E, Pecorari E, Rampazzo G, Pavoni B (2012) A procedure to assess local and long-range transport contributions to PM_{2.5} and secondary inorganic aerosol. *Aerosol Science* 46:64–76.

ter Braak CJF (1985) Correspondence-Analysis Of Incidence And Abundance Data - Properties In Terms Of A Unimodal Response Model. *Biometrics* 41, 859-873.

Therneau TM, Atkinson B, Ripley B, Oksanen J, De'ath G (2007) mvpart: Multivariate partitioning. R package version 1.2-6. <http://cran.r-project.org/web/packages/mvpart/index.html>.

Tisa LS, Sekelsky JJ, Adler J (2000) Effects of Organic Antagonists of Ca²⁺, Na⁺, and K⁺ on Chemotaxis and Motility of *Escherichia coli*. J. Bacteriol. 182: 4856-4861.

van Buuren S and Groothuis-Oudshoorn K (2011) mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45.

Wang Y & Qian P (2009). Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies. PLoS ONE 4:e7401.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microb 73: 5261–5267.

Wainwright M, Wickramasinghe NC, Narlicar JV, Rajaratnam P (2003) Microorganisms cultured from strato spheric air samples obtained at 41 KM. FEMS Microbiol Lett. 218: 161-165.

Woo AC, Brar MS, Chan Y, Lau MCY, Leung FCC, Scott JA, Vrijmoed LLP, Zafar-Reza P, Pointing SB (2013) Temporal variation in airborne microbial populations and microbially-derived allergens in a tropical urban landscape. Atmos. Environ. 74: 291-300.

Yunyun F, Keats KF, Rivkin RB, Lang AL (2013) Water mass and depth determine the distribution and diversity of Rhodobacterales in an Arctic marine system. FEMS Microbiol Ecol 84: 564-576.

CHAPTER 5

**Unraveling the bacterial diversity in the
atmosphere**

5.1 Introduction

The term 'bioaerosol' refers to airborne biological particles such as bacterial cells, fungal spores, viruses, pollens and their by-products such as endotoxins (Stetzenbach et al. 2004). Most of the first studies on aerosols focused on the non-biological fraction of particles, but, in recent years, interest has grown on the biological fraction, and in particular in disclosing the composition and the effects of bioaerosol. This interest has been boosted by an increasing amount of evidence suggesting that this fraction may play a critical role in the negative effects of aerosols on biological systems. For example, it has been proved that airborne endotoxins are partly responsible for inflammation processes caused by aerosol inhalation (Becker et al. 2005; Camatini et al. 2012; Finnerty et al. 2007; Longhin et al. 2013), while some biological particles, such as pollens and fungal spores, can cause asthma and allergies (D'Amato 2002; Moorman et al. 2011).

The atmosphere can be considered an extreme environment for microorganisms because of its chemical and physical characteristics, such as high solar radiation, low moisture and nutrients, and its large dispersing capability. Nevertheless, first investigations demonstrated that microorganisms do exist in the atmosphere and can be metabolically active (Cote et al. 2008; Fuzzi et al. 1997; Lighthart 1997; Lighthart and Shaffer 1995a,b; Sattler et al. 2001; Tong and Lighthart 1998; Womack et al. 2010). For this reason, it has been speculated that bacteria might be actively involved in the detrimental effects of bioaerosols. Despite that it is still unclear whether airborne bacteria are an actual ecological community or are simply a pool of organisms passively gathered together from different sources, many authors refer to the bacteria in the atmosphere as to the 'microbial community'. In this paper the same conventional term will be adopted, without implying that ecological processes occur among bacteria.

The main interests in air microbiology currently lay on the detection of potentially harmful microorganisms in indoor environments and several reviews have been published on this topic (D'Arcy et al. 2012; Eames et al. 2009; Eduard et al. 2012; Tang

2009), while studies on bioaerosols in outdoor environments have lagged behind. In addition, research on this latter topic mainly focused on the atmosphere as a source of pathogenic microbes (Polymenakou 2012). Particularly, many studies assessed microorganism dispersion from bio-hazardous agricultural or industrial plants such as wastewater treatment plants, composting plants and breeding farms. Reviews on this topic have been recently published (Dungan 2010; Dungan and Leytem 2009; Korzeniewska 2011; O'Hara and Rubin 2005; Millner 2009). However, a small but significant body of literature exists on the diversity and variability of outdoor, near-surface airborne microbial assemblages, whose composition is not influenced by the presence of specific point sources of potentially hazardous microbes.

This paper aims at reviewing the current literature on this particular topic of air microbiology, and on the analytical methods used in this research field. We also emphasize how the use of culture-independent techniques and, more recently, of the Next-Generation Sequencing-based methods (NGS), has improved the ability of the scientific community to address the main research questions in this field, which are: 1) the description of the temporal and spatial variability of airborne microbial communities, 2) the identification of the environmental factors affecting this variability and the diversity of airborne microbes, 3) the study of the potential terrestrial and marine sources of atmospheric bacteria. Critical evaluations of the state of the knowledge as well as suggestions about research needs are also provided.

5.2 Sampling methods

A wide variety of bioaerosol sampling methods is currently available and no standardized protocols have been established so far. Different devices have been developed for particulate matter (PM) sampling; they all have both, advantages and disadvantages, and the choice is mainly determined by the aims of the study (Fahlgren et al. 2011) and the analyses that will be performed on the samples. An accurate sampling campaign design is also crucial for data analysis, and differences in sampling strategies can hamper data comparison. Sampling devices can be divided in

four principal categories on the basis of the method used for bioaerosol collection: gravity, impaction, liquid impingement and filtration.

In the first method, also called 'depositional sampling', airborne microorganisms are collected by gravity deposition on an agar medium exposed to the environment. This method is cheap and simple to perform, but it is non-quantitative and biased towards the collection of large particles, which settle more than small ones (Buttner et al. 2002). Since the volume of air involved in bioaerosol deposition is also unknown, depositional sampling does not allow determining the concentration of organisms. However, it is used for cultivation-based studies, because it preserves the viability and cultivability of sampled microorganisms.

The other three sampling methods use air forced samplers, and allow determining the concentration of organisms (Stetzenbach et al. 2004). Impactors drive the deposition of particles onto a solid collection surface, generally agar medium, exploiting the particle inertia (Fang et al. 2007). In addition, they can deposit the particles onto an adhesive-coated surface that can be analyzed microscopically. The airflow rates of impactors usually range from 10 to 700 l min⁻¹. This sampling method is suitable for cultivation-based studies, but the impaction stress generated on the collected bioaerosols may reduce the retrieval of culturable microorganisms.

The impingement consists in the collection of airborne particles into a liquid. The typical airflow rate of these samplers ranges from 12.5 to 20 l min⁻¹. Recently, new high-velocity devices based on impingement, referred to as 'aerosol-to-hydrosol samplers', have also been developed (Xu et al. 2011). The main advantage of this technique is the possibility to process samples by dilution or concentration and to use samples for different types of analysis such as cultivation, microscopy, flow cytometry or molecular methods (Cho and Hwang 2011; Fierer et al. 2008). The main disadvantage is a limit in collection time to a few hours with current devices, due to the fact that an extended sampling may increase the sampling stress and reduce the viability of the collected bacteria (Fahlgren et al. 2011).

Devices based on filtration force the passage of air through a porous filter membrane where bioaerosol is collected. The highly variable features of these devices allow the collection of samples with different characteristics, which depend, for example, on the chosen bioaerosol fraction, on filter pore size and material, and on airflow rate. The last usually varies from 1 to hundreds of l min⁻¹. This technique is simple, low-cost, versatile and suitable for many different downstream analyses of the samples. However, filtration may determine cell desiccation and compromise the viability of vegetative cells (Griffin et al. 2011). For this reason it is generally used in studies that do not imply cultivation-dependent analyses (Bowers et al. 2012).

5.3 Description of airborne microorganisms

5.3.1 Quantification of airborne microorganisms

The assessment of the bacterial load in the air is a central task for the description of airborne microbial communities. Traditionally, quantification has been carried out by culture-dependent methods such as direct plate count. The precision of this method clearly depends on the success of cultivation. In fact, a relevant fraction of microorganisms are supposed to be viable, but they do not form colonies on agar plates because they entered the viable-but-not-culturable state (Lighthart 1997; Peccia and Hernandez 2006). Studies whereby both culture-dependent and culture-independent microbial counts were performed on different atmospheric samples allowed estimating that the percentage of culturable microorganisms varied from less than 1% to a maximum of 20% in the atmosphere (Cho and Hwang 2011; Peccia and Hernandez 2006; Temkiv et al. 2012; Vaitilingom et al. 2012). For this reason, the use of culture-independent techniques has significantly increased in recent years. Among these, total direct count has become usual (Rinsoz et al. 2008). It involves DNA staining with fluorescent dyes, such as DAPI (4',6-diamidino-2-phenylindole) or SYBR Green, and a subsequent epifluorescence microscopy observation. In addition, if specific dyes such as acridine orange are used, viable cells can also be recognized and quantified (Maron et al. 2005). However, a considerable amount of bacteria must be

present in samples (over 10^4 cells l^{-1}) to obtain reliable results. Moreover, it is also possible to misidentify microorganisms with debris of non-biological origin. Flow cytometry can also be used for cell enumeration (Bowers et al. 2011b). This method offers great speed, automation and more accuracy in sample processing, but it requests a minimum bioaerosol concentration of 1000 cells ml^{-1} . This restriction can be overcome by increasing the sampling time (Georgakopoulos et al. 2009).

The development of PCR-based methods enhanced the ability to detect and identify microorganisms (An et al. 2006). The main advantages of these techniques are better sensitivity and higher speed than other methods. In addition, PCR-based methods do not depend on viable or culturable cells for enumeration, so that a wide variety of samplers can be used for bacteria collection. In particular, the quantitative PCR (qPCR) is a method that measures in real-time the increasing amount of amplification products in PCR reactions using a fluorogenic compound (SYBR Green, TaqMan Probes, etc.). Either total bacteria or a peculiar group of microorganisms can be quantified depending on the primer set used. For instance, Brodie and colleagues (2007) used the qPCR to validate PhyloChip detection of changes in abundance of *Pseudomonas oleovorans*. However, qPCR may underestimate the actual amount of airborne bacteria, mainly because of primer mismatches or inefficiency in DNA extraction from samples (DeLeon-Rodriguez et al. 2013). Moreover, a recent study pointed out that qPCR is able to discriminate between microbial abundances only when they differ more than 1.3 to 3.2 times (Hospodsky et al. 2010).

This heterogeneity in the quantification techniques used in different studies hampers comparison of results, as reported abundances can differ by several orders of magnitude depending on the enumeration method used. For example, cultivation-dependent counts revealed a bacterial abundance that ranged from 0 to $\approx 10^4$ CFU m^{-3} (Fahlgren et al. 2010; Fang et al. 2007; Haas et al. 2013; Jeon et al. 2011; Ravva et al. 2012), while studies that estimated total bacterial load by epifluorescence returned values that ranged from $\approx 10^3$ to $\approx 10^7$ cells m^{-3} (Bowers et al. 2009, 2011a, 2012; Harrison et al. 2005; Vařtilingom et al. 2012), and abundance estimates from qPCR

ranged between 10^1 and 10^6 cells m^{-3} (chapter 3; Cho and Hwang 2011; Lee et al. 2010; Li et al. 2010; Smith et al. 2012). The lower number of CFU counts reported in cultivation-based studies is probably due not only to the aforementioned cultivation bias, but also to a loss of cell viability consequent to their exposure to atmospheric stresses (desiccation, temperature UV radiation, etc.) (Zweifel et al. 2012), so that currently there is no definitive method for assessing the total number of viable bacteria in the atmosphere.

5.3.2 Cultivation-based surveys

Cultivation was the method used in the first aerobiological studies that aimed at investigating airborne microbial communities in outdoor environments. These studies provided a first description of airborne microorganisms by a taxonomic classification of a variable number of isolates from different environments, such as urban, rural, forest, coastal and remote sites (see e.g. Di Giorgio et al. 1996; Fang et al. 2007; Mahdy and El-Sehrawi 1997; Mancinelli and Schulls 1978; Shaffer and Lighthart 1997). A number of surveys was also conducted on dust events in different geographical regions. They were extensively reviewed elsewhere (Griffin 2007) and will not be treated here. Although studies on airborne microorganisms were conducted at sites that largely differ in general ecological and environmental conditions, 70- 90% of isolates were generally Gram-positive bacteria, particularly belonging to the genera *Bacillus*, *Micrococcus* and *Staphylococcus*. The only notable exception was an urban site in Marseilles (France), where 60% of bacteria were instead Gram-negative (Di Giorgio et al.1996). The authors of these studies hypothesized that Gram-positive bacteria were more abundant because their spores are more resistant to the harsh conditions of desiccation and high exposure to solar radiation, which characterize the atmospheric environment. However, cultivation methods heavily underestimate both the number and the diversity of microorganisms in a given environment. In addition, some studies reported large differences in the composition of airborne microbial communities assessed through culture-based or molecular methods (Cho and Hwang

2011; Ravva et al. 2012), while others observed substantially overlapping results (Fahlgren et al. 2010; Urbano et al. 2011).

Despite all the biases and limits described so far, cultivation is the method of choice for the isolation of microbial strains, which is still the most direct method to obtain reliable information on physiologic and metabolic characteristics of bacteria. For example, cultivation-based studies revealed that a high percentage of colonies grown on plates are pigmented (Cho and Hwang 2011; Fahlgren et al. 2010; Fang et al. 2007; Lighthart 1997; Mahdy and El-Sehrawi 1997; Mancinelli and Schulls 1978; Shaffer and Lighthart 1997; Temkiv et al. 2012; Väitilingom et al. 2012). Pigmentation is a very common feature in airborne microorganisms, probably because it protects cells from atmospheric UV exposure (Tong and Lighthart 1997) and contributes to survival at low temperatures (Fong et al. 2001). Kuske (2006) also reminded that culture-based analyses are still fundamental for some studies, such as epidemiological and sanitary surveys, because the successful isolation of microorganisms allows further investigation of their pathogenic traits. For example, Gandolfi et al. (2011, chapter 2) determined the antibiotic resistance of several microbial strains, and in particular of *Staphylococcus* strains, after isolation from PM10. Cultivation-based methods are also preferred when the focus of the study is on viable microorganisms only. For example, Temkiv et al. (2012) isolated *Methylobacterium* and *Bradyrhizobium* strains from hailstones, and successively tested their ability to utilize a range of different organic compounds as carbon sources. They demonstrated that *Methylobacterium* isolates were commonly able to metabolize some carboxylic acids that constitute PM. Moreover, Al-Bader et al. (2012) isolated from air dust consortia of phototrophs and heterotrophs able to use crude oil as carbon source. These are among the first evidences of metabolic activity of airborne bacteria. Recently, some culture-based surveys also dealt with the identification of potential sources of airborne bacteria (see also paragraph 3.3.4), especially in marine and coastal environments (Cho and Hwang 2011; Fahlgren et al. 2010; Urbano et al. 2011), often with the aid of the analysis of backward trajectories of air masses. These studies demonstrated that, when air

masses or aerosols are from marine origin, Gram-negative isolates generally predominate, while Gram-positive bacteria predominate in continental air masses. Studies based on culturable bacteria collected in air masses of the troposphere that moved over a long-range suggested that viable microorganisms retrieved in North America may be originated from regions as far as China and Japan (Smith et al. 2012).

5.3.3 Molecular characterization of microbial communities

The increasing awareness that cultivation biases pose an obstacle to a full and comprehensive description of microbial communities prompted microbiologists to turn to PCR-based methods also in the field of aerobiology. The first molecular-based works on airborne microorganisms were extensively reviewed by Peccia and Hernandez (2006). Since then, several other surveys have been conducted with the aim of giving a thorough description of airborne microbial communities (Table 1). Furthermore, the recent development of Next- Generation Sequencing-based methods (NGS) greatly improved the ability of microbiologists to describe microbial communities in different environments (Caporaso et al. 2012). In the atmosphere, the high-throughput performances of NGS allowed obtaining taxonomic profiles of multiple airborne communities by sequencing thousands of 16S rRNA gene amplicons per sample. Thus, even rare bacterial populations could be detected and the robustness of statistical methods applied to the results increased.

Even within the studies that used molecular investigation methods, there are many heterogeneities in sampling methods, sampled PM fraction, season when samples were collected, and target environment examined (e.g. urban, rural, marine, etc.), as well as in molecular and statistical methods used in the analyses. These heterogeneities in experimental design often hamper the comparison of results from different studies. However, most authors focused their research on a few specific topics, so that, probably, the current information on airborne microbial communities can be better summarized according to them. They are: 1) the evaluation of biodiversity of airborne microbial communities; 2) the assessment of the temporal

and spatial variability of the structure of airborne communities; 3) the influence of environmental and meteorological parameters on airborne communities; 4) the identification of potential sources of airborne bacteria.

Table 1 Main papers since 2005 dealing with descriptions of airborne microbial communities through molecular-based methods; TSP = Total Suspended Particulate.

| Reference | PM fraction | Target environment | PCR-based methods |
|-------------------------|-------------------|--------------------------------|--------------------|
| Harrison et al. 2005 | PM10 | Urban, rural, remote | Cloning |
| Maron et al. 2005 | TSP | Rural | Cloning, ARISA |
| Kuske 2006 | TSP | Urban | T-RFLP |
| Maron et al. 2006 | TSP | Urban | ARISA |
| Brodie et al. 2007 | TSP | Urban | Cloning, PhyloChip |
| Després et al. 2007 | PM2.5 | Urban, rural, high-alpine | Cloning, T-RFLP |
| Fierer et al. 2008 | TSP | University campus | Cloning |
| Polymenakou et al. 2008 | Several fractions | Coastal during dust storm | Cloning, RFLP |
| Bowers et al. 2009 | TSP | High-elevation site | Cloning, NGS |
| Fahlgren et al. 2010 | TSP | Coastal city | Cloning |
| Lee et al. 2010 | TSP | Urban | T-RFLP |
| Li et al. 2010 | TSP | Urban | DGGE |
| Bowers et al. 2011a | TSP | Agricultural, suburban, forest | NGS |
| Bowers et al. 2011b | PM2.5 | Urban, small town | NGS |
| Franzetti et al. 2011 | PM10, PM2.5 | Urban | NGS |
| Jeon et al. 2011 | TSP | Suburban during dust storm | Cloning, DGGE |
| Bowers et al. 2012 | TSP | High-elevation site | NGS |
| Zweifel et al. 2012 | TSP | Atmospheric boundary layer | Cloning |
| Bertolini et al. 2013 | TSP | Urban | NGS |

5.3.3.1 Evaluation of biodiversity of airborne microbial communities

One of the primary aims of several researches using molecular methods of analysis is to obtain a more detailed and more comprehensive description of microbial taxa associated with PM than that achieved in previous cultivation-based surveys. Sanger sequencing-based investigations revealed a different scenario than that depicted by cultivation-based methods, since in several cases they showed a predominance of Gram-negative microorganisms, mainly belonging to *Alpha*, *Beta*- and *Gammaproteobacteria*, and *Bacteroidetes* (Després et al. 2007; Fahlgren et al. 2010, 2011; Fierer et al. 2008; Lee et al. 2010; Maron et al. 2005; Zweifel et al. 2012). At lower taxonomic levels, the most frequently observed genera were *Sphingomonas* for *Alphaproteobacteria* and *Pseudomonas* for *Gammaproteobacteria* (Fahlgren et al. 2010; Harrison et al. 2005; Li et al. 2010; Zweifel et al. 2012). Only in few occasions were Gram-positive taxa more represented than Gram-negative; in these cases, the most abundant genus was *Bacillus* (Brodie et al. 2007; Harrison et al. 2005; Kuske 2006). Similar results were obtained with the use of NGS techniques. In fact, in the first paper reporting an NGS-based survey for the analysis of airborne bacteria, Bowers et al. (2009) described the bacterial community collected near-surface at a high-elevation site in Colorado (USA) using a barcoded 454 pyrosequencing protocol. They found that all air samples collected in summer had similar bacterial communities regardless of atmospheric conditions and were dominated by *Betaproteobacteria* (mainly *Burkholderiales*) and *Gammaproteobacteria* (mainly *Moraxellaceae*). Also in a more recent study conducted at the same site, *Proteobacteria* dominated microbial communities of summer samples, with *Moraxellaceae* and *Comamonadaceae* being the dominant groups within *Betaproteobacteria* (Bowers et al. 2012). Other recent studies using NGS techniques confirmed that *Proteobacteria* (particularly *Burkholderiales*) are among the dominant phyla, but pointed out that also

Actinobacteria and *Firmicutes* were among the most abundant bacterial groups (chapter 3; Bowers et al. 2011a,b; Franzetti et al. 2011). Overall, the majority of culturable airborne bacteria are Gram-positive, rather than negative (paragraph 3.2), while the same holds true in a few cases when molecular methods are used. Therefore, these results clearly indicate that cultivation-based methods are unable to provide a thorough description of the structure of airborne bacterial communities.

The composition of airborne microbial community can dramatically change during extreme meteorological events. For example, Polymenakou et al. (2008) analysed the PM collected during a dust event in Crete (Greece) and found a sharp prevalence of *Firmicutes* in the coarser fractions, which were gradually replaced by *Actinobacteria* in the finer fractions. Jeon et al. (2011) also compared airborne microbial communities between dust and non-dust event days and observed a clear difference in the structure of microbial communities. *Firmicutes* dominated the communities during dust events, accounting for 53% of sequences, while *Proteobacteria* dominated in non-dust event days (53% of sequences). These data may indicate a prevalence of spore-forming bacteria in desert dust, maybe due to their higher ability of surviving to long-range transport (Marchant et al. 2008). DeLeon-Rodriguez et al. (2013) observed that microbial communities in the upper troposphere (8-15 km altitude) dramatically changed after a hurricane, with a majority of the alphaproteobacterial genus *Afipia* being replaced by members of the *Burkholderiales*.

Some authors also analyzed the archaeal component of airborne microbial communities, either by universal (Bowers et al. 2009, 2011a,b, 2012; Fierer et al. 2008) or by domain-specific primers (Cho and Hwang 2011; Després et al. 2007) for PCR amplification, while Brodie et al. (2007) included 307 archaeal taxa in their PhyloChip. Sequences classified in the Archaea domain were never retrieved, except for one sample of marine aerosol (Cho and Hwang 2011), in which a phylotype belonging to *Euryarchaeota* was detected.

Chloroplast sequences are commonly retrieved when analyzing airborne bacterial communities and, notwithstanding the use of bacterial-specific primer pairs, they can

represent up to 79% of sequences, because of partially overlapping sequences with bacterial 16S rRNA gene. The issue of whether or not to take them into account in the subsequent analyses is still debated, and has been addressed in both ways. For example, Brodie et al. (2007), Zweifel et al. (2012), Fahlgren et al. (2011), Franzetti et al. (2011) and our previous work reported in chapter 3 included chloroplast sequences in their analyses, while Fahlgren et al. (2010), Jeon et al. (2011) and Maron et al. (2005) excluded them.

Biodiversity of airborne microbial communities was estimated in several papers through indices, such as Shannon diversity or evenness (Fahlgren et al. 2011; Jeon et al. 2011; Maron et al. 2005; Ravva et al. 2012), and/or rarefaction curves (Bowers et al. 2009, 2012; Brodie et al. 2007; Ravva et al. 2012). When the latter method was applied to the large number of sequences provided by pyrosequencing, Franzetti et al. (2011) observed that OTU richness in the near-surface atmosphere of an urban area of Italy was greater than any other value previously reported for air bacterial communities and close to those reported for soil. The same analysis revealed that, on the contrary, the upper troposphere harbours less complex communities compared with several other environments, probably due to the extreme conditions at high altitudes (DeLeon-Rodriguez et al. 2013). The same analyses can be performed also on community fingerprints rather than on proper sequences. For example, Polymenakou et al. (2008) constructed rarefaction curves from an RFLP-based distribution of clones in OTUs. The appropriateness of both approaches has been widely debated. In fact, Lee et al. (2010) calculated a set of diversity indices on T-RFLP profiles and found lower values than Maron et al. (2005), which conversely based their analyses on a clone library. Such discrepancy can be due to an underestimation of taxon diversity of T-RFLP technique. On the contrary, Després et al. (2007) found that T-RFLP analysis was able to reveal a larger bacterial diversity than the sequence analysis. Nevertheless, they observed a lower biodiversity than that generally retrieved in other environments, such as soil.

5.3.3.2 Temporal and spatial variability in the structure of airborne microbial communities

Airborne bacterial communities seem to be in permanent change. Indeed, most studies using Sanger sequencing-based methods found large changes in the structure of airborne communities between consecutive days or even between hours within the same day (Fierer et al. 2008). In addition, seasonal variability was suggested to be even higher (Brodie et al. 2007; Lee et al. 2010; Maron et al. 2005, 2006). Molecular fingerprint techniques, such as DGGE, T-RFLP and ARISA, proved to be particularly useful for the investigation of the temporal and spatial variability of airborne microbial communities. For example, Jeon et al. (2011) found very similar DGGE profiles, both for consecutive days and for different locations of the Seoul (South Korea) suburban area in the same day, which abruptly changed only when a dust event occurred. Interestingly, Li et al. (2010) demonstrated that microorganisms retrieved at ground level are not significantly different from those found at higher elevation (238 m), thus suggesting the occurrence of a wide atmospheric mixing.

The temporal and spatial variability of bacterial communities has also been investigated in different studies based on NGS methods, and most of these studies showed significant differences among the airborne microbial communities in different seasons (chapter 3; Bowers et al. 2011b, 2012; Franzetti et al. 2011). The taxa most responsible for this seasonal shift were *Firmicutes* and *Actinobacteria*, whose abundance was higher in cold seasons, and *Sphingomonadales*, whose abundance was higher in summer (Bowers et al. 2009; Franzetti et al. 2011). Interestingly, the observed seasonal variability in the composition of bacterial communities was lower in Total Suspended Particle (TSP) than in finer fractions (PM₁₀ and PM_{2.5}) in Milan (Italy), thus suggesting that airborne bacterial communities associated with different dimensional fractions of PM may differ (chapter 3; Franzetti et al. 2011). The airborne community showed also significant changes according to location and land-use around sampling point, as were the cases of agricultural, forest and suburban

environments in northern Colorado (USA) (Bowers et al. 2011a), and of three cities and one small town in Midwestern United States (Bowers et al. 2011b).

The statistical methods used in these investigations were essentially multivariate. Principal Coordinate Analysis (PCoA) is commonly used to represent multivariate microbial communities in a few dimensions, and is the method of choice when non-Euclidean distances (as the commonly used UniFrac distance) are involved. Analysis of similarity (ANOSIM) and multivariate analysis of variance (MANOVA) were commonly used to investigate differences in the structure of microbial communities among seasons and environments. PRIMER (Clarke and Gorley 2006) and R (R Core Team 2012) are the most commonly used softwares for these analyses.

5.3.3.3 Influence of environmental and meteorological factors on airborne bacterial communities

The identification of the environmental factors and meteorological conditions that affect the structure of the airborne microbial communities and the abundance of microorganisms in the atmosphere is another major topic of air microbiology. However, only few studies using both NGS-based and fingerprinting/cloning methods were able to identify single meteorological factors that affect the structure of microbial communities (chapter 3; Bowers et al. 2012; Brodie et al. 2007; Maron et al. 2006). The statistical method most commonly used to investigate this topic is the Mantel test albeit Multivariate Regression Trees (MRT) have also been employed (chapter 3; Brodie et al. 2007).

Air temperature was most often the factor that significantly shaped the microbial community in different environments. For example, Maron et al. (2006) showed that temperature and relative humidity mostly contributed to the differences in bacterial community structure among weeks in the urban area of Nancy (France). In the near-surface atmosphere of Colorado the temperature was the only predictor of bacterial community structure along the four seasons and it significantly explained the intra-seasonal variability in autumn and spring bacterial communities (Bowers et al. 2012).

MRT disclosed cut-off values of mean daily temperature that separated the bacterial communities in different clusters. Particularly, air samples collected in days with temperatures below 22°C in Texas (Austin and San Antonio) or below 6°C Italy (Milan) were characterized by higher abundances of *Actinomycetales* (chapter 3; Brodie et al. 2007). However, most authors recognized that both the spatial and the temporal variability of airborne bacterial communities are more likely due to the combined effects of differences in several meteorological factors, in the chemical composition of PM and in the relative importance of the main sources of the air particles, rather than to temperature per se (chapter 3; Brodie et al. 2007).

5.3.3.4 Identification of potential sources of airborne bacteria

Potential sources of airborne bacteria have commonly been hypothesized on the basis of the most abundant taxa retrieved in PM samples (chapter 3; Bowers et al. 2011a,b, 2012; Brodie et al. 2007; Franzetti et al. 2011; Maron et al. 2005). In some papers, the Similarity Percentage (SIMPER) function of PRIMER (Clarke and Gorley 2006) was used to identify taxa that can serve as ‘indicators’ of different sources of airborne microorganisms. Bacteria were first collected from different putative source environments and SIMPER was used to identify the taxa that primarily provide a discrimination between bacterial communities at each environment. These ‘indicator taxa’ were then used to assess the relative contribution of each source to the community in each airborne sample (Bowers et al. 2011a,b, 2012).

Consistent results from different studies allowed putting forward that, in continental and temperate regions, plants are important sources of atmospheric bacteria, especially in warm seasons. This is suggested by the abundance of plant-associated bacteria, such as *Sphingomonadales* and *Rhizobiales*, and the high number of chloroplast sequences retrieved in samples collected in warm periods. Conversely, soil-inhabiting bacteria, such as *Actinobacteriales*, *Firmicutes* and *Burkholderiales*, prevailed in the atmosphere during cold seasons. Bacteria of marine origin were obviously abundant in coastal environments (Cho and Hwang 2011; Fahlgren et al.

2010). Moreover, even in outdoor environments researchers are increasingly focusing their studies on the identification of the sources of specific bacterial populations that may negatively affect human health. For example, obligate anaerobes (*including Bacteroidales, Clostridiales and Fusobacteria*) were interpreted as indicator taxa for aerosolization of dog faeces in urban areas of the Midwestern United States (Bowers et al. 2011b). However, the overall similarity between bacterial communities sampled in different sites and seasons suggested that a fraction of airborne microorganisms may be ubiquitously distributed due to long range transport of bacteria (Bowers et al. 2011a; Marchant et al. 2008). DeLeon-Rodriguez et al. (2013) hypothesized the existence of a 'background' bacterial community at high-altitude in the upper troposphere. Furthermore, Fierer et al. (2008) remarked that abundant airborne taxa are also common in other cold and oligotrophic environments, such as ice and polar soils, thus suggesting that bacteria living in these extreme environments share particular features which promote their ability to survive in the atmosphere.

In some cases, the source area of airborne bacteria was identified through backward trajectory analyses. Indeed, temporal shifts in microbial community composition could sometimes be related to the different origin of air masses (Fahlgren et al. 2010; Jeon et al. 2011; Zweifel et al. 2012), but this was possible only when air masses at the same location originated from clearly different environments (e.g. marine or continental areas, agricultural, forest or desert areas) in different days.

5.4 Future perspectives

Despite that the use of NGS technologies has significantly improved our knowledge on the dynamics and sources of airborne microbial populations, researches in this field of microbiology are still in their infancy. Common patterns in dominant bacterial populations in the atmosphere have been independently retrieved in different studies. Therefore, a microbial background seems to exist in outdoor near-surface atmosphere and seems to be shared among different environments (e.g. urban, rural, marine, etc.) and geographic locations. Importantly, these dominant bacterial

populations were consistently identified in studies based on different sampling and analysis methods. However, the adoption of homogeneous protocols and experimental designs would be helpful for data comparison among different research groups. Several papers have also identified recurrent local sources of bacteria in outdoor near-surface atmosphere. A shared method to perform this task is still lacking, but the identification of bacterial taxa as indicators of specific source environments may represent a valid tool for future investigations (Bowers et al. 2011b).

Long-range transport of microorganisms in the atmosphere is another topic of air microbiology that is still open to further research advances. In fact, it affects the structure of local microbial communities and their ecological role, as well as the biogeography of microorganisms. Particularly, transport of microorganisms in the upper troposphere may have meteorological and climatic implications, which are still far to be fully elucidated (DeLeon-Rodriguez et al. 2013; Morris et al. 2011; Smith et al. 2012). The existence of large biogeographic air regions, from which different bacteria might be originated, has also been hypothesized but not definitively proved (Womack et al. 2010).

Although the presence of a metabolic activity of some airborne bacteria appears to be plausible (see Womack et al. 2010 and references therein), the occurrence of ecological interactions among them is still debated. Thus, significant gaps remain in our knowledge of the potential functions of the airborne microbial community and of the processes through which these functions are affected by environmental conditions. Metagenomics and metatranscriptomics have been increasingly used to address similar questions in several environments other than the atmosphere. On the contrary, to the best of our knowledge, only one airborne metagenome has been published so far in a study describing an urban indoor environment in Singapore (Tringe et al. 2008). The application of these techniques also to air microbiology has therefore the potential to supply ecological information on genetic adaptations of microorganisms to the atmospheric environment. Furthermore, it would also disclose

whether the extreme conditions of the atmosphere may have selected bacterial genes with potential biotechnological applications.

5.5 Acknowledgements

Our research on airborne bacteria was partially funded by Cariplo Foundation in the frame of the Project TOSCA (Toxicity of particulate matter and molecular markers of risk), and by the Italian Ministry of Research – MIUR (Grant: PRIN 2010-2011 - 2010WLNFY2_005).

5.6 References

Al-Bader D, Elias M, Rayan R, Radwan S (2012) Air-dust-borne associations of phototrophic and hydrocarbon-utilizing microorganisms: promising consortia in volatile hydrocarbon bioremediation. *Environ Sci Pollut Res* 19, 3997-4005.

An HR, Mainelis G, White L (2006) Development and calibration of real-time PCR for quantification of airborne microorganisms in air samples. *Atmos Environ* 40, 7924-7939.

Becker S, Dailey LA, Soukup JM, Grambow SC, Devlin RB, Huang YCT (2005) Seasonal variations in air pollution particle-induced inflammatory mediator release and oxidative stress. *Environ Health Persp* 113, 1032-1038.

Bertolini V, Gandolfi I, Ambrosini R, Bestetti G, Innocente E, Rampazzo G, Franzetti A (2013) Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy. *Appl Microbiol Biotechnol* online first doi: 10.1007/s00253-012-4450-0

Bowers RM, Lauber CL, Wiedinmyer C, Hamady M, Hallar AG, Fall R, Knight R, Fierer N (2009) Characterization of airborne microbial communities at a high-elevation site

and their potential to act as atmospheric ice nuclei. *Appl Environ Microbiol* 75, 5121-5130.

Bowers RM, McCubbin IB, Hallar AG, Fierer N (2012) Seasonal variability in airborne bacterial communities at a high-elevation site. *Atmos Environ* 50, 41-49.

Bowers RM, McLetchie S, Knight R, Fierer N (2011a) Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of the potential source environments. *ISME J* 5, 601-612.

Bowers RM, Sullivan AP, Costello EK, Collett Jr JL, Knight R, Fierer N (2011b) Sources of bacteria in outdoor air across cities in the Midwestern United States. *Appl Environ Microbiol* 77, 6350-6356.

Brodie EL, De Santis TZ, Moberg Parker JP, Zubietta IX, Piceno YM, Andersen GL (2007) Urban aerosols harbor diverse and dynamic bacterial populations. *P Natl Acad Sci USA* 104,299-304.

Buttner MP, Willeke K, Grinshpun SA (2002) Sampling and analysis of airborne microorganisms. In: Hurst CJ, Crawford RL, Knudsen G, McInerney M, Stetzenbach LD (eds) *Manual of environmental microbiology*, 2nd edn. ASM Press, Washington DC, pp. 814-826.

Camatini M, Corvaja V, Pezzolato E, Mantecca P, Gualtieri M (2012) PM10-biogenic fraction drives the seasonal variation of proinflammatory response in A549 cells. *Environ Toxicol* 27, 63-73.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-

throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621–1624.

Cho BC, Hwang CY (2011) Prokaryotic abundance and 16S rRNA gene sequences detected in marine aerosols on the East Sea (Korea). *FEMS Microbiol Ecol* 76, 327–341.

Clarke KR, Gorley RN (2006) *PRIMER v6: User Manual/Tutorial*. PRIMER-E, Plymouth.

Cote V, Kos G, Mortazavi R, Ariya PA (2008) Microbial and de novo transformation of dicarboxylic acids by three airborne fungi. *Sci Total Environ* 390, 530–537.

D'Amato G (2002) Environmental urban factors (air pollution and allergens) and the rising trends in allergic respiratory diseases. *Allergy* 57 Suppl 72, 30–33.

D'Arcy N, Canales M, Spratt DA, Lai KM (2012) Healthy schools: standardisation of culturing methods for seeking airborne pathogens in bioaerosols emitted from human sources. *Aerobiologia* 28, 413–422.

DeLeon-Rodriguez N, Lathem TL, Rodriguez-R LM, Barazesh JM, Anderson BE, Beyersdorf AJ, Ziemba LD, Bergin M, Nenes A, Konstantinidis T (2013) Microbiome of the upper troposphere: species composition and prevalence, effects of tropical storms, and atmospheric implications. *P Natl Acad Sci USA* online first.

Després VR, Nowoisky JF, Klose M, Conrad R, Andreae MO, Pöschl U (2007) Characterization of primary biogenic aerosol particles in urban, rural, and high-alpine air by DNA sequence and restriction fragment analysis of ribosomal RNA genes. *Biogeosciences* 4, 1127–1141.

Di Giorgio C, Krempff A, Guiraud H, Binder P, Tiret C, Dumenil G (1996) Atmospheric pollution by airborne microorganisms in the city of Marseilles. *Atmos Environ* 30, 155-160.

Dungan RS (2010) Board-invited review: Fate and transport of bioaerosols associated with livestock operations and manures. *J Anim Sci* 88, 3693-3706.

Dungan RS, Leytem AB (2009) Qualitative and quantitative methodologies for determination of airborne microorganisms at concentrated animal-feeding operations. *World J Microb Biot* 25, 1505-1518.

Eames I, Tang JW, Li Y, Wilson P (2009) Airborne transmission of disease in hospitals. *J R Soc Interface* 6, S697-S702.

Eduard W, Heederik D, Duchaine C, Green BJ (2012) Bioaerosol exposure assessment in the workplace: the past, present and recent advances. *J Environ Monitor* 14, 334-339.

Fahlgren C, Bratbak G, Sandaa RA, Thyraug R, Zweifel UL (2011) Diversity of airborne bacteria in samples collected using different devices for aerosol collection. *Aerobiologia* 27, 107-120.

Fahlgren C, Hagström Å, Nilsson D, Zweifel UL (2010) Annual variations in the diversity, viability, and origin of airborne bacteria. *Appl Environ Microbiol* 76, 3015-3025.

Fang Z, Ouyang Z, Zheng H, Wang X, Hu L (2007) Culturable airborne bacteria in outdoor environments in Beijing, China. *Microb Ecol* 54, 487-496.

Fierer N, Liu Z, Rodríguez-Hernández M, Knight R, Henn M, Hernandez MT (2008) Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microbiol* 74, 200-207.

Finnerty K, Choi JE, Lau A, Davis-Gorman G, Diven C, Seaver N, Linak WP, Witten M, McDonagh PF (2007) Instillation of coarse ash particulate matter and lipopolysaccharide produces a systemic inflammatory response in mice. *J Toxicol Environ Health A* 70, 1957–1966.

Fong NJC, Burgess ML, Barrow KD, Glenn DR (2001) Carotenoid accumulation in the psychrotrophic bacterium *Arthrobacter agilis* in response to thermal and salt stress. *Appl Microbiol Biotechnol* 56, 750-756.

Franzetti A, Gandolfi I, Gaspari E, Ambrosini R, Bestetti G (2011) Seasonal variability of bacteria in fine and coarse urban air particulate matter. *Appl Microbiol Biotechnol* 90, 745–753.

Fuzzi S, Mandrioli P, Perfetto A (1997) Fog droplets—an atmospheric source of secondary biological aerosol particles. *Atmos Environ* 31, 287–290.

Gandolfi I, Franzetti A, Bertolini V, Gaspari E, Bestetti G (2011) Antibiotic resistance in bacteria associated with coarse atmospheric particulate matter in an urban area. *J Appl Microbiol* 110, 1612-1620.

Georgakopoulos DG, Després V, Fröhlich-Nowoisky J, Psenner R, Ariya PA, Pósfai M, Ahern HE, Moffett BF, Hill TCJ (2009) Microbiology and atmospheric processes: biological, physical and chemical characterization of aerosol particles. *Biogeosciences* 6, 721-737.

Griffin DW (2007) Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clin Microbiol Rev* 20, 459-477.

Griffin DW, Gonzalez C, Teigell N, Petrosky T, Northup DE, Lyles M (2011) Observations on the use of membrane filtration and liquid impingement to collect airborne microorganisms in various atmospheric environments. *Aerobiologia* 27, 25-35.

Haas D, Galler H, Luxner J, Zarfel G, Buzina W, Friedl H, Marth E, Habib J, Reinthaler FF (2013) The concentrations of culturable microorganisms in relation to particulate matter in urban air. *Atmos Environ* 65, 215-222.

Harrison RM, Jones AM, Biggins PDE, Pomeroy N, Cox CS, Kidd SP, Hobman JL, Brown NL, Beswick A (2005) Climate factors influencing bacterial count in background air samples. *Int J Biometeorol* 49, 167-178.

Hospodsky D, Yamamoto N, Peccia J (2010) Accuracy, precision and method detection limits of quantitative PCR for airborne bacteria and fungi. *Appl Environ Microbiol* 76, 7004-7012.

Jeon EM, Kim HJ, Jung K, Kim JH, Kim MY, Kim YP, Ka JO (2011) Impact of Asian dust events on airborne bacterial community assessed by molecular analyses. *Atmos Environ* 45, 4313-4321.

Korzeniewska E (2011) Emission of bacteria and fungi in the air from wastewater treatment plants - a review. *Front Biosci* 3, 393-407.

Kuske CR (2006) Current and emerging technologies for the study of bacteria in the outdoor air. *Curr Opin Biotechnol* 17, 291-296.

Lee SH, Lee HJ, Kim SJ, Lee HM, Kang H, Kim YP (2010) Identification of airborne bacterial and fungal community structures in an urban area by T-RFLP analysis and quantitative real-time PCR. *Sci Total Environ* 408, 1349-1357.

Li K, Dong S, Wu Y, Yao M (2010) Comparison of the biological content of air samples collected at ground level and at higher elevation. *Aerobiologia* 26, 233-244.

Lighthart B (1997) The ecology of bacteria in the alfresco atmosphere. *FEMS Microbiol Ecol* 23, 263-274.

Lighthart B, Shaffer BT (1995a) Viable bacterial aerosol particle size distributions in the midsummer atmosphere at an isolated location in the high desert chaparral. *Aerobiologia* 11, 19–25.

Lighthart B, Shaffer BT (1995b) Airborne bacteria in the atmosphere surface layer: temporal distribution above a grass seed field. *Appl Environ Microbiol* 61, 1492–1496.

Longhin E, Pezzolato E, Mantecca P, Holme JA, Franzetti A, Camatini M, Gualtieri M (2013) Season linked responses to fine and quasi-ultrafine Milan PM in cultured cells. *Toxicol in Vitro* 27, 551–559.

Mahdy HM, El-Sehrawi MH (1997) Airborne bacteria in the atmosphere of El-Taif region, Saudi Arabia. *Water Air Soil Poll* 98, 317-324.

Mancinelli RL, Schulls WA (1978) Airborne bacteria in an urban environment. *Appl Environ Microbiol* 35, 1095-1101.

Marchant R, Franzetti A, Pavlostathis SG, Tas DO, Erdbrugger I, Unyayar A, Mazmanci MA, Banat IM (2008) Thermophilic bacteria in cool temperate soils: are they

metabolically active or continually added by global atmospheric transport? *Appl Microbiol Biotechnol* 78, 841-852.

Maron PA, Lejon DPH, Carvalho E, Bizet K, Lemanceau P, Ranjard L, Mougel C (2005) Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. *Atmos Environ* 39, 3687-3695.

Maron PA, Mougel C, Lejon DPH, Carvalho E, Bizet K, Marck G, Cubito N, Lemanceau P, Ranjard L (2006) Temporal variability of airborne bacterial community structure in an urban area. *Atmos Environ* 40, 8074-8080.

Millner PD (2009) Bioaerosols associated with animal production operations. *Bioresource Technol* 100, 5379-5385.

Moorman JE, Zahran H, Truman BI, Molla MT (2011) Current asthma prevalence—United States, 2006–2008. *MMWR Surveill Summ* 60 Suppl:84–86.

Morris CE, Sands DC, Bardin M, Jaenicke R, Vogel B, Leyronas C, Ariya PA, Psenner R (2011) Microbiology and atmospheric processes: research challenges concerning the impact of airborne micro-organisms on the atmosphere and climate. *Biogeosciences* 8, 17-25.

O'Hara RE, Rubin R (2005) Reducing bioaerosol dispersion from wastewater treatment and its land application: a review and analysis. *J Environ Health* 68, 24-29.

Peccia J, Hernandez M (2006) Incorporating polymerase chain reaction-based identification, population characterization, and quantification of microorganisms into aerosol science: a review. *Atmos Environ* 40, 3941-3961.

Polymenakou PN (2012) Atmosphere: a source of pathogenic or beneficial microbes? Atmosphere 3, 87-102.

Polymenakou PN, Mandalakis M, Stephanou EG, Tselepidis A (2008) Particle size distribution of airborne microorganisms and pathogens during an intense African dust event in the Eastern Mediterranean. Environ Health Persp 116, 292-296.

Ravva SV, Hernlem BJ, Sarreal CZ, Mandrell RE (2012) Bacterial communities in urban aerosols collected with wetted-wall cyclonic samplers and seasonal fluctuations of live and culturable airborne bacteria. J Environ Monit 14, 473-481.

Rinsoz T, Duquenne P, Greff-Mirguet G, Oppliger A (2008) Application of real-time PCR for total airborne bacterial assessment: comparison with epifluorescence microscopy and culture dependent methods. Atmos Environ 42, 6767-6774.

Sattler B, Puxbaum H, Psenner R (2001) Bacterial growth in supercooled cloud droplets. Geophys Res Lett 28, 239-242.

Shaffer BT, Lighthart B (1997) Survey of culturable airborne bacteria at four diverse locations in Oregon: urban, rural, forest, and coastal. Microb Ecol 34, 167-177.

Smith DJ, Jaffe DA, Birmele MN, Griffin DW, Schuerger AC, Hee J, Roberts MS (2012) Free tropospheric transport of microorganisms from Asia to North America. Microb Ecol 64, 973-985.

Stetzenbach LD, Buttner MP, Cruz P (2004) Detection and enumeration of airborne biocontaminants. Curr Opin Microbiol 15, 170-174.

Tang JW (2009) The effect of environmental parameters on the survival of airborne infectious agents. J R Soc Interface 6, S737-S746.

Temkiv TS, Finster K, Hansen BM, Nielsen NW, Karlson UG (2012). The microbial diversity of a storm cloud as assessed by hailstones. *FEMS Microbiol Ecol* 81, 684-695.

Tong YY, Lighthart B (1997) Solar radiation is shown to select for pigmented bacteria in the ambient outdoor atmosphere. *Photochem Photobiol* 65, 103-106.

Tringe SG, Zhang T, Liu X, Yu Y, Lee WH, Yap J, Yao F, Suan ST, Ing SK, Haynes M, Rohwer F, Wei CL, Tan P, Bristow J, Rubin EM, Ruan Y (2008) The airborne metagenome in an indoor urban environment. *PLoS ONE* 3, 10.

Urbano R, Palenik B, Gaston CJ, Prahter KA (2011) Detection and phylogenetic analysis of coastal bioaerosols using culture dependent and independent techniques. *Biogeosciences* 8, 301-309.

Väitilingom M, Attard E, Gaiani N, Sancelme M, Deguillaume L, Flossmann AI, Amato P, Delort AM (2012) Long-term features of cloud microbiology at the puy de Dôme (France). *Atmos Environ* 56, 88-100.

Womack AM, Bohannan BJM, Green JL (2010) Biodiversity and biogeography of the atmosphere. *Philos Trans R Soc B Biol Sci* 365, 3645–3653.

Xu Z, Wu Y, Shen F, Chen Q, Tan M Yao M (2011) Bioaerosol science, technology, and engineering: past, present, and future. *Aerosol Sci Tech* 45, 1337–1349.

Zweifel UL, Hagström Å, Holmfeldt K, Thyrraug R, Geels C, Frohn LM, Skjøth CA, Karlson UG (2012) High bacterial 16SrRNA gene diversity above the atmospheric boundary layer. *Aerobiologia* 28, 481-498.

CHAPTER 6

**Bacterial community structure on two
alpine debris-covered glaciers and
biogeography of *Polaromonas*
phylotypes**

6.1 Introduction

Debris-covered glaciers (DCGs) are mountain glaciers (that is, all the glaciers but ice sheets) whose ablation area is covered mostly by debris (Benn and Evans 2010). DCGs are common on mountain ranges of Asia (Himalaya, Karakoram and Tien Shan), South America (Andes), Alaska, New Zealand and Alps (Smiraglia et al. 2000; Diolaiuti et al. 2003; Mihalcea et al. 2008; Benn and Evans 2010). Thick supraglacial debris reduces the rate and magnitude of buried ice ablation (Østrem 1959; Nakawo and Rana 1999), but the debris surface can be heated by solar radiation to temperatures that can exceed $\pm 30\text{ }^{\circ}\text{C}$ (Brock et al. 2010). The debris cover of DCGs is characterized by continuous rock substrate inputs at the highest part of the glacier. The debris is mainly composed of clasts, ranging in size from millimetres to metres, which fall on the glacier surface from surrounding mountains. The debris of DCGs therefore differs from the fine windblown particles that constitute the cryoconite, the most abundant debris in the central parts of debris-free glaciers (Laybourn-Parry et al. 2012). The debris is then transported down valley for times that, on some glaciers, can be as long as a century (Pelfini et al. 2007). As a result, the glacier surface is covered by a continuous debris layer, whose thickness generally increases toward the glacier terminus. The long transport on the glacier surface allows debris weathering and alteration, and its colonization not only by microorganisms but also by animals (for example, arthropods) and plants (Pelfini et al. 2007, 2012; Caccianiga et al. 2011; Gobbi et al. 2011). Ecological communities therefore exist on the surface of DCGs, and may be structured according to a chronosequence, with communities increasing in complexity towards the glacier terminus (Gobbi et al. 2011).

The rock coverage of DCGs, by virtue of its ecological features, allows the investigation of ecological hypotheses about microbial colonization and succession of microbial communities in extreme environments, but has been neglected so far in ecological and microbiological studies. Indeed, in Alpine environments, microbial

succession studies have typically been conducted only on glacial forefields, where the distance from the receding glacier front served as a proxy for time since organism colonization (Sigler and Zeyer 2002; Nicol et al. 2005; Nemergut et al. 2007; Schmidt et al. 2008; Sattin et al. 2009; Schutte et al. 2009,2010; Philippot et al. 2011).

In this study, we collected, at increasing distances from glacier terminus, 38 samples of the debris coverage of two Italian DCGs, the Miage (Valle d' Aosta—Italy) and the Belvedere (Piemonte—Italy) glaciers, and used the Illumina technology to provide a thorough description of the microbial communities, by 16S rRNA gene tag-deep sequencing. The first aim of this study was therefore to verify two hypotheses about microbial communities in the extreme environment of DCGs: (1) that the structure of the microbial communities along the glacier surface changes according to time of exposure of the debris, and (2) that microbial successions of these environments are dominated by autotrophic bacteria in the first stages and by heterotrophic bacteria in later stages.

High-elevation cold environments are considered ideal model systems for microbial biogeographical studies, as they are 'extreme' ecosystems, are geographically widespread and separated to one another, usually contain low microbial diversity and are linked to one another by the movements of cold dry air masses (Darcy et al. 2011). The bacterial genus *Polaromonas* is among the dominant bacterial taxa in glacial ice and sediments worldwide, and its presence has been reported particularly in recently deglaciated substrates (Frey et al. 2010; Darcy et al. 2011; Michaud et al. 2012). It is therefore a suitable model taxon for investigating biogeographical patterns of distribution of microorganisms. This study also aims at analysing *Polaromonas* sequences found on the Miage and Belvedere glaciers together with others published in the literature (Darcy et al. 2011) to investigate whether *Polaromonas* phylotypes do present a biogeographical distribution at global scale, as suggested recently (Darcy et al. 2011).

6.2 Material and Methods

6.2.1 Study sites and environmental data

The ablation zone of Miage glacier (451 470N, 061 520E, Mont Blanc massif) is debris covered from about 2400 m above sea level (a.s.l.) to the terminus at 1720 m a.s.l. The debris cover is mainly composed by gneisses, schistes and granites (Brock et al., 2010), and ranges in thickness between a few centimetres to more than 1 m at the terminus. Belvedere glacier (451 570N, 41 340E, Monte Rosa massif) is debris covered from about 2300 m a.s.l. to the terminus at 1750 m a.s.l. Rock debris is mainly composed by gneisses, micaschists and granites, and ranges in thickness between a few centimetres to more than 1 m at the terminus (Diolaiuti et al. 2003). Debris thickness on both glaciers was estimated by analysing, according to empirical models, kinetic surface temperature maps (pixel size 90 x 90 m²) acquired by the ASTER satellite on 5 August 2009 (Belvedere glacier) and 12 August 2009 (Miage glacier) at 10:40 am local solar time (Taschner and Ranzi 2002; Mihalcea et al. 2008). Annual surface velocity of both glaciers was measured by the differential global positioning system method (Diolaiuti et al. 2005; Caccianiga et al. 2011). Surface elevation was extracted from 2003 Digital Elevation Models. The organic carbon (OC) content of sampled debris was quantified by the dichromate oxidation and titration (Walkley and Black method; Schumacher 2002) on the clastic fraction 0.2 mm, and pH was measured on the same clastic fraction in a distilled water soil suspension (1 : 2.5 soil : water ratio) by a portable pH meter (mod. HI 991300; Hanna Instruments, Woonsocket, RI, USA).

The two glaciers significantly differ in surface temperature, slope at sampling sites and pH values of sampled debris (see Supplementary Information for further details on glaciological features).

6.2.2 Debris sampling and DNA extraction

On both glaciers, samples were aseptically collected at the surface of the debris cover at approximately constant intervals (elevation and distance) from 1794 to 2043 m a.s.l. on the Miage glacier and from 1803 to 2085 m a.s.l. on the Belvedere glacier (Supplementary Table S1, Figure 1) both in July 2009 (the 28th on Miage and the 2nd on Belvedere) and September 2009 (the 10th on Miage and the 18th on Belvedere). Large clasts were discarded during sampling. The sampling point position was determined in July by a global positioning system (Oregon 300; Garmin Ltd, Olate, KS, USA). Global positioning system was used to recover the sampling sites in September. In the field, each sampling site was also visually recognized by the aid of a picture of the sampling location taken during the first visit in July. Samples were, therefore, taken at (approximately) the same site, with the only exception of sample 4 on the Miage glacier, which was sampled in different positions in July and September because the sampling point of July could not be reached in September because of a crevasse. In all, 9 samples were aseptically collected on Miage glacier and 10 on Belvedere glacier. Debris samples were kept at 4 °C during transport to the laboratory, which occurred within 12 h. Samples were then divided into two aliquots: one was stored at -20°C for DNA extraction and the other at +4 °C for the analyses of OC and pH. No replicated samples were taken at each site because multivariate regression-like analyses (see Statistical methods) were used to investigate the change in the structure of microbial communities along an ecological gradient (Lennon 2011). DNA was extracted from 0.5 g of debris using FastDNA Spin for soil Kit (MP Biomedicals, Solon, OH, USA) following the manufacturer's instructions. Amplification tests were carried out on the extracts and ensured that no inhibition occurred during amplification of the different DNA samples.

Further information on field and laboratory procedures is available in the Supplementary Information.

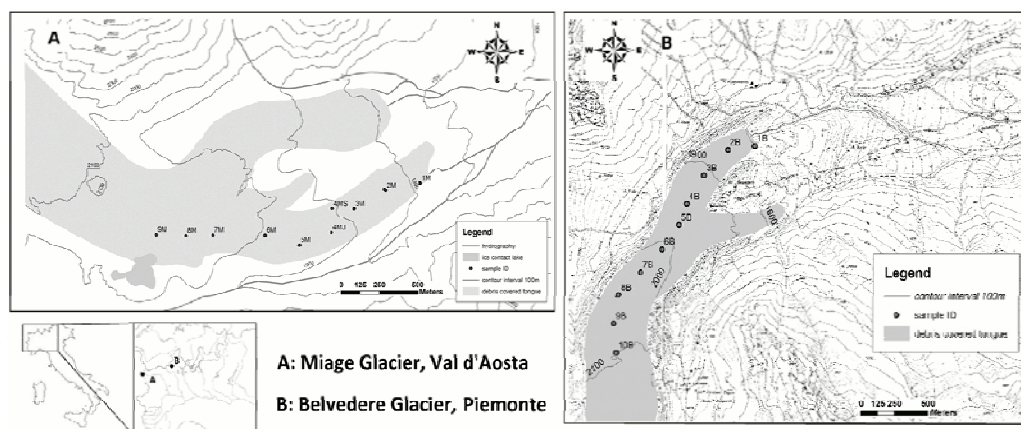


Figure 1 The Miage (A) and Belvedere (B) glacier, with sampling points.

6.2.3 16S rRNA gene fragment sequencing

The V5 and V6 hypervariable regions of the 16S rRNA gene were polymerase chain reaction amplified using 783F and 1027R primers (Huber et al. 2007; Wang and Qian 2009; see also the Supplementary Information), for Genome AnalyzerIIx sequencing. Six to eight purified amplicons with different barcodes at 50 end of the forward primer (Supplementary Table S2) were pooled in 100 ml samples with a DNA concentration of 50 ng ml⁻¹. Multiplexed sequencing of all the pooled samples was performed on a single Illumina GA-IIx lane, using a paired-end 76 bp protocol and the 4.0 sequencing chemistry.

Illumina Real-Time Analysis software, version 1.8.7 (San Diego, CA, USA), was used to perform the cluster extraction and base-calling processing analyses.

6.2.4 Sequence processing and data analysis

Each sequence was assigned to its original sample according to its barcode. After sorting, the reverse read of each paired-end sequence was reverse-complemented and merged with the corresponding forward read, inserting 20 Ns in between (Claesson et al. 2010). A quality cutoff was then applied to remove sequences that did not contain the barcode or with an average base quality value (Q) lower than 30. The barcode was removed from sequences before further processing.

The taxonomic attribution of the filtered sequences was carried out using the stand-alone version of Ribosomal Database Project Bayesian classifier, using 50% confidence cutoff as recently suggested for sequences shorter than 200 bp (Wang et al. 2007; Claesson et al. 2009). Operational taxonomic units (OTUs) were defined on the basis of the Ribosomal Database Project classification, considering the fourth taxonomic level, which for most bacteria corresponds to the Order.

The sequences classified as *Polaromonas* spp. by Ribosomal Database Project classifier with a confidence level $\geq 80\%$ were selected and clustered by UCLUST algorithm through QIIME interface (Caporaso et al., 2010) with the following parameter values: max_accept ≥ 0 , max_reject ≥ 0 , sort by abundance and sequence identity 99%. The GenBank *Polaromonas* sequences (Darcy et al. 2011) were used as reference database. Among the sequences clustered on the seed reference sequences, we selected the unique *Polaromonas* sequences from the Miage and Belvedere glaciers, respectively. The use of unique sequences at each glacier avoids inflating the effect of samples collected at very short geographical distances. The 43 unique within-glacier *Polaromonas* sequences obtained were used in the analyses of the biogeography of *Polaromonas* genus (see Statistical methods) along with 55 sequences from 15 glaciers throughout the world published in Darcy et al. (2011) and obtained from the GenBank (Benson et al. 2010). *Polaromonas* sequences were aligned using SILVA database as template and the genetic distances between them were computed with MOTHUR (Schloss et al. 2009). The source codes used for sequence processing and data formatting are available upon request from the authors.

6.2.5 Statistical methods

To compare samples that largely differ in the number of sequences (6047–42 864, see also Results), 6000 reads were randomly selected from all libraries and used in the analyses of the structure of microbial communities. Such number of sequences allows

an accurate determination of b-diversity among microbial communities (Caporaso et al. 2012). In addition, accumulation curves at each sampling site suggest that a-diversity could be captured with 6000 reads per sample, at least at the fourth taxonomic level (details not shown).

The number of sequences belonging to each OTU was considered an estimate of bacterial abundance at each sampling site. Ecological determinants of the structure of microbial communities were investigated by redundancy analysis (Legendre and Legendre 1998; Borcard et al. 2011) and hierarchical cluster analysis with the complete linkage method, on Hellinger distances between bacterial communities. This distance was chosen as it is metric, depends on the difference in the proportion of OTUs between samples, decreases the importance of OTU abundance over occurrence and avoids the double-zero problem when comparing OTU composition between samples (Legendre and Legendre 1998; De Caceres et al. 2010).

We preliminarily investigated the correlations among environmental variables to avoid entering highly correlated predictors in the statistical models (see Supplementary Information for details on variables excluded from the analyses). Predictors entered in models were distance from the glacier terminus, pH, two dichotomous variables indicating the glacier (Belvedere or Miage) and the month (July or September), slope, ablation rate, debris thickness and amount of OC in the sample. Models were simplified by backward selection of non-significant variables. Analyses were performed with the VEGAN package (Oksanen et al. 2009) and the HCLUST procedure in R 2.8.1 (R Development Core Team 2008).

Difference in OTU abundance among glaciers or months was tested by t-tests, and variation in OTU abundance according to other environmental variables was tested by correlations. P-values were corrected for multiple testing according to the false discovery rate (FDR) procedure (Benjamini and Yekutieli 2001) using the MULTTEST package in R 2.8.1.

For the analysis of the biogeography of *Polaromonas*, we calculated the matrix of great-circle distances among glaciers using in-house procedures (see Ambrosini et al.

2009). The matrix of genetic distances among sequences (see above) was then related to the matrix of geographic distances by means of Mantel tests and Mantel correlograms (corrected by the FDR procedure). Sampling sites were unevenly distributed worldwide (Supplementary Table S4 and Supplementary Figure S3). Breaks among distance classes in the Mantel correlogram were therefore arbitrary set to values corresponding to gaps in the distribution of distances among glaciers. The number of genetic distances in each geographical distance class ranged between 315 and 945, thus assuring adequate power to all statistical tests. See also Supplementary Information for why this procedure was preferred over alternative ones.

6.2.6 Data accessibility

Demultiplexed fastq-formatted DNA sequences have been deposited to Sequence Read Archive, study accession number ERP001386 (<http://www.ebi.ac.uk/ena/data/view/ERP001386>).

Fasta-formatted DNA sequences of *Polaromonas* have been uploaded as Supplementary Information.

6.3 Results

After sorting and quality filtering, the number of 2 x 76 bp paired-end sequences obtained from the sequencing was 829 657, while the number of sequences of each library ranged from 6047 to 42 864 (Supplementary Table S3).

As reported in previous works, the number of sequences classified at 50% confidence level by the Bayesian classifier at a given taxonomic level sharply decreased from Domain to Genus level owing to the detection of unknown bacteria (Roesch et al. 2007; Franzetti et al. 2011). The classification of all the filtered sequences is reported at the fourth taxonomic level (Order; Figure 2 and Supplementary Figure S1) to take into account most of the sequences and to assure a classification accuracy 48.0% (Claesson et al. 2010). The total number of classified taxa in the sample ranged from 46 to 68 (Supplementary Table S3). At each sample, 33.9–51.7% of sequences were

not classified. In all the communities, *Actinobacteridae*, *Sphingobacteriales* and *Burkholderiales* were the dominant taxa and accounted for 450% of the classified sequences. As reported in Figure 2, at lower taxonomic levels most of the *Actinobacteridae* belonged to suborder Micrococccineae and *Propionibacterineae*; almost all (74–80%) of the *Sphingobacteriales* were classified as *Chitinophagaceae*, while *Comamonadaceae* family was the dominant taxon within the order Burkholderiales. Furthermore, most of the *Comamonadaceae* sequences belonged to genera *Polaromonas* (61% for Miage and 63% for Belvedere) and *Variovorax* (25% for Miage and 24% for Belvedere).

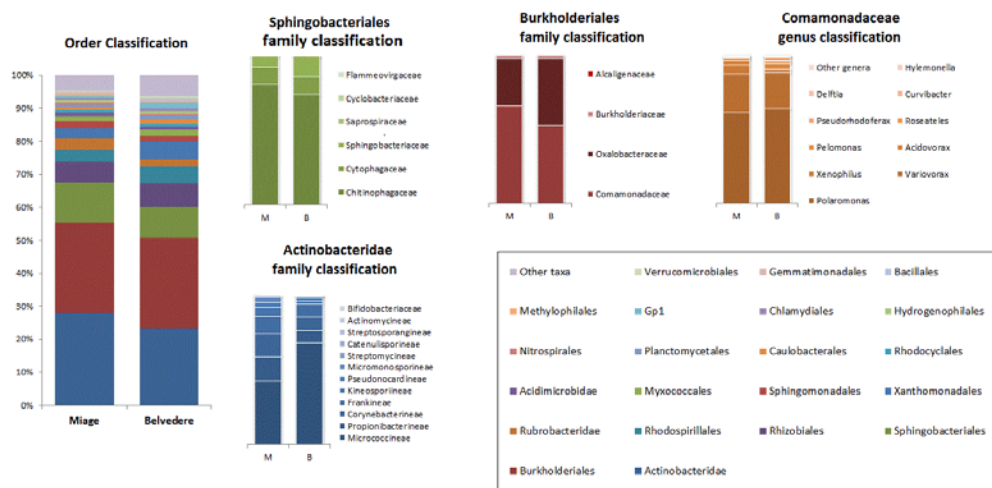


Figure 2 Taxonomic classification using a RDP Bayesian classifier (50% confidence) of the filtered sequences. Classification of all the filtered sequences is reported at the fourth taxonomic level. Classification at lower taxonomic levels is reported for *Actinobacteridae*, *Burkholderiales*, and *Sphingobacteriales*.

6.3.1 Structure of microbial communities

Cluster analysis partitioned samples into three main groups, which however did not correspond to any obvious subdivision according to glacier and/or season

(Supplementary Figure S2). Redundancy analysis indicated that the structure of microbial communities differed significantly between glaciers and months, and changed with increasing distance from the glacier terminus, amount of OC in the debris and ablation rate at sampling site (Table 1 and Figure 3).

Table 1: Final redundancy analysis of the structure of the microbial communities in relation to environmental factors. Significance was assessed by a randomization procedure (Legendre and Legendre 1998).

| Effect | df | Variance | pseudo-F | P |
|----------------------------|----|----------|----------|-------|
| Distance from the terminus | 1 | 0.001 | 1.809 | 0.047 |
| Glacier | 1 | 0.003 | 8.289 | 0.005 |
| Month | 1 | 0.001 | 1.916 | 0.035 |
| Organic C | 1 | 0.001 | 2.373 | 0.015 |
| Ablation rate | 1 | 0.001 | 2.247 | 0.010 |
| Residual | 32 | 0.011 | | |

Overall model significance: pseudo- $F_{5,32} = 4.137$, $P = 0.005$

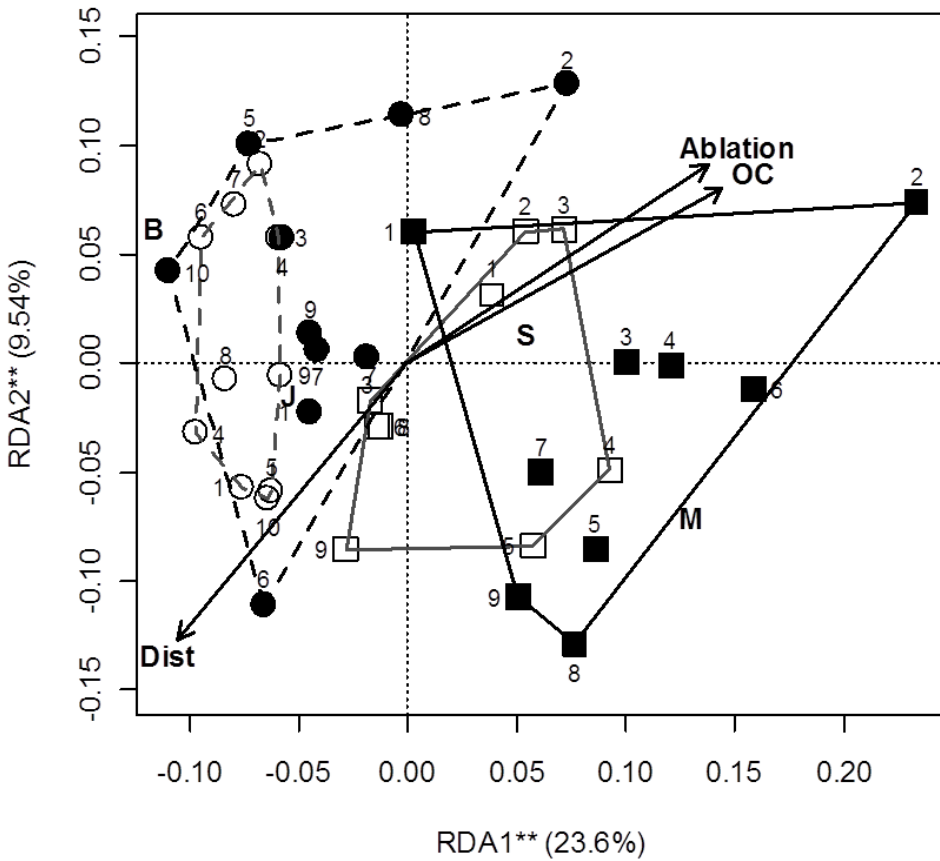


Figure 3 Two-dimensional redundancy analysis plot of microbial communities observed on Belvedere (circles and dashed lines) or Miage (squares and solid lines) glacier in July (open symbols and grey lines) and September (filled symbols and black lines). The proportion of variance explained by each axis is shown. Asterisks indicate statistical significance of axes (** = $P < 0.01$). Numbers denote sampling site at each glacier and month. Arrows represent vectors of increasing distance from the glacier terminus (Dist), organic carbon content in the debris (OC), or ablation rate (Ablation). Letters indicate centroids of categorical variables entered in the analyses (B = Belvedere glacier, M = Miage glacier, J = July, S = September).

Relative abundance of *Acidimicrobidae*, *Rubrobacteridae*, *Sphingobacteriales* and *Sphingomonadales* increased ($r \geq 0.515$, $n = 38$, $\text{PFDR} \leq 0.024$) at increasing levels of OC in the debris, and that of *Xantomonadales* significantly decreased ($r = -0.501$, $n = 38$, $\text{PFDR} \leq 0.031$) at increasing ablation rate. Distance from the terminus did not significantly affect the abundance of any OTU ($|r| \leq 0.453$, $n = 38$, $\text{PFDR} \geq 0.072$). Abundance of 16 OTUs significantly differed between Belvedere and Miage glaciers: *Acidimicrobidae*, *Actinobacteridae*, *Gemmatimonadales*, *Gp4*, *Nitrospirales*, *Rubrobacteridae*, *Sphingobacteriales* and *Sphingomonadales* were more abundant in the Miage glacier than in the Belvedere glacier ($t_{36} \leq -3.390$, $\text{PFDR} \leq 0.033$), while the reverse held true for *Caulobacterales*, *Chlamydiales*, *Clostridiales*, *Gp1*, *Opitutales*, *Planctomycetales*, *Rhodospirillales* and *Xanthomonadales* ($t_{36} \geq 3.600$, $\text{PFDR} \leq 0.024$). Finally, abundance of *Bacillales* differed significantly between months, being larger in July than in September ($t_{36} = 3.446$, $\text{PFDR} = 0.021$).

6.3.2 Biogeography of *Polaromonas* genus

The 43 *Polaromonas* sequences retrieved in this study were combined with the 55 sequences from Darcy et al. (2011), who reported *Polaromonas* sequences from 15 sites from the entire world (Supplementary Figure S3). We re-run the analyses carried out by Darcy with this expanded data set. Genetic distance among unique within-site *Polaromonas* sequences was positively correlated with geographic distance among sites ($r_M = 0.202$, $P < 0.001$, one-sided test). Mantel correlogram indicated that sequences sampled at ≤ 100 km was significantly more similar than those sampled at larger distance ($r_M = 0.117$, $\text{PFDR} = 0.007$, one-sided tests). The same held true for sequences retrieved in glaciers between 9400 and 13 500 km to each other ($r_M = 0.109$, $\text{PFDR} = 0.007$, one-sided tests; Figure 4). Shifting the cutoff limits of these two distance classes did not qualitatively change the results of the correlogram analysis (see Supplementary Information for details).

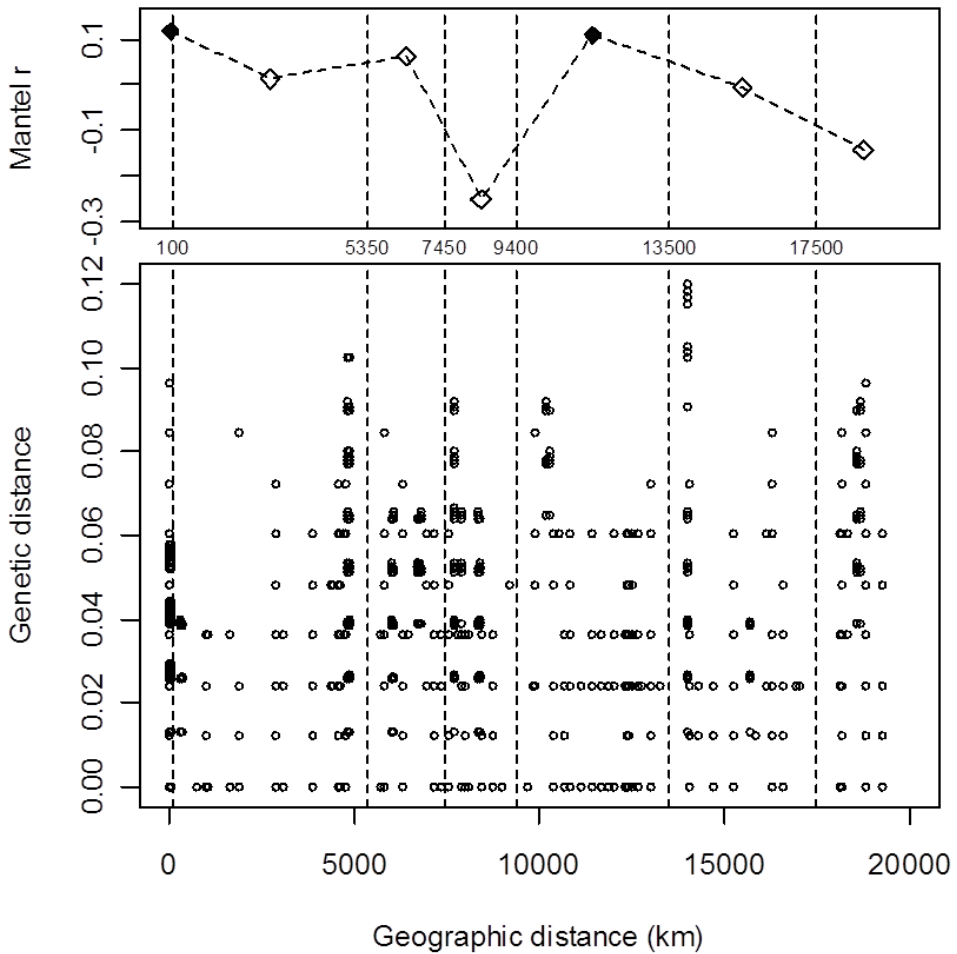


Figure 4 Genetic distance among *Polaromonas* sequences by geographic distance between glaciers. Lower panel shows all pairwise comparisons of genetic distance among *Polaromonas* sequences ($n = 4753$) and geographic distance among glaciers ($n = 17$). Upper panel shows the Mantel correlogram of genetic distance among the same sequences separated in six distance classes. The midpoint of each distance class is plotted. Shaded diamonds indicate significant tests in the Mantel correlogram. Vertical dashed lines indicate breakpoints between distance classes. The small numbers between the panels indicate the exact position of breakpoints.

6.4 Discussion

This work represents the first attempt to characterize the microbial communities on the debris cover of alpine DCGs and to determine possible environmental factors affecting their diversity.

Microbial communities of the debris were dominated by *Actinobacteridae*, *Sphingobacteriales* and *Burkholderiales*. The dominant taxon was the *Actinobacteridae* that represented 27% and 22% of the classified sequences at Miage and Belvedere glaciers, respectively. Most of these sequences were classified as *Micrococcineae* and *Propionibacterineae*, in accordance with recent studies specifically addressing the structure of *Actinobacteria* in cold environments (Jiang et al. 2010; Sanyika et al. 2012). Moreover, *Actinobacteridae* are ubiquitous in soils under very different climatic conditions (Madigan et al. 2009) and their high abundance is consistent with the presence of OC in the sediments, despite that the average OC content of debris sampled in Miage and Belvedere glaciers was 0.83 (0.72 s.d.) g kg⁻¹. This value is much lower than that of young soils (1–100 years) in European Alps, which range from 1.9 to 147.8 g kg⁻¹ (Egli et al., 2012).

Sphingobacteriales, which belongs to the phylum *Bacteroidetes*, is the only major taxon whose relative abundance significantly increased with increasing level of OC in the debris. This confirms the results from Fierer et al. (2007), where *Bacteroidetes* were positively correlated with C mineralization rates, which is a proxy of OC available to microorganisms. In addition, most of *Sphingobacteriales* sequences (74–80%) were classified as *Chitinophagaceae* at the family level (Figure 2). These microorganisms are able to degrade chitin and other complex polymeric organic matter (del Rio et al. 2010). *Chitinophagaceae* and other chitinolytic bacteria dominated the microbial communities of Antarctic soils and glacier forefields (Brankatschk et al. 2011; Ganzert et al. 2011). Chitin provides both C and N to the microorganisms (Moorhead et al. 2012) and may be supplied by arthropods, whose presence was already documented on the Miage glacier (Gobbi et al. 2011), and by fungi, which occur in recently deglaciated soils (Zumsteg et al. 2012). Interestingly, arthropods' communities

showed a turnover along the Miage glacier tongue, and richer and more structured communities occurred close to glacier terminus (Gobbi et al. 2011). Despite the fact that no data on the concentration of chitin in the debris cover of DCGs are available, we speculate that chitin might represent a relevant fraction of the OC in the debris of DCGs.

The majority of the *Burkholderiales* sequences retrieved in this study belonged to the *Comamonadaceae* family (65%) and, in particular, to the genus *Polaromonas*, which represented 5.6% (Miage glacier) and 5.8% (Belvedere glacier) of the total sequences classified at genus level (Figure 2). This confirms the results of Nemergut et al. (2007), who found high abundances of *Polaromonas* in recently deglaciated areas. According to the genomic and metagenomic data, *Polaromonas* possesses the dormancy-inducing gene *hipA*, which might enhance its ability to survive to low temperature and long-range transport (Darcy et al. 2011). Moreover, *Polaromonas* strains from cold soils are able to oxidize a wide range of substrates (see Darcy et al. (2011) and references therein) and it has been recently demonstrated that a *Polaromonas* strain is able to weather recently exposed rocks (Frey et al. 2010). Despite the fact that we have no direct evidence of these abilities in the local *Polaromonas* strains, we can speculate that all these features might provide *Polaromonas* with selective advantages over other microorganisms in colonizing new substrates and therefore contribute to its high abundances.

It has been postulated that bacterial autotrophs may play a crucial role in the initial stages of microbial colonization and ecosystem development on rock substrates due to the low nutrient content, few available OC and scarce vegetation of fresh substrates (Sigler et al. 2002). Autotrophic bacteria processes may have significant rock-weathering activity. For example, chemolithotrophic iron/sulphur oxidizers may contribute to weather pyritebearing rocks by lowering the pH even when present at low abundances in the microbial communities (Borin et al. 2010; Mapelli et al. 2011). Endolithic Cyanobacteria may also contribute to silica dissolution through alkalinization (Mapelli et al. 2012). Interestingly, a red surface patina can be observed

on Miage debris, whose lithology shows general similarities with the one described by Borin et al. (2010), but is less widespread on the Belvedere, whose debris has a slightly different lithology and a more acidic substrate (see Supplementary Information). However, no sequence classified as *Acidithiobacillales* (Order) was found either on Belvedere glacier or Miage glacier, and pH values of the samples (pH4.5) suggested a negligible contribution of these bacteria to rock-weathering.

The most described role of autotrophs in early succession stages is the CO₂ fixation. Particularly, some works suggested that *Cyanobacteria* are important sources of C in very early succession stages and some surveys did report high abundances of *Cyanobacteria* in early colonization stages (Nemergut et al. 2007; Schmidt et al. 2008; Zumsteg et al. 2012). On the contrary, other investigations showed low abundances of *Cyanobacteria* in glacier forefields without any relation between their abundance and soil age (Sigler et al. 2002; Sattin et al. 2009; Goransson et al. 2011; Knelman et al. 2012). Hence, the role of *Cyanobacteria* in the very early stages of ecosystem development is currently controversial. Our results add another piece of evidence against the old paradigm stating that these microorganisms are early pioneer organisms. In fact, *Cyanobacteria* were not among the dominant taxa in the debris, as they represented only 2.1% and 0.9% of the sequences classified at Class level at Belvedere and Miage glaciers, respectively.

Redundancy analysis showed that glacier, month, distance from the terminus, ablation rate and OC content significantly influenced the structure of the bacterial communities. Abundance of 16 taxa significantly differed from Belvedere and Miage glaciers. This is not surprising since the importance of local factors such as climatic conditions, bedrock composition, soil texture and pH in influencing bacterial diversity has been already documented (Mannisto et al. 2007; Lazzaro et al. 2009; Zumsteg et al. 2012). Particularly, the influence of pH was reported in several papers, although with contrasting results (Eskelinen et al. 2009; Lauber et al. 2009; Borin et al. 2010; King et al. 2010; Philippot et al. 2011; Knelman et al. 2012). Despite the fact that the microbial community structure did not show any significant relation with pH, the

debris of the Belvedere glacier was on average more acidic than that of the Miage glacier (see Supplementary Information). Hence, the observed difference in microbial community structure between the two glaciers may reflect a difference in average pH of the debris.

On a DCG, distance from glacier terminus, ablation rate and OC content of the debris are intimately related. Indeed, the rock coverage of a DCG is characterized by a continuous input of fresh, poor- in-OC debris from surrounding rocks in the upper part of the glacier. During downwards transport by glacier movements, OC content in the debris increases due to several concomitant mechanisms as a longer exposure of sediments to wind-blown organic matter, increased organic activity at lower altitudes and greater debris colonization by plants and animals. This leads to a gradient in the OC content of debris. Hence, distance from the terminus is a proxy of the time during which debris weathering and bacterial and vegetation colonization occur. Local ablation rate also increases toward the terminus, and is higher in areas where glacier movements determine a finer debris cover and even exposure of ice, for example, in areas with crevasses (Pelfini et al. 2012). Hence, OC content, ablation rate and distance from the terminus may probably be considered proxies of variation in general ecological features of the supraglacial debris that occurs during down valley transport. This interpretation is corroborated by the observations that distance from the terminus is represented by a vector opposite to those representing the other two variables in the redundancy analysis plot (Figure 3). The results from the multivariate analysis therefore suggest that glacier debris cover may host chronosequences similar to those observed in glacier forefields (Sigler and Zeyer 2002; Nicol et al. 2005; Nemergut et al. 2007; Schmidt et al. 2008; Sattin et al. 2009; Schutte et al. 2009, 2010; Philippot et al. 2011; Knelman et al. 2012; Zumsteg et al. 2012).

Overall, the results of this study of the variation in the structure of microbial communities along DCGs and the detailed analysis of the taxa that predominate in the communities suggest that bacteria with heterotrophic metabolisms dominate the microbial communities of the debris cover of the studied glaciers. Therefore, during

the initial stages of soil development in these extreme environments, organic C seems to be mainly provided by allochthonous deposition of organic matter (for example, plant debris, insects, animal excrements) rather than by CO₂ incorporation, as already supposed for glacier forefields (Schmidt et al. 2008; Brankatschk et al. 2011). In addition, the nutrient bioavailability in such extreme environments might be enhanced by the activity of specialized rockweathering bacteria, and chitinolytic bacterial activities might further contribute to enrich supraglacial debris in organic matter.

The study of the biogeographical pattern of *Polaromonas* distribution revealed that the genetic distance among phlotypes was positively correlated with geographic distance. In addition, phlotypes collected at sites <100 km and between 9400 and 13500 km to each other were significantly more similar than those collected at other distance classes. The significant similarity of sequences retrieved in glacial areas separated by <100 km and the significance of the Mantel test, which indicates a genetic by distance divergence of *Polaromonas* strains, suggests that geographic distances contribute to the genetic structuring of *Polaromonas* phlotypes. However, the Mantel correlogram showed that sequences retrieved at glaciers separated by 9400–13 500 km to each other were significantly more similar to each other than sequences found at glaciers in other distance classes. On the one hand, this result is not surprising, as it is similar to the findings of Darcy et al. (2011), from which most of the *Polaromonas* sequences used in this paper were obtained, where a significant similarity between sequences at thousands of kilometres to each other was observed. On the other hand, this result is difficult to explain, as we have no clear hypothesis about why glaciers at this distance class should host *Polaromonas* strains that are significantly more similar to each other than those found on glaciers at other distances. Maybe, *Polaromonas* strains do are ubiquitous, as suggested by Darcy et al. (2011) and in contrast with the previous results. However, if *Polaromonas* strains had a really cosmopolitan distribution, the Mantel correlogram would have identified no significant difference in the genetic distance between sequences at any distance class.

Further investigation on a larger number of *Polaromonas* sequences from more glacial areas of the world are needed to assess whether strains of this bacterial genus do show a biogeographical distribution or are cosmopolitan.

In summary, the results of this study on bacterial communities of DCGs add pieces of evidence supporting the hypotheses that (1) the structure of microbial communities changes according to the time of exposure of the sediment; (2) autotrophic organisms are not the only important early pioneers of recently exposed substrates, but heterotrophic and rock-weathering organisms may play a crucial role in soil formation by exploiting allochthonous deposition of organic matter and enhancing nutrient availability. Moreover, contrasting evidences about the hypothesis on the biogeographical distribution of *Polaromonas* phylotypes were found, and it is still unclear whether strains of this cosmopolitan bacterial genus do show a biogeographical pattern of distribution or they have a worldwide distribution.

6.5 Acknowledgements

We gratefully thank Alessia Sacchetti for help during field work, Rocco Piazza and Alessandra Pirola for their support during sequencing and bioinformatic analyses, Chiara Compostella for chemical analysis of debris samples and Daniel Said Pullicino for helpful discussion on OC sources. Comments from two anonymous referees greatly improved the quality of the manuscript. The work was partially funded by the 2008 MIUR PRIN grant (Grant No. 2008723SYJ_001) and by the 2010-2011 MIUR PRIN grant (Grant No. 2010AYKTAB_006) to CS, and by the 2009 FAR grant to RA.

6.6 References

Ambrosini R, Møller A, Saino N (2009) A quantitative measure of migratory connectivity. *J Theor Biol* 257, 203–211.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 29, 1165–1188.

Benn D, Evans J (2010) *Glaciers and Glaciation*, 2nd edn. Hodder Education: London, UK.

Benson D, Karsch-Mizrachi I, Lipman D, Ostell J, Sayers E (2010) GenBank. *Nucleic Acids Res* 38: D46–D51. Borchard D, Gillet F, Legendre F. (2011). *Numerical Ecology with R*. Springer: New York, NY, USA.

Borin S, Ventura S, Tambone F, Mapelli F, Schubotz F, Brusetti L et al. (2010) Rock weathering creates oases of life in a high Arctic desert. *Environ Microbiol* 12, 293–303.

Brankatschk R, Towe S, Kleineidam K, Schloter M, Zeyer J (2011) Abundances and potential activities of nitrogen cycling microbial communities along a chronosequence of a glacier forefield. *ISME J* 5, 1025–1037.

Brock BW, Mihalcea C, Kirkbride MP, Diolaiuti G, Cutler MEJ, Smiraglia C (2010) Meteorology and surface energy fluxes in the 2005–2007 ablation seasons at the Miage debris-covered glacier, Mont Blanc Massif, Italian Alps. *J Geophys Res* 115, D09106.

Caccianiga M, Andreis C, Diolaiuti G, D'Agata C, Mihalcea C, Smiraglia C (2011) Alpine debris-covered glaciers as a habitat for plant life. *Holocene* 21, 1011–1020.

Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al. (2010) QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 7, 335–336.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N et al. (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J* 6, 1621–1624.

Claesson MJ, O’Sullivan O, Wang Q, Nikkila J, Marchesi JR, Smidt H et al. (2009) Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS One* 4, e6669.

Claesson MJ, Wang QO, O’Sullivan O, Greene-Diniz R, Cole JR, Ross RP et al. (2010) Comparison of two nextgeneration sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic Acids Res* 38, e200.

Darcy JL, Lynch RC, King AJ, Robeson MS, Schmidt SK (2011) Global distribution of *Polaromonas* phylotypes—evidence for a highly successful dispersal capacity. *PLoS One* 6, e23742.

De Caceres M, Legendre P, Moretti M (2010) Improving indicator species analysis by combining groups of sites. *Oikos* 119, 1674–1684.

del Rio TG, Abt B, Spring S, Lapidus A, Nolan M, Tice H et al. (2010) Complete genome sequence of *Chitinophaga pinensis* type strain (UQM 2034T). *Stand Genomic Sci* 2, 87–95.

Diolaiuti G, D’Agata C, Smiraglia C (2003) Belvedere Glacier, Monte Rosa, Italian Alps: tongue thickness and volume variations in the second half of the 20th century. *Arct Antarct Alp Res* 35, 255–263.

Diolaiuti G, Kirkbride MP, Smiraglia C, Benn DI, D'Agata C, Nicholson L (2005) Calving processes and lake evolution at Miage glacier, Mont Blanc, Italian Alps. *Ann Glaciol* 40, 207–214.

Egli M, Favilli F, Krebs R, Pichler B, Dahms D (2012) Soil organic carbon and nitrogen accumulation rates in cold and alpine environments over 1 Ma. *Geoderma* 183–184, 109–123.

Eskelinen A, Stark S, Mannisto M (2009) Links between plant community composition, soil organic matter quality and microbial communities in contrasting tundra habitats. *Oecologia* 161, 113–123.

Fierer N, Bradford M, Jackson R (2007) Toward an ecological classification of soil bacteria. *Ecology* 88, 1354–1364.

Franzetti A, Gandolfi I, Gaspari E, Ambrosini R, Bestetti G (2011) Seasonal variability of bacteria in fine and coarse urban air particulate matter. *Appl Microbiol Biotechnol* 90, 745–753.

Frey B, Rieder S, Brunner I, Plotze M, Koetzsch S, Lapanje A et al. (2010) Weathering-associated bacteria from the Damma glacier forefield: physiological capabilities and impact on granite dissolution. *Appl Environ Microbiol* 76, 4788–4796.

Ganzert L, Lipski A, Hubberten H, Wagner D (2011) The impact of different soil parameters on the community structure of dominant bacteria from nine different soils located on Livingston Island, South Shetland Archipelago, Antarctica. *FEMS Microbiol Ecol* 76, 476–491.

Gobbi M, Isaia M, De Bernardi F (2011) Arthropod colonisation of a debris-covered glacier. *Holocene* 21, 343–349.

Goransson H, Venterink HO, Baath L (2011) Soil bacterial growth and nutrient limitation along a chronosequence from a glacier forefield. *Soil Biol Biochem* 43, 1333–1340.

Huber J, Mark Welch D, Morrison H, Huse S, Neal P, Butterfield D et al. (2007) Microbial population structures in the deep marine biosphere. *Science* 318, 97–100.

Jiang H, Huang Q, Deng S, Dong H, Yu B (2010) Planktonic actinobacterial diversity along a salinity gradient of a river and five lakes on the Tibetan Plateau. *Extremophiles* 14, 367–376.

King AJ, Freeman KR, McCormick KF, Lynch RC, Lozupone C, Knight R et al. (2010) Biogeography and habitat modelling of high-alpine bacteria. *Nat Commun* 1, 3.

Knelman JE, Legg TM, O'Neill SP, Washenberger CL, Gonzalez A, Cleveland CC et al. (2012) Bacterial community structure and function change in association with colonizer plants during early primary succession in a glacier forefield. *Soil Biol Biochem* 46, 172–180.

Lauber C, Hamady M, Knight R, Fierer N (2009) Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl Environ Microbiol* 75, 5111–5120.

Laybourn-Parry J, Tranter M, Hodson AJ (2012) *The Ecology of Snow and Ice Environments*. Oxford University Press: Oxford, UK.

Lazzaro A, Abegg C, Zeyer J (2009) Bacterial community structure of glacier forefields on siliceous and calcareous bedrock. *Eur J Soil Sci* 60, 860–870.

Legendre P, Legendre L (1998) *Numerical Ecology*, 2nd English edn. Elsevier: Amsterdam, The Netherlands.

Lennon J (2011) Replication, lies and lesser-known truths regarding experimental design in environmental microbiology. *Environ Microbiol* 13, 1383–1386.

Madigan MT, Martinko JM, Dunlap PV, Clark DP (2009) *Brock Biology of Microorganisms*, 12th edn. Pearson International Edition: San Francisco, CA, USA.

Mannisto M, Tiirola M, Haggblom M (2007) Bacterial communities in Arctic fields of Finnish Lapland are stable but highly pH-dependent. *FEMS Microbiol Ecol* 59, 452–465.

Mapelli F, Marasco R, Balloi A, Rolli E, Cappitelli F, Daffonchio D et al. (2012) Mineral–microbe interactions: biotechnological potential of bioweathering. *J Biotechnol* 157, 473–481.

Mapelli F, Marasco R, Rizz A, Baldi F, Ventura S, Daffonchio D et al. (2011) Bacterial communities involved in soil formation and plant establishment triggered by pyrite bioweathering on Arctic moraines. *Microbial Ecol* 61, 438–447.

Michaud L, Caruso C, Mangano S, Interdonato F, Bruni V, Lo Giudice A (2012) Predominance of *Flavobacterium*, *Pseudomonas*, and *Polaromonas* within the prokaryotic community of freshwater shallow lakes in the northern Victoria Land, East Antarctica. *FEMS Microbiol Ecol* 82, 391–404.

Mihalcea C, Brock BW, Diolaiuti G, D’Agata C, Citterio M, Kirkbride MP et al. (2008) Using ASTER satellite and ground-based surface temperature measurements to derive

supraglacial debris cover and thickness patterns on Miage Glacier (Mont Blanc Massif, Italy). *Cold Regions Sci Technol* 52, 341–354.

Moorhead DL, Lashermes G, Sinsabaugh RL (2012) A theoretical model of C- and N-acquiring exoenzyme activities, which balances microbial demands during decomposition. *Soil Biol Biochem* 53, 133–141.

Nakawo M, Rana B (1999) Estimate of ablation rate of glacier ice under a supraglacial debris layer. *Geografiska Annaler* 81, 695–701.

Nemergut DR, Anderson SP, Cleveland CC, Martin AP, Miller AE, Seimon A et al. (2007) Microbial community succession in an unvegetated, recently deglaciated soil. *Microbial Ecol* 53, 110–122.

Nicol GW, Tscherko D, Embley TM, Prosser JI (2005) Primary succession of soil Crenarchaeota across a receding glacier foreland. *Environ Microbiol* 7, 337–347.

Oksanen J, Kindt R, Legendre P, O'Hara B, Simpson G, Solymos P et al. (2009) *Vegan: Community Ecology Package*, R package version 1.15-3.

Østrem G (1959) Ice melting under a thin layer of moraine and the existence of ice cores in moraine ridges. *Geograf Ann* 41, 228–230.

Pelfini M, Santilli M, Leonelli G, Bozzoni M (2007) Investigating surface movements of debris-covered Miage glacier, Western Italian Alps, using dendroglaciological analysis. *J Glaciol* 53, 141–152.

Pelfini M, Diolaiuti G, Leonelli G, Bozzoni M, Bressan N, Brioschi D et al. (2012) The influence of glacier surface processes on the short-term evolution of supraglacial tree vegetation: a case study of the Miage Glacier, Italian Alps. *Holocene* 22, 847–856.

Philippot L, Tscherko D, Bru D, Kandeler E (2011) Distribution of high bacterial taxa across the chronosequence of two Alpine glacier forelands. *Microbial Ecol* 61, 303–312.

R Development Core Team (2008) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing: Vienna, Austria.

Roesch L, Fulthorpe R, Riva A, Casella G, Hadwin A, Kent A et al. (2007) Pyrosequencing enumerates and contrasts soil microbial diversity. *ISME J* 1, 283–290.

Sanyika TW, Stafford W, Cowan DA (2012) The soil and plant determinants of community structures of the dominant actinobacteria in Marion Island terrestrial habitats, Sub-Antarctica. *Polar Biol* 35, 1–13.

Sattin SR, Cleveland CC, Hood E, Reed SC, King AJ, Schmidt SK et al. (2009) Functional shifts in unvegetated, perhumid, recently-deglaciated soils do not correlate with shifts in soil bacterial community composition. *J Microbiol* 47, 673–681.

Schloss P, Westcott S, Ryabin T, Hall J, Hartmann M, Hollister E et al. (2009) Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75, 7537–7541.

Schmidt SK, Reed SC, Nemergut DR, Grandy AS, Cleveland CC, Weintraub MN et al. (2008). The earliest stages of ecosystem succession in high-elevation (5000 meters above sea level), recently deglaciated soils. *Proc R Soc Ser B* 275, 2793–2802.

Schumacher BA (2002) Methods of the Determination of Total Organic Carbon (TOC) in Soils and Sediments. NCEA-C-1282. United States Environmental Protection Agency: Las Vegas, NV, USA.

Schutte U, Abdo Z, Bent S, Williams C, Schneider G, Solheim B et al. (2009) Bacterial succession in a glacier foreland of the high Arctic. *ISME J* 3, 1258–1268.

Schutte U, Abdo Z, Foster J, Ravel J, Bunge J, Solheim B et al. (2010) Bacterial diversity in a glacier foreland of the high Arctic. *Mol Ecol* 19, 54–66.

Sigler W, Crivii S, Zeyer J (2002) Bacterial succession in glacial forefield soils characterized by community structure, activity and opportunistic growth dynamics. *Microb Ecol* 44, 306–316.

Sigler WV, Zeyer J (2002) Microbial diversity and activity along the forefields of two receding glaciers. *Microb Ecol* 43, 397–407.

Smiraglia C, Diolaiuti G, Casati D, Kirkbride MP (2000) Recent areal and altimetric variations of Miage Glacier (Monte Bianco massif, Italian Alps). In Nakawo M, Raymond CF, Fountain A (eds) *Debris Covered Glaciers*. IAHS: Wallingford, Seattle, WA, USA, pp 227–233.

Taschner S, Ranzi R(2002) Comparing the opportunities of LANDSAT-TM and ASTER data for monitoring a debris covered glacier in the Italian Alps within the GLIMS project. *Proceedings of the 22nd International Geoscience and Remote Sensing Symposium (IGRASS 2002) Vol. 2: 24–28 June 2002, Toronto, Canada*, pp 1044–1046.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73, 5261–5267.

Wang Y, Qian P (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4, e7401.

Zumsteg A, Luster J, Goransson H, Smittenberg RH, Brunner I, Bernasconi SM et al. (2012) Bacterial, archaeal and fungal succession in the forefield of a receding glacier. *Microb Ecol* 63, 552–564.

CHAPTER 7

Perspectives

7.1 Future perspectives

The use of NGS technologies has significantly improved our knowledge on the processes and dynamics that govern airborne microbial community structure. Independent studies, conducted among different environments (e.g. urban, rural, marine, etc.) individuated common patterns of microbial distribution in the atmosphere. These surveys independently retrieved a conserved group of dominant populations (Brodie et al, 2007; Fierer et al, 2008; Bowers et al, 2011a; 2012; 2013). This fact suggests the existence of a shared microbial background in outdoor near-surface atmosphere even among different environments and geographic locations.

Moreover, several shared sources of bacteria were identified such as soil, seawater, leaf sediments and animal feces (Bowers et al, 2011b; 2013, chapters 3 and 4). These findings, on one hand disclose the possibility of common origins of airborne bacteria, and on the other hand, open the research to the bacterial long-range transport study. Long-range transport of microorganisms in the atmosphere affects the structure of local microbial communities and their ecological role, as well as the biogeography of microorganisms. Particularly, transport of microorganisms in the upper troposphere may have meteorological and climatic implications, which are still far to be fully elucidated (DeLeon-Rodriguez et al, 2013; Morris et al, 2011; Smith et al, 2012). The existence of different origins of air masses, from which different bacteria might be originated, has also been hypothesized, but not definitively proved (Womack et al, 2010).

Furthermore, the presence of a metabolic activity of some airborne bacteria appears to be plausible, although it is not already be definitively demonstrated (see Womack et al. 2010 and references therein). Moreover, the occurrence of ecological interactions among bacteria is still debated. Thus, there is still a significant lack of knowledge of the potential functions of the airborne microbial community and of the processes through which these functions are affected by environmental conditions.

However so far, only one airborne metagenome has been published in a study describing an urban indoor environment in Singapore (Tringe et al. 2008). The

application of metagenomics and metatranscriptomics techniques to air microbiology could supply ecological information on genetic adaptations of microorganisms to the atmospheric environment. Furthermore, it would also disclose whether the extreme conditions of the atmosphere might have selected some bacterial populations rather than others on the basis of their genetic characteristics.

Despite these encouraging results and perspectives, researches in this field of microbiology are still incomplete and heterogeneous. The adoption of shared methods, homogeneous protocols and experimental designs would be helpful for data comparison among different research groups.

7.2 References

Bowers RM, Clements N, Emerson JB, Wiedinmyer C, Hannigan MP, Fierer N (2013) Seasonal variability in bacteria and fungal diversity of near-surface atmosphere. *Environ Sci Technol* DOI: 10.1021/es402970s.

Bowers RM, McCubbin IB, Hallar AG, Fierer N (2012) Seasonal variability in airborne bacterial communities at a high-elevation site. *Atmos Environ* 50, 41-49.

Bowers RM, McLetchie S, Knight R, Fierer N (2011a) Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of the potential source environments. *ISME J* 5, 601-612.

Bowers RM, Sullivan AP, Costello EK, Collett Jr JL, Knight R, Fierer N (2011b) Sources of bacteria in outdoor air across cities in the Midwestern United States. *Appl Environ Microbiol* 77, 6350-6356.

Brodie EL, De Santis TZ, Moberg Parker JP, Zubietta IX, Piceno YM, Andersen GL (2007) Urban aerosols harbor diverse and dynamic bacterial populations. *P Natl Acad Sci USA* 104,299-304.

DeLeon-Rodriguez N, Latham TL, Rodriguez-R LM, Barazesh JM, Anderson BE, Beyersdorf AJ, Ziemba LD, Bergin M, Nenes A, Konstantinidis T (2013) Microbiome of the upper troposphere: species composition and prevalence, effects of tropical storms, and atmospheric implications. *P Natl Acad Sci USA* online first.

Morris CE, Sands DC, Bardin M, Jaenicke R, Vogel B, Leyronas C, Ariya PA, Psenner R (2011) Microbiology and atmospheric processes: research challenges concerning the impact of airborne micro-organisms on the atmosphere and climate. *Biogeosciences* 8, 17-25.

Smith DJ, Jaffe DA, Birmele MN, Griffin DW, Schuerger AC, Hee J, Roberts MS (2012) Free tropospheric transport of microorganisms from Asia to North America. *Microb Ecol* 64, 973-985.

Tringe SG, Zhang T, Liu X, Yu Y, Lee WH, Yap J, Yao F, Suan ST, Ing SK, Haynes M, Rohwer F, Wei CL, Tan P, Bristow J, Rubin EM, Ruan Y (2008) The airborne metagenome in an indoor urban environment. *PLoS ONE* 3, 10.

Womack AM, Bohannan BJM, Green JL (2010) Biodiversity and biogeography of the atmosphere. *Philos Trans R Soc B Biol Sci* 365, 3645–3653.

CHAPTER 8

Appendix

8.1 Supporting information of Chapter 2



Figure S1 - Torre Sarca sampling site.



Figure S2 - Torre Sarca surroundings.

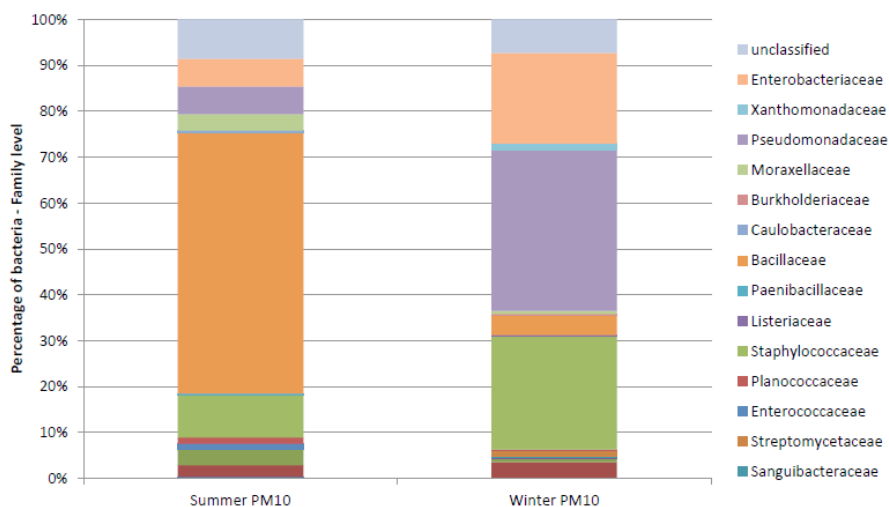


Figure S3 - Histogram chart showing classification at family level of summer and winter isolates.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

8.2 Supporting Information of Chapter 3

8.2.1 Details on statistical methods

8.2.1.1 Missing PM2.5 data

Measures of PM2.5 for the first two days of the winter session of sampling were unavailable. Since complete datasets are required in most multivariate statistical analyses, we imputed these missing data by regressing PM2.5 concentrations on mean PM10 concentration. We used a second order polynomial regression ($R^2 = 0.920$) as this model fitted the data significantly better than a linear regression ($F_{1,35} = 7.43$, $P = 0.001$). Use of higher order polynomials did not significantly improve the fit of the model (details not shown).

8.2.1.2 ANOSIM

ANOSIM is a nonparametric randomization-based method of multivariate analysis widely used to compare the variation in species abundance and composition among groups of sampling units (Kuramae et al, 2011). ANOSIM provides an R statistics, which can range from -1 to 1 (albeit values < 0 are rarely obtained). R-values > 0.75 indicate well separated microbial compositions; R > 0.5 overlapping, but clearly different communities; and R < 0.25 practically not separable communities (Clarke and Gorley, 2006). The significance of ANOSIM tests was assessed using a randomization procedure whereby R was recomputed 9999 times.

8.2.1.3 Short-term variation in bacterial community composition

If the structure of bacterial communities slightly modifies over a few days, then communities sampled in consecutive days, or a few days apart, should be more similar than those sampled several days after. A positive association between distance between communities and the number of days between samplings is therefore expected in a Mantel test. This justified the use of one-tailed tests. Sources of PM in urban areas include dust from road and industrial activities. Inputs from these sources may show cyclical patterns of variation at the time scale of a week, according e.g. to the reduced traffic during weekends. This cyclical variation in inputs may traduce to similarity in bacterial communities sampled at particular time-lags. Mantel correlograms on samples taken within each season was used to investigate this hypothesis (Legendre and Legendre, 1998).

Airborne microbial communities may show large variation in their structure within a few days (Fierer et al, 2008). To compare variation in community structure among samples taken in the same season (within-season variation) and among samples taken in different seasons (between-season variation), we used goodness-of-fit Mantel tests (*sensu* Legendre and Legendre, 1998) whereby model matrix indicated whether samples were taken in the same or in different seasons. Goodness-of-fit Mantel tests were also used to compare variability in community structure within each season to

that observed in all the other samples. All Mantel tests were based on Spearman correlation to better compare the results with ANOSIM tests (Legendre and Legendre, 1998).

8.2.2 References

Clarke KR, Gorley RN. (2006) PRIMER v6: User Manual/Tutorial. PRIMER-E, Plymouth.

Fierer N, Liu Z, Rodriguez-Hernandez M, Knight R, Henn M, Hernandez MT. (2008). Short-term temporal variability in airborne bacterial and fungal populations. *Appl Environ Microb* 74: 200-207.

Kuramae E, Gamper H, van Veer J, Kovalchuk G (2011) Soil and plant factors driving the soil-borne microorganisms across chronosequences of secondary succession of chalk grassland with a neutral pH. *FEMS Microbiol Ecol* **77**: 285-294.

Legendre P, Legendre L. (1998). *Numerical ecology* (2nd english ed.). - Elsevier. Amsterdam, The Netherlands.

8.2.3 Tables

Table S1. 6bp-barcodes for Illumina sequencing and relative samples.

| Barcode | Sample ID |
|--------------|-------------------------------|
| 5'-AGCAGC-3' | 1 Wi/2 Wi/3 Wi/4 Wi/6 Wi |
| 5'-AGAGAG-3' | 5 Wi/7 Wi/8 Wi/9 Wi/10 Wi |
| 5'-AGATGC-3' | 11 Sp/12 Sp/13 Sp/14 Sp/15 Sp |
| 5'-ATGATG-3' | 16 Sp/17 Sp/18 Sp/19 Sp/20 Sp |
| 5'-CAGAGC-3' | 21 Su/22 Su/23 Su/24 Su/25 Su |
| 5'-CAGCAG-3' | 26 Su/27 Su/28 Su/29 Su/30 Su |
| 5'-CATGAG-3' | 31 Au/32 Au/33 Au/34 Au/35 Au |
| 5'-CTGCTG-3' | 36 Au/37 Au/38 Au/39 Au/40 Au |

Figure S2. Heat-plot of the abundances (blue = low abundance, red = high abundance) of the most abundant OTUs detected in the 40 samples (numbers on the right). The dendrogram represents the hierarchical clustering analyses. A and B are the two main clusters identified. Colours at the base of the dendrogram represent the season when the sample was taken (blue = winter, green = spring, red = summer, orange = autumn).

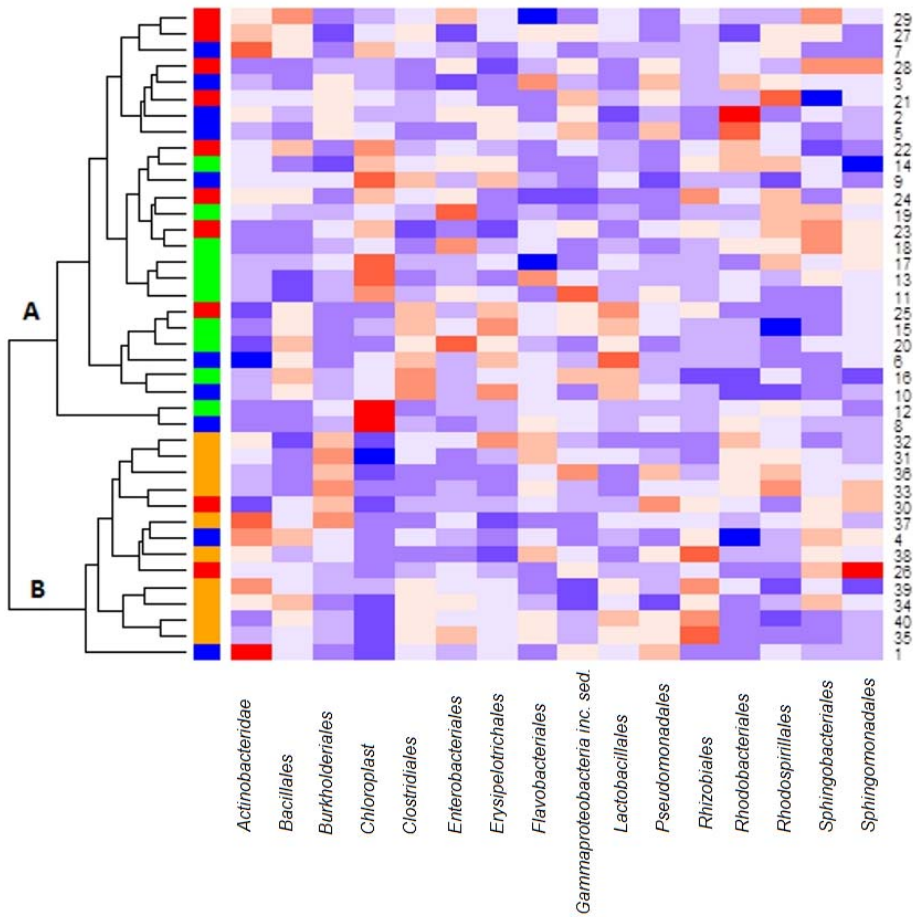


Figure S3. Two dimensional PCA plot of microbial communities sampled in different seasons (blue squares: winter; green circles: spring; red triangles: summer; orange diamonds: autumn) based on OTUs obtained by RDP classifier. Only the most abundant OTUs were considered. The percentage of variance explained by each axis is shown.

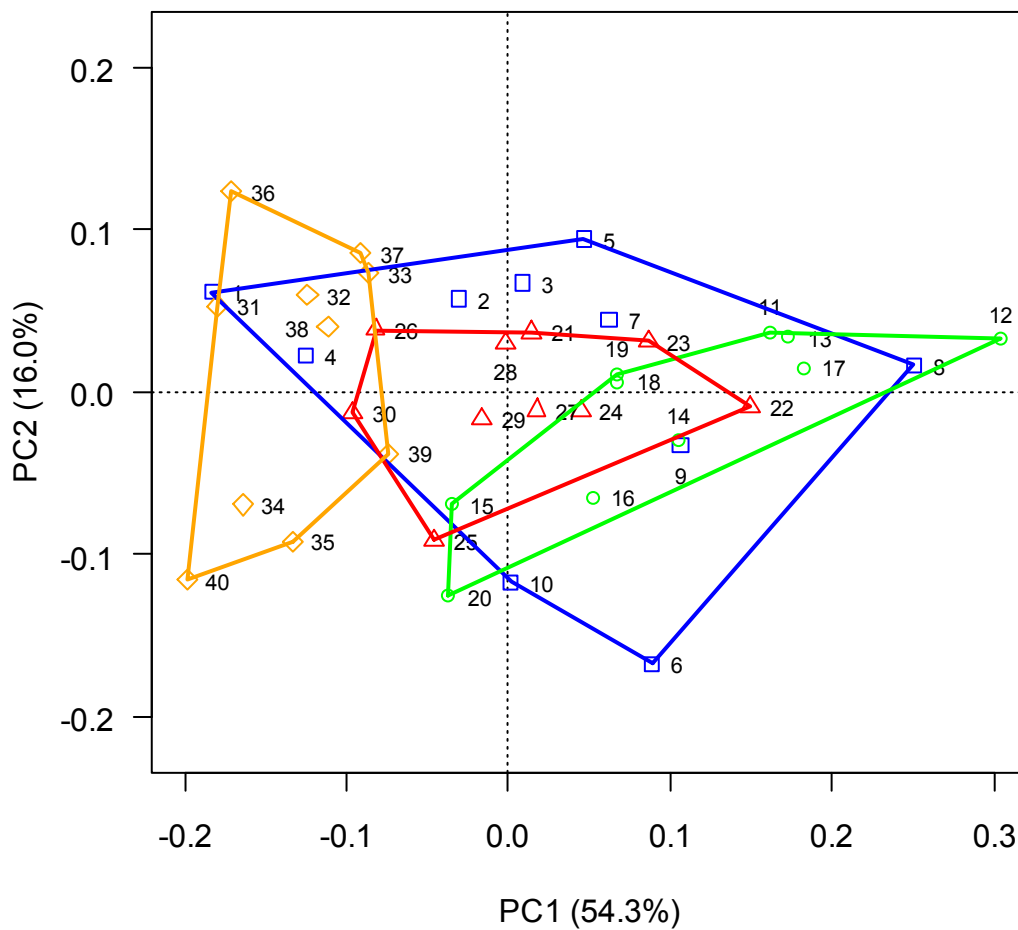
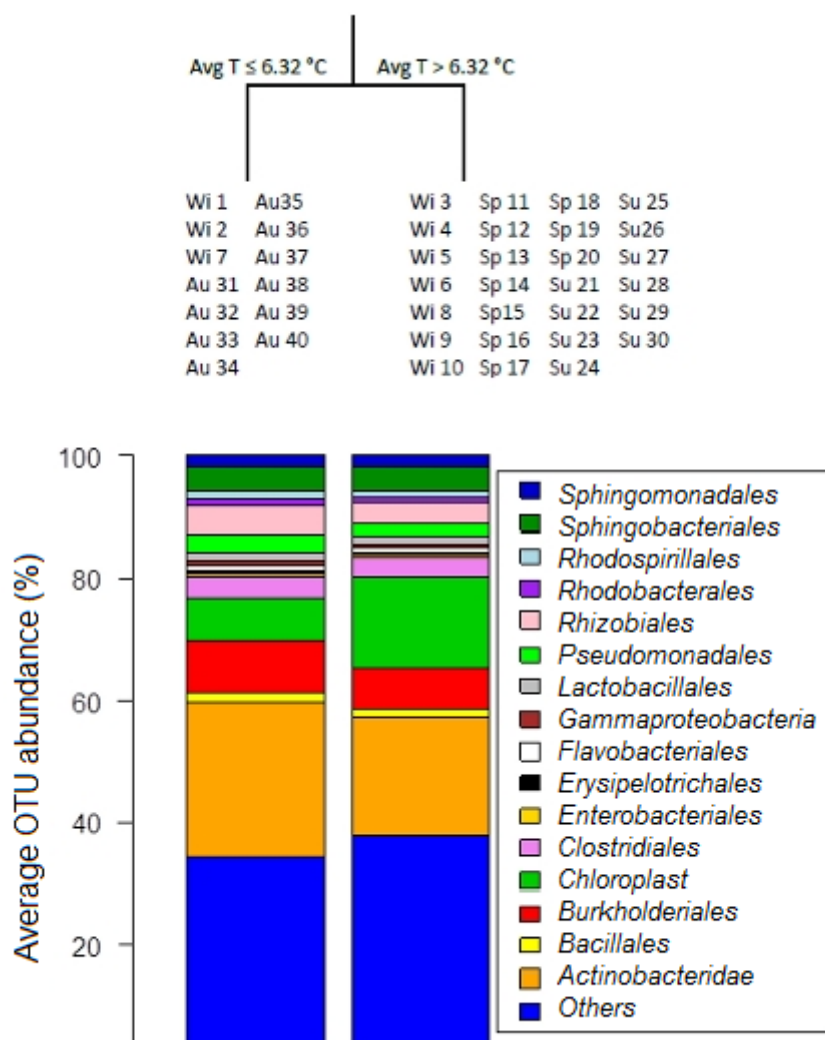


Figure S4. MRT analysis of OTU abundance and air and PM factors. Bar-plots show the average abundance of all the common OTUs (all the other OTUs are clumped in the “Other” category) in samples collected in days with average temperature lower (left bar) or higher (right bar) than 6.32°C.



8.3 Supporting Information of Chapter 4

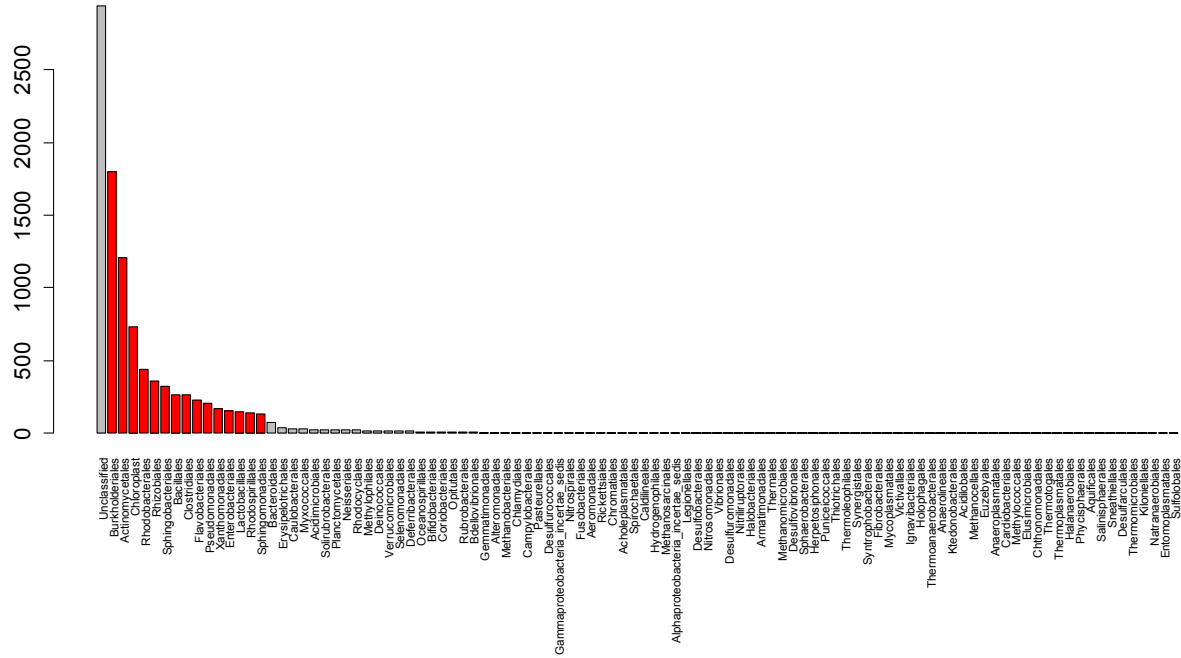


Figure S1 Bar plot of OUT's mean relative abundance. Abundant OTUs are highlighted in red.

8.4 Supporting Information of chapter 6

8.4.1 Further information on glaciological and environmental features of the miage and belvedere glaciers

8.4.1.1 Miage glacier

Miage Glacier (11 km²) is located in the Mont Blanc Massif; Valle d'Aosta, Italy. The ablation zone of this glacier (45° 47' N, 06° 52' E) is debris-covered from about 2400 m above sea level (a.s.l.) downwards to the terminus at 1720 m a.s.l. The lower zone of Miage glacier is covered by a continuous thick debris layer (Mihalcea et al. 2008), mainly composed by igneous and metamorphic rocks (gneisses, schistes and granites; Brock et al. 2010). The rock debris coverage ranges in thickness between a few centimetres of sparse debris up to more than 1 m at the terminus. The glacier tongue is divided into three lobes (North, Central and South lobe). Trees (mainly European Larch *Larix decidua* Mill. up to 2 m high) grow on the glacier debris cover up to 2000 m a.s.l. (Pelfini et al. 2007; 2012). The debris is also colonized by herbaceous vegetation (Caccianiga et al. 2011), arthropods (Gobbi et al. 2011) and nematodes (A. Zullini, personal communication). Mammals (e.g. Chamois *Rupicapra rupicapra*) can be observed on the glacier (C. Mihalcea, personal observations).

Ground-measured temperatures of debris cover during a sunny summer day exceed 30 °C, and are about 1-4 °C during night (Brock et al. 2010). Daily temperature excursion is therefore 28-33 °C. Close to the terminus debris temperature is positive throughout the vertical profile during almost the entire ablation season, and a continuous ice melt occurs at the bottom.

ASTER kinetic surface temperatures (12 August 2009 10:40 am UTC +1:00) are similar to ground measured temperatures: mean surface temperature over areas 90 × 90 m wide varies between 25 °C at 2059 m (site 9M) and 32 °C close to the terminus (site 1M). The Pearson's correlation coefficient between ASTER and the field surface temperature data is 0.94 over the analyzed area.

Debris thickness was estimated from ASTER data according to empirical models (Taschner and Ranzi, 2002; Mihalcea et al. 2008). Previous researches indeed

demonstrated that surface temperatures derived from remote sensing data are good proxies for debris thicknesses (Taschner and Ranzi, 2002; Mihalcea et al. 2008). However, calculated debris thickness and ASTER temperatures are not collinear ($r = 0.146$) and can therefore be entered simultaneously in statistical analyses (see Statistical analyses). The calculated debris thickness varies between 18 cm up to 55 cm (average value over 90×90 m areas). Overall, debris thickness increases downwards, but is thinner in crevassed areas. The lowest thickness was measured at site 5M located at about 1956 m a.s.l. in a crevassed area where the glacier divides into three lobes. Thickest debris corresponds to areas close to the terminus (site 1M and 2M). On the south lobe the debris layer is generally more than 30 cm thick and stratified. Thick and stable debris cover presents approximately 3-4 layers: a fine layer at the bottom with melting water along the first centimetres followed by a mixed layer of fine and coarse debris, a layer of coarse debris with clasts of 1-10 cm, and a final layer at the surface with clasts larger than 10 cm.

Generally, over the analyzed area the slope is between 9 and 34 degrees. The annual glacier velocity along the longitudinal profile, measured by the Differential Global Positioning System method (Diolaiuti et al. 2005; Caccianiga et al. 2011) in the area where samples were collected, decreases downwards, from 23 m y^{-1} in the upper parts to 2 m y^{-1} at the terminus on the South lobe. South lobe is generally more stable, presents a thick layer of debris and, probably, favourable conditions for microorganisms and vegetation development (Pelfini et al. 2007; Caccianiga et al. 2011).

The ablation rate is influenced by debris distribution with rates that varies from 0.8 cm d^{-1} with more than 55 cm of debris to 3 cm d^{-1} where the debris is less than 20 cm thick. These values were calculated over time intervals of approximately 100 days during the 2010 ablation period. The ablation rate generally decreases from the top site (9M) to the lowest one (1M) with higher values during June-July and lower values during August-September.

8.4.1.2 Belvedere glacier

Belvedere debris-covered glacier (45° 57' N, 4° 34' E) descends the steep eastern slope of Monte Rosa, Piemonte, Italy. The glacier snout formed two main lobes ending at about 1750 m a.s.l. The glacier tongue is debris covered from about 2300 m a.s.l. to the terminus (Diolaiuti et al. 2003). The debris layer is mainly composed by gneiss, micaschists and granites. Trees (European Larch and Willows *Salix* ssp. up to 1 m high) grow on the glacier debris cover up to 2000 m a.s.l. (M. Pelfini, personal observation). Herbaceous vegetation, arthropods, (Gobbi et al. 2011) and nematodes (A. Zullini, personal communication) colonize the debris cover of the Belvedere glacier, and Chamois can be observed on the glacier (R. Ambrosini personal observations).

The slope is generally lower than that of Miage glacier with maximum values of 16 degrees.

Surface temperatures are available from an ASTER image (5 august 2009 10:40 UTC +1:00) and are generally lower than those on the Miage glacier. The maximum temperature is 29 °C at 1959 m a.s.l. (site 4B). Generally, the surface temperature is 23 °C close to the terminus and decreases to 21 °C above 2000 m with some exceptions: 25 °C at site 3B and 26 °C at site 9B. Remote sensing data did not disclose a clear pattern of variation of surface temperature with elevation as that observed on Miage glacier.

Debris thickness, calculated from ASTER images, ranges from 23 cm at the highest sampling site (9B) to 41 cm (site 1B) at the glacier snout, but shows no clear trend over the sampling area. This is probably due to a surging event that occurred in 2003 in this glacier and to rock avalanches from the steep East face of Monte Rosa, which altered the distribution of the debris at the glacier surface.

The ablation rate is higher than on Miage glacier and varies from 6 cm d⁻¹ (June-July) at 1920-1970 m a.s.l. (sites 3B-5B) to 2 cm d⁻¹ (June-July) at 2080 m (site 9B). At lower elevation (1850 m), the ablation rate is about 5 cm d⁻¹. Melt rate is probably increased by the presence of crevassed areas between 1920 and 1970 m a.s.l. At

lower elevation the surface is more regular with continuous debris cover and decreasing melt rates.

Annual surface velocity, measured by the Differential Global Positioning System method (Diolaiuti et al. 2005; Caccianiga et al. 2011), is highest in the upper zone at 2000 m a.s.l. with values of 40-45 m y⁻¹ and decreases down valley to 30 m y⁻¹ at sites 7B and 6B, 15 m y⁻¹ at site 5B and 4B. The lowest velocity of 5 m y⁻¹ was observed at site 2B (1878 m a.s.l.) near the terminus.

8.4.1.3 Comparison of ecological conditions at sampling sites on Miage and Belvedere glaciers

Surface temperatures obtained by ASTER images were on average higher on the Miage (28.22 °C ± 0.86 SE) than on the Belvedere glacier (23.63 °C ± 0.93 SE; t-test: $t_{17} = -3.595$, $P = 0.002$). Slope at sampling sites was higher on the Miage (21.11 degrees ± 2.93 SE) than on the Belvedere glacier (9.40 degrees ± 1.69 SE; t-test for inequality of variances: $t_{12.97} = -3.463$, $P = 0.004$). pH of sampled debris was more acidic on the Belvedere (5.53 ± 0.13 SE) than on the Miage glacier (6.98 ± 0.14 SE; t-test: $t_{36} = -7.793$, $P < 0.001$). Annual ablation rate, altitude of sampling point, annual surface velocity, debris thickness, distance from the glacier terminus and organic carbon (OC) content did not differ significantly between glaciers ($|t_{17}| \leq 1.518$, $P \geq 0.147$ in all cases).

8.4.2 Further information on sampling methods

At each sampling site approximately 50 ml of debris were aseptically collected by a garden scoop sterilized by denatured alcohol and fire. Scoop sterilization was repeated before taking each sample. Large clasts were discarded. Debris was put in sterilized falcon tubes and temporarily stored in a portable ice chest at approximately +4 °C. Within 12 h from collection, samples were brought to the lab where they were stored in a refrigerator at -20 °C until molecular analyses, and at +4 °C for determination of OC content and pH.

8.4.3 Further information on molecular analyses

Total DNA was extracted once for each sample according to the manufacturer's instructions for the product, which assure an optimal extraction yield. However, even in the case DNA extraction would be not complete, this extraction bias can be assumed to be at the same extent for all samples, which still allows comparison among samples.

The V5 and V6 hypervariable regions of the 16S rRNA gene were PCR-amplified for Genome Analyzer-IIx sequencing. The PCR was performed in 3 reactions with a volume of 75 μ L for each sample with @Taq[®] Hot start Taq polymerase (Euroclone, Pero, Italy), 4 mM MgCl₂, 0.2 mM dNTPs mix, 1 μ M of each primer. 783F and 1046R primers were used (Huber et al. 2007, Wang and Qian, 2009) and the cycling conditions were: initial denaturation at 94 °C for 5 min; 29 cycles at 94 °C for 50 s, 47 °C for 30 s, and 72 °C for 30 s and final extension at 72 °C for 5 min. At the 5' end of the 783F primer, one of eight 6-bp barcodes was also included to allow sample pooling and subsequent sequence sorting (Table S2). After amplification, the 3 PCR reactions per each sample were pooled together before further processing. The amplified products of 288 bp were purified with the Wizard SV PCR purification kit (Promega Corporation, Madison, WI, USA) and DNA quantity and purity was evaluated spectrophotometrically by NanoDrop™ (Thermo Scientific, Waltham, MA, USA). Multiplexed sequencing of all the pooled samples were performed on a single Illumina GA-IIx lane, using a paired-end 76 base-pair protocol and the 4.0 sequencing chemistry. Illumina Real Time Analysis software, version 1.8.7, was used to perform the cluster extraction and base-calling processing analyses. The Illumina Multiplexing Sample Preparation Oligonucleotide Kit provides 12 index oligos for pooling up to 12 samples per lane. Therefore, the further customised tagging system with 8 barcodes at the 5' end of the 783F primer allowed pooling up to 96 samples per lane. Illumina

GA-IIx sequencing was carried out at the Department of Clinical and Preventive Medicine, University of Milano-Bicocca, Italy.

8.4.4 Further details on statistical analyses

Preliminary analyses on independent variables entered in redundancy analyses

We preliminarily investigated the correlations among environmental variables in order to avoid entering highly correlated predictors in the statistical models. To this end we calculated all pairwise correlations among all variables that could be entered in the analyses as predictors, namely ablation rate, altitude of sampling point, amount of OC in the sample, ASTER kinetic surface temperature, annual surface velocity, debris thickness, distance from the glacier terminus, pH and slope. This analysis showed that altitude of sampling point, distance from the terminus, and annual surface velocity were highly correlated ($r \geq 0.85$), as well as temperature and pH ($r = 0.64$). Hence, altitude, surface velocity and temperature were disregarded from the analyses.

Use of arbitrary cutoffs in the Mantel correlogram over alternative procedures and robustness of the results

Mantel correlogram analysis needs a priori identification of distance classes. This task is often performed by the Sturges' rule, which identified the number of distance classes according to the formula $N = 1 + 3.322 \times \log_{10}(m)$.

Where N is the number of distance classes identified by the Sturges' rule and m is the number of pairwise distances in the correlogram. Cutoffs between distance classes are then chosen so to identify N classes of equal width (De Caceres et al. 2010, Legendre and Legendre 1998). If sampling sites are not equally spaced, distance classes chosen according to the Sturges' rule contain different number of pairwise differences, and therefore statistical power of test differ across distance classes.

An alternative procedure would be to create distance classes of unequal widths, but containing exactly the same number of pairwise differences. This procedure would assure the same statistical power across all distance classes, but the choice of cutoffs

is more difficult (De Caceres et al. 2010, Legendre and Legendre 1998). In addition, the construction of distance classes containing an equal number of comparisons is prevented when ties occur in distances among sampling units, as was the case in the dataset used in the present analysis, where sequences found on the same glaciers were compared.

We therefore preferred to set cutoffs between distance classes to values corresponding to gaps in the distribution of distances among glaciers (see Figure 4 in the main text).

As detailed in the Results of the main text, we found that sequences retrieved on glaciers less than 100 km apart (first distance class in Figure 4 of the main text) were more similar than those found on glaciers at larger distances. Similarly, sequences on glaciers between 9400 to 13500 km (fifth distance class in Figure 4 of the main text) to each other were more similar to one another than those found on glaciers at other distances. We investigated the robustness of these results by re-running the correlogram analysis while shifting the cutoff limits of these distance classes. In particular we shifted 1) the upper limit of the first distance class up to 1200 km; 2) the lower limit of the fifth distance class between 8000 and 11000 km; 3) the upper limit of the fifth distance class between 12000 and 15000 km.

In all cases results did not change qualitatively (details not shown), and we are therefore confident that the arbitrary choice of cutoffs has not altered the general conclusions that we draw from the Mantel correlogram analysis.

8.4.5 References

Brock BW, Mihalcea C, Kirkbride MP, Diolaiuti G, Cutler MEJ, Smiraglia C. (2010). Meteorology and surface energy fluxes in the 2005-2007 ablation seasons at the Miage debris-covered glacier, Mont Blanc Massif, Italian Alps. *J Geophys Res - Atmos* 115:D09106.

Caccianiga M, Andreis C, Diolaiuti G, D'Agata C, Mihalcea C, Smiraglia C. (2011). Alpine debris-covered glaciers as a habitat for plant life. *Holocene* 21:1011-1020.

De Caceres M, Legendre P, Moretti M. (2010). Improving indicator species analysis by combining groups of sites. *Oikos* 119:1674-1684.

Diolaiuti G, D'Agata C, Smiraglia C (2003). Belvedere Glacier, Monte Rosa, Italian Alps: Tongue thickness and volume variations in the second half of the 20th century. *Arct Antarct Alp Res* 35:255-263.

Diolaiuti G, Kirkbride MP, Smiraglia C, Benn DI, D'Agata C, Nicholson L. (2005). Calving processes and lake evolution at Miage glacier, Mont Blanc, Italian Alps. *Ann Glaciology*, 40:207-214.

Gobbi M, Isaia M, De Bernardi F. (2011). Arthropod colonisation of a debris-covered glacier. *Holocene* 21:343-349.

Huber J, Mark Welch D, Morrison H, Huse S, Neal P, Butterfield D et al. (2007). Microbial population structures in the deep marine biosphere. *Science* 318:97-100.

Legendre P, Legendre L. (1998). *Numerical Ecology*, 2nd English edn. Elsevier:Amsterdam.

Mihalcea C, Brock BW, Diolaiuti G, D'Agata C, Citterio M, Kirkbride MP et al. (2008). Using ASTER satellite and ground-based surface temperature measurements to derive supraglacial debris cover and thickness patterns on Miage Glacier (Mont Blanc Massif, Italy). *Cold Regions Sci Technol* 52:341-354.

Pelfini M, Diolaiuti G, Leonelli G, Bozzoni M, Bressan N, Brioschi D, et al. (2012). The influence of glacier surface processes on the short-term evolution of supraglacial tree vegetation: A case study of the Miage Glacier, Italian Alps. *Holocene*. 22:847-856.

Pelfini M, Santilli M, Leonelli G, Bozzoni M. (2007) Investigating surface movements of debris-covered Miage glacier, Western Italian Alps, using dendroglaciological analysis. *J Glaciol* 53:141-152.

Taschner S, Ranzi R. (2002). Comparing the opportunities of LANDSAT-TM and ASTER data for monitoring a debris covered glacier in the italian alps within the GLIMS project In Proceedings of the 22nd International Geoscience and Remote Sensing Symposium (IGRASS 2002), 24-28 June 2002, Toronto Canada. Vol. 2:1044-1046.

Wang Y, Qian P. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS One* 4:e7401.

8.4.6 Tables

Table S1 - UTM coordinates and elevation of sampling sites.

| Glacier | Month | Sample ID | UTM y | UTM x | h |
|---------|-------|-----------|---------|--------|------|
| Miage | July | 1MJ | 5072431 | 336380 | 1780 |
| Miage | July | 2MJ | 5072383 | 336153 | 1844 |
| Miage | July | 3MJ | 5072265 | 335947 | 1888 |
| Miage | July | 4MJ | 5072109 | 335798 | 1926 |
| Miage | July | 5MJ | 5072027 | 335587 | 1956 |
| Miage | July | 6MJ | 5072091 | 335365 | 1993 |
| Miage | July | 7MJ | 5072092 | 335022 | 2038 |
| Miage | July | 8MJ | 5072088 | 334846 | 2049 |
| Miage | July | 9MJ | 5072094 | 334650 | 2059 |

Chapter 8

| | | | | | |
|-----------|-----------|------|---------|--------|------|
| Miage | September | 1MS | 5072431 | 336376 | 1780 |
| Miage | September | 2MS | 5072389 | 336145 | 1842 |
| Miage | September | 3MS | 5072261 | 335946 | 1888 |
| Miage | September | 4MS | 5072264 | 335803 | 1908 |
| Miage | September | 5MS | 5072026 | 335585 | 1956 |
| Miage | September | 6MS | 5072091 | 335363 | 1994 |
| Miage | September | 7MS | 5072093 | 335021 | 2038 |
| Miage | September | 8MS | 5072088 | 334846 | 2049 |
| Miage | September | 9MS | 5072095 | 334649 | 2059 |
| Belvedere | July | 1BJ | 5091144 | 416670 | 1803 |
| Belvedere | July | 2BJ | 5091117 | 416466 | 1878 |
| Belvedere | July | 3BJ | 5090937 | 416291 | 1926 |
| Belvedere | July | 4BJ | 5090737 | 416159 | 1959 |
| Belvedere | July | 5BJ | 5090588 | 416102 | 1971 |
| Belvedere | July | 6BJ | 5090416 | 415976 | 2002 |
| Belvedere | July | 7BJ | 5090253 | 415815 | 2017 |
| Belvedere | July | 8BJ | 5090094 | 415650 | 2033 |
| Belvedere | July | 9BJ | 5089636 | 415811 | 2085 |
| Belvedere | July | 10BJ | 5089685 | 415638 | 2079 |
| Belvedere | September | 1BS | 5091144 | 416670 | 1803 |
| Belvedere | September | 2BS | 5091117 | 416466 | 1878 |
| Belvedere | September | 3BS | 5090937 | 416291 | 1926 |

Chapter 8

| | | | | | |
|-----------|-----------|------|---------|--------|------|
| Belvedere | September | 4BS | 5090737 | 416159 | 1959 |
| Belvedere | September | 5BJ | 5090588 | 416102 | 1971 |
| Belvedere | September | 6BJ | 5090416 | 415976 | 2002 |
| Belvedere | September | 7BJ | 5090253 | 415815 | 2017 |
| Belvedere | September | 8BS | 5090094 | 415650 | 2033 |
| Belvedere | September | 9BS | 5089636 | 415811 | 2085 |
| Belvedere | September | 10BS | 5089685 | 415638 | 2079 |

Table S2 - 6 bp-barcodes for Illumina sequencing and related samples.

| Barcode ID | Sequence 5'-3' | Sample IDs |
|------------|----------------|--------------------------|
| 1B | AGCAGC | 3MJ;6MJ;9MJ;12MJ;15MJ |
| 3B | AGAGAC | 18MJ;21MJ;24MJ;6MS;10MS |
| 4B | AGATGC | 27MJ;14MS;18MS;1BJ;4BJ |
| 9B | ATGATG | 22MS;26MS;10BJ;13BJ;37BS |
| 12B | CAGAGC | 2MS;30MS;36MS;16BJ;19BJ |
| 14B | CAGCAG | 22BJ;25BJ;1BS;5BS;9BS |
| 17B | CATGAG | 28BJ;13BS;17BS;29BS;33BS |
| 20B | CTGCTG | 7BJ;21BS;25BS |

Table S3 – Number of filtered sequences and of Operational Taxonomic Units in each sample.

Chapter 8

| Sample ID | n° filtered sequences | % unclassified sequences | n° classified taxa (OTUs) |
|------------------|------------------------------|---------------------------------|----------------------------------|
| 1 MJ | 37176 | 47 | 62 |
| 2 MJ | 28925 | 44 | 63 |
| 3 MJ | 10391 | 47 | 54 |
| 4 MJ | 25127 | 41 | 61 |
| 5 MJ | 17668 | 44 | 54 |
| 6 MJ | 24484 | 47 | 64 |
| 7 MJ | 17974 | 45 | 57 |
| 8 MJ | 6047 | 45 | 50 |
| 9 MJ | 30496 | 43 | 63 |
| 1 BS | 7480 | 40 | 54 |
| 2 BS | 20723 | 46 | 56 |
| 3 BS | 13758 | 49 | 56 |
| 4 BS | 32072 | 47 | 63 |
| 5 BS | 8676 | 48 | 57 |
| 6 BS | 24543 | 36 | 52 |
| 7 BS | 6452 | 46 | 46 |
| 8 BS | 23169 | 50 | 63 |
| 9 BS | 15749 | 50 | 56 |
| 10 BS | 39852 | 51 | 68 |
| 1 BJ | 24085 | 41 | 61 |

Chapter 8

| | | | |
|--------------|-------|----|----|
| 2 BJ | 7962 | 48 | 52 |
| 3 BJ | 36436 | 49 | 61 |
| 4 BJ | 27611 | 41 | 63 |
| 5 BJ | 8189 | 40 | 49 |
| 6 BJ | 33019 | 52 | 63 |
| 7 BJ | 25998 | 51 | 61 |
| 8 BJ | 30973 | 48 | 63 |
| 9 BJ | 21402 | 43 | 60 |
| 10 BJ | 42864 | 39 | 63 |
| 1 MS | 7308 | 50 | 51 |
| 2 MS | 27811 | 37 | 57 |
| 3 MS | 19235 | 44 | 62 |
| 4 MS | 22297 | 41 | 56 |
| 5 MS | 14971 | 41 | 54 |
| 6 MS | 19985 | 39 | 56 |
| 7 MS | 12634 | 44 | 59 |
| 8 MS | 34892 | 34 | 61 |
| 9 MS | 21222 | 41 | 54 |

Chapter 8

Table S4 - Distance in kilometres among sampling sites of *Polaromonas* sequences in figure S3. Distance smaller than 100 km are boldfaced, distance between 9400 and 13500 km are underlined and in italic.

| | Belvedere glacier | Miage glacier | Toklat glacier | Mendenhall glacier | Arikaree glacier | Cordillera Vilcanota | John Evans glacier 1 | John Evans glacier 2 | Schneeferner glacier | Tain-Shen Glaciers 1 & 2 | Zuntal glacial valley | Puruogangri ice field | Fox glacier | Franz Josef glacier | Kamb ice stream |
|--------------------------|-------------------|---------------|----------------|--------------------|------------------|----------------------|----------------------|----------------------|----------------------|--------------------------|-----------------------|-----------------------|-------------|---------------------|-----------------|
| Miage glacier | 82 | | | | | | | | | | | | | | |
| Toklat glacier | 7700 | 7704 | | | | | | | | | | | | | |
| Mendenhall glacier | 7906 | 7897 | 992 | | | | | | | | | | | | |
| Arikaree glacier | 8398 | 8353 | 3872 | 2887 | | | | | | | | | | | |
| Cordillera Vilcanota | <u>10276</u> | <u>10194</u> | <u>10829</u> | <u>9854</u> | 6970 | | | | | | | | | | |
| John Evans glacier 1 | 4843 | 4841 | 2886 | 3094 | 4604 | <u>10390</u> | | | | | | | | | |
| John Evans glacier 2 | 4851 | 4849 | 2879 | 3085 | 4597 | <u>10387</u> | 8 | | | | | | | | |
| Schneeferner glacier | 284 | 362 | 7579 | 7829 | 8452 | <u>10546</u> | 4747 | 4755 | | | | | | | |
| Tain-Shen Glaciers 1 & 2 | 5996 | 6076 | 7158 | 8091 | <u>10678</u> | 16127 | 6304 | 6310 | 5715 | | | | | | |
| Zuntal glacial val. | 6724 | 6806 | 8739 | <u>9689</u> | <u>12304</u> | 16947 | 7886 | 7892 | 6464 | 1626 | | | | | |

Chapter 8

| | | | | | | | | | | | | | | | |
|-----------------------|-------|-------|--------------|--------------|--------------|--------------|-------|-------|-------|--------------|--------------|--------------|------|------|------|
| Puruogangri ice field | 6754 | 6836 | 8017 | 8977 | <u>11653</u> | 17018 | 7344 | 7350 | 6481 | 1048 | 760 | | | | |
| Fox glacier | 18588 | 18669 | <u>12398</u> | <u>12426</u> | <u>12540</u> | <u>11131</u> | 15259 | 15253 | 18338 | <u>12691</u> | <u>11875</u> | <u>11879</u> | | | |
| Franz Josef glacier | 18592 | 18674 | <u>12393</u> | <u>12420</u> | <u>12534</u> | <u>11127</u> | 15255 | 15248 | 18342 | <u>12693</u> | <u>11879</u> | <u>11883</u> | 7 | | |
| Kamb ice stream | 15717 | 15709 | 16609 | 16275 | 14701 | 9177 | 19293 | 19285 | 15842 | 14315 | <u>12762</u> | <u>13268</u> | 4180 | 4405 | |
| Collins Glacier | 13988 | 13994 | 18195 | 18127 | 16307 | <u>9885</u> | 18808 | 18816 | 14072 | <u>13028</u> | <u>11406</u> | <u>12028</u> | 5816 | 5822 | 1882 |

8.4.7 Figures

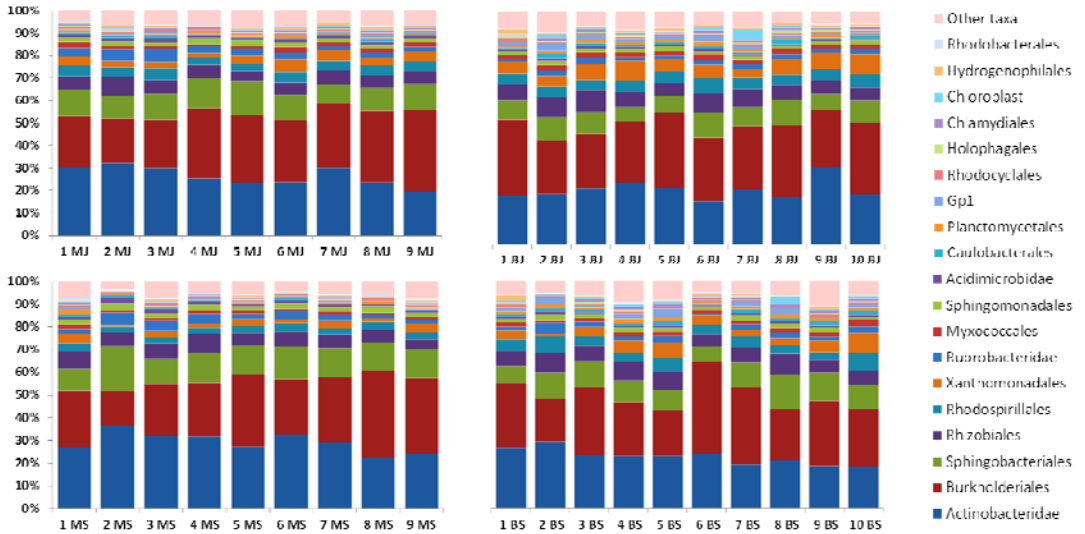


Figure S1 - Taxonomic classification using an RDP Bayesian classifier (50% confidence) at the fourth taxonomic level. a) Miage July; b) Belvedere July; c) Miage September; d) Belvedere September. Unclassified sequences ranged from 33.9% to 51.7% and were not shown.

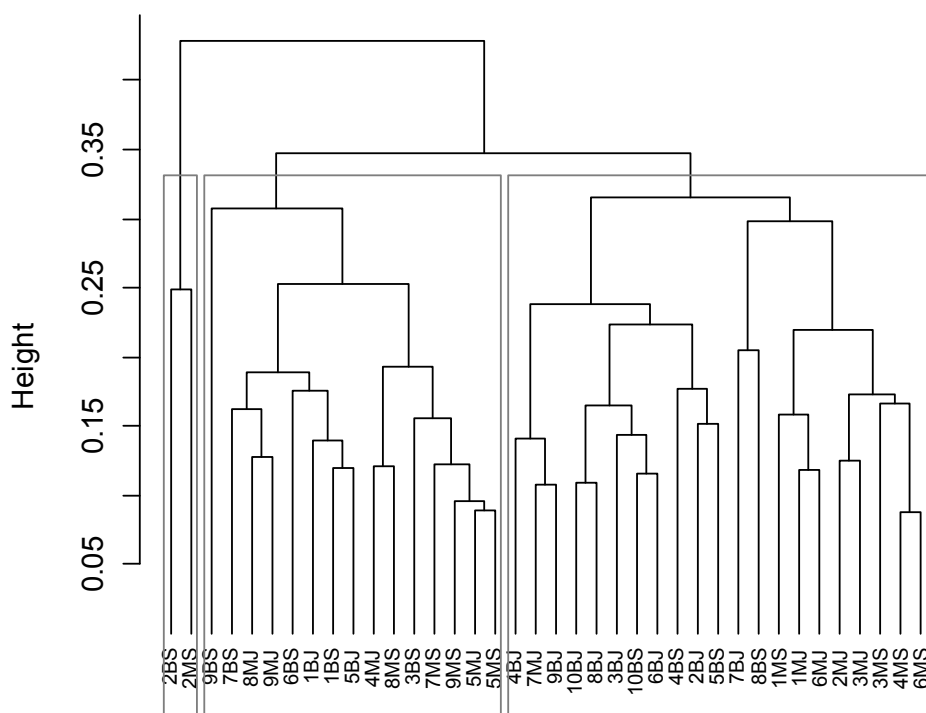


Figure S2 - Dendrogram of hierarchical cluster analyses of microbial communities in the 38 samples collected on Miage (M) and Belvedere (B) glaciers in July (J) or September (S). Numbers indicate sampling locations ordered starting from the glacier terminus. Hellinger distance among communities was used for the analysis. The three main clusters are highlighted by grey boxes.

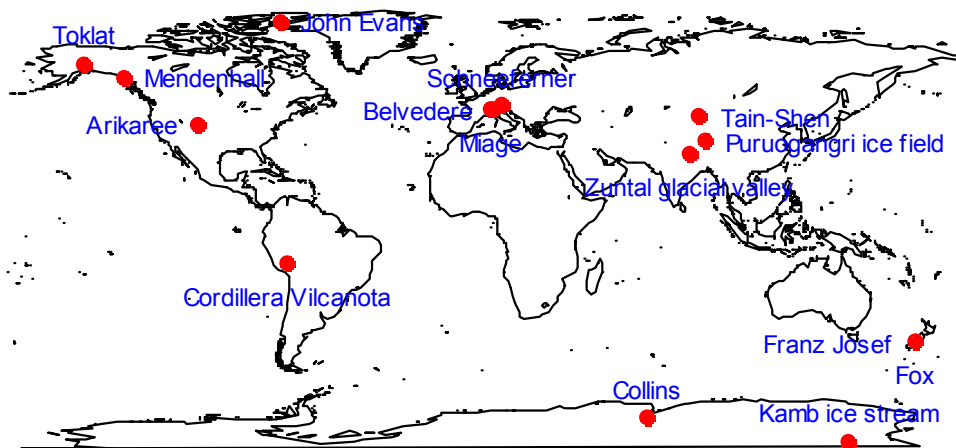


Figure S3 - Location of sampling sites of *Polaromonas* strains (see also Darcy et al. 2011).

CHAPTER 9

Ringraziamenti

Desidero innanzitutto ringraziare il Professor Tortora, per aver accettato il ruolo di Tutor del mio dottorato, per la sua amicizia e per la sua pronta disponibilità ad aiutarmi. Grazie alla Professoressa Bestetti, responsabile ultima del mio progetto, per avermi accolto nel suo laboratorio fin dall'inizio della mia carriera lavorativa. La sua presenza discreta, ma certa ha permesso al nostro gruppo di ricerca di crescere e migliorare. Grazie infinite al Dottor Andrea Franzetti, un caro amico oltre che un grande maestro; grazie di cuore per tutto l'aiuto che mi hai dato, grazie del sostegno ricevuto e grazie per essere così affascinantemente appassionato al tuo lavoro. Guardarti all'opera è uno spettacolo! Grazie mille!!!

Grazie anche al Professor Ambrosini per il suo scrupoloso e attento lavoro indispensabile per il mio dottorato. Grazie perché mi insegni ad essere precisa e accurata, la tua passione traspare dalla tua precisione, è proprio bello imparare da gente così. Grazie a Elena Innocente e al Professor Rampazzo, dell'Università di Venezia, per aver collaborato a questo progetto.

Grazie a tutti i miei colleghi che hanno contribuito a rendere il mio lavoro lieto e interessante. Grazie per i giorni condivisi assieme nella quotidianità della ricerca, che belli! Grazie agli studenti passati dal laboratorio, a quelli interessati e incapaci, a quelli non interessati e capaci e a quelli non interessati e incapaci, perché la loro sola presenza è stata sempre occasione di verifica e autoesame del mio modo di lavorare e di trasmettere le mie, seppur ancora poche, conoscenze.

Grazie alla mia famiglia, da mia nonna Giuseppina fino ad arrivare alla famiglia di mio marito, che mi ha accolto come una figlia, una sorella, una zia, per l'incommensurabile amore che hanno per me. Grazie ai miei amici, indispensabili per la mia vita e alcuni anche per la mia tesi, per il sostegno immancabile e la condivisione della vita. Grazie a mio marito, per aver accettato di essermi al fianco sempre e per avere l'immensa carità di farmi sempre rialzare lo sguardo. Grazie a mia figlia, che non conosco ancora bene, per il semplice fatto di essere stata il dono più grande della mia vita. Il fatto che tu sia arrivata proprio in questo momento così intenso della mia vita ti rende ancora più misteriosa e meravigliosa.