Science of the Total Environment 599–600 (2017) 2095–2104

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Bioaerosol sampling and detection methods based on molecular approaches: No pain no gain

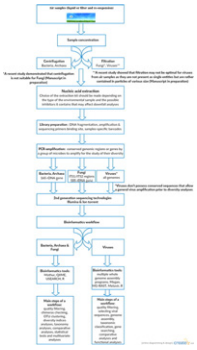
Hamza Mbareche [a,b], Evelyne Brisebois [a,b], Marc Veillette [b], Caroline Duchaine [a,b,*]

[a] Département de biochimie, microbiologie et bio-informatique, Faculté des sciences et de génie, Université Laval, Canada
[b] Centre de recherche de l'Institut universitaire de cardiologie et de pneumologie, Université Laval, Canada

HIGHLIGHTS

- This paper presents a review of the challenges of applying molecular methods to bioaerosols studies
- Comparison of studies using culture-dependent and culture-independent approaches to determine the microbial diversity over the past seven years on a scale of 50 randomly selected studies
- Particular attention should be given to moulds and viruses when recovered from air samples prior to molecular analysis
- The biases introduced by nucleic acid extraction should be given a particular attention when applying molecular approaches, especially next-generation sequencing approaches
- When examining bioaerosols, optimization of environmental microbiology analyses approaches according to the needs of each project is mandatory and ignoring these critical steps will greatly decrease the accuracy of the result

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 1 December 2016
Received in revised form 5 May 2017
Accepted 8 May 2017
Available online 25 May 2017

Keywords:
Bioaerosols
Molecular biology
Air sampling
DNA extraction
Fungi
Bacteria
Viruses

ABSTRACT

Bioaerosols are among the less studied particles in the environment. The lack of standardization in sampling procedures, difficulties related to the effect of sampling processes on the integrity of microorganisms, and challenges associated with the application of environmental microbiology analyses and molecular and culture methods frighten many young scientists. Every microorganism has its own particularities and acts differently when aerosolized in various conditions. Because the air is an extremely biologically diluted environment, it is necessary to concentrate its content before any analysis is performed. Challenges faced when applying molecular methods to air samples reveal the need for a better standardization of approaches for cell and nucleic acid recovery, the choice of genetic markers, and interpretation of data. This paper presents a few of the limits and difficulties tackled when molecular methods are applied to bioaerosols, suggests some improvements by specifying the critical stages that should be considered when studying the microbial ecology of bioaerosols, and provides thoughtful insights on how to overcome the challenges encountered.

* Corresponding author at: Département de biochimie, microbiologie et bio-informatique, Université Laval, CRIUCPQ A-21 2725 CHemin Ste-Foy, G1V4G5 Quebec, PQ, Canada.
E-mail addresses: hamza.mbareche.1@criucpq.ulaval.ca (H. Mbareche), caroline.duchaine@bcm.ulaval.ca (C. Duchaine).

http://dx.doi.org/10.1016/j.scitotenv.2017.05.076
0048-9697/

1. Introduction: challenges in using molecular methods for bioaerosol studies

The use of molecular methods in microbial studies is based on the detection of genetic material of organisms present in a given sample. By applying these methods to different environments, it is possible to identify the microorganisms present and to better understand the environmental impacts and ecological roles of the microbial community. Molecular methods are criticized because they do not relate to organisms' viability or integrity. Propidium monoazide staining was applied to bioaerosol studies of viruses (Bonifait et al., 2015) and bacteria (Bonifait et al., 2014), thereby allowing the specific PCR amplification of targeted genes extracted from intact cells only. Other major challenges associated with the detection of nucleic acid to identify and quantify microorganisms can be summarized as follows: (Bonifait et al., 2015) concentrating and retrieving the biological material from aerosol samples prior to analyses and (Bonifait et al., 2014) choosing the right genomic regions to amplify and/or to sequence to obtain the most accurate information about the microbial content of the environment samB19pled. Bacteria can be quantified and identified using the conserved 16S rRNA prokaryotic gene, after centrifugation of liquid or filtered and re-suspended air samples. For molds, two major challenges are confronted: they cannot be efficiently recovered from liquid air samples by using centrifugation protocols similar to those used for bacteria and archaea (data not shown, manuscript in preparation and see Fig. 1) and it is difficult to choose a genomic region that will lead to a representative image of bioaerosol content. The eukaryotic conserved regions, 18S gene and internal transcribed spacer (ITS) regions, can be used depending on the context of the study. Even though these regions have highly conserved functions in their respective organisms, there are some distinct differences between their variability in eukaryotes and those in prokaryotes. For example, while the V6 region in 16S has been considered as a variable and well suited for assessing bacterial diversity (Huse et al., 2008), its equivalent in the 18S region is more conserved in eukaryotes, which is often avoided for this reason (Sogin, 1991; Hadziavdic et al., 2014). In addition, one of the disadvantages in using molecular methods is that the majority of protocols are specific to each project and differ from one study to another. Specific genes from different organisms can also be detected using molecular approaches. Virulence factor genes, metabolic operons, and antibiotic resistance genes can be quantified from environmental samples (Chizhikov et al., 2001; Martínez, 2008; Richardson et al., 2004). David Stahl developed the concept of molecular microbial ecology on the basis of the sequencing of the 16S rRNA gene region and compared it with different databases (Amann et al., 1990). This concept was used in the last decade to study the microbial diversity in environmental samples. However, we need to be aware of the particularities and pitfalls of these molecular methods, especially when using them with air samples. Viruses do not possess universal and conserved markers, thereby making it impossible to use a blinded molecular approach, unless viral metagenome is retrieved by next-generation sequencing (NGS) methods (Posada-Cespedes et al., 2016). Filtration steps for concentrating the biological material in air samples showed little efficacy when recovering viruses from air samples. The size of particles carrying airborne viruses is not known as it may vary depending on conditions linked to the aerosolization. In addition, the viruses contained in the droplet nuclei, the dried residue formed by evaporation of aerosolized droplets, may not be resuspended during liquid sampling and could remain as a complex virus-containing particle. In general, it is not possible to concentrate viruses because they can be sheltered in a particle whose size is not predictable, not known, or may vary from sample to sample or even within a given sample or the size of the viruses may vary from that of the particle they are trapped in. For example, Lindsley and collaborators showed that although influenza viruses are 80–120 nm in size, they could be found in 1- to 4-μm particles in hospital settings (Lindsley et al., 2010; Verreault et al., 2008).

Although molecular approaches might not provide direct information on dissemination and transmission of diseases, crucial information can be obtained when used to detect DNA sequences of microorganisms in an environment. Indeed, these approaches enable us to directly link aerosol content to the source and to better evaluate air contamination. For example, detection of high concentrations of *Aspergillus fumigatus* in a composting environment by using qPCR is a good indicator that workers might be at higher risk of developing health problems, regardless of the viability of *A. fumigatus* in our samples. Thus, using the NGS approach, the most abundant microorganisms in an environment can be found. Subsequently, qPCR can be applied to obtain a specific concentration per cubic meter of a given genus. Finally, the role of this genus in the transmission and dissemination of diseases can be investigated.

Fig. 1 shows the proposed standardized experimental protocol applied for studying the microbial ecology of an aerial environment by using molecular approaches, particularly the NGS approach. Depending on the types of microorganisms assessed and the environments analyzed, every step is critical. The sampling strategy including the number and types of replicates, volume of air sampled, microorganism concentration method, nucleic acid extraction protocol, genomic targets to be amplified, sequencing technology, and bioinformatics work flow should be carefully chosen and specifically adapted to meet the requirements of the project.

2. Molecular approaches *vs.* culture-based methods

Exposure to bioaerosols, including airborne infectious agents, indoor allergens, fungal agents or viruses, is usually determined by collecting such particles in solid, liquid, or agar media. This is followed by the qualitative and/or quantitative analyses of the samples by using microscopic, microbiological, biochemical, and immunochemical techniques, which offer different degrees of sensitivity and specificity. Unfortunately, not all microorganisms can be cultured. Therefore, examining only the cultivable fraction leads to an underestimation of the total airborne microorganism concentration in a sample. In addition to the variations in culturing ability of microorganisms in an environment, aerosolization and sampling processes create additional stresses that affect the capacity of microorganisms to grow in a culture (Heidelberg et al., 1997). Molecular-based methods can lead to a bacterial concentration of up to 3 orders of magnitude higher than culture-based methods (Nehme et al., 2008). Conversely, an important factor that should be considered when using molecular detection is that not all organisms possess a known unique region catalogued in databases; therefore, not all organisms can be identified by the same way. Viruses represent the principal example of such cases. Moreover, contrary to the current dogma, Shade and his collaborators (Shade et al., 2012) demonstrated that culture-based methods could reveal rare bacteria present in a community. Their results showed that bacteria detected by culture-based methods were either less abundant or absent when studied by molecular approaches. These observations suggest that, to have a true portrayal of microbial richness, culture-based methods should be performed in combination with molecular methods.

Because of the less sequencing costs, multiple studies have tried to establish the relationship between the culture-based and molecular approaches when studying the microbial diversity. This democratization can be visualized in Fig. 2 that shows the distribution of the number of studies using culture-dependent and culture-independent approaches to determine the microbial diversity over the past 7 years on a scale of 50 randomly selected studies. The selection criteria were based on studies that used either culture or sequencing approaches to study a broad range of microorganisms in an environment. Studies that focused on a particular phylum were eliminated. As shown in the graph, the use of culture-independent methods has continuously increased since the NGS methods have invaded the market. However, some studies used both the approaches on the same samples to determine their common potential and relevance. Although the culture-dependent techniques

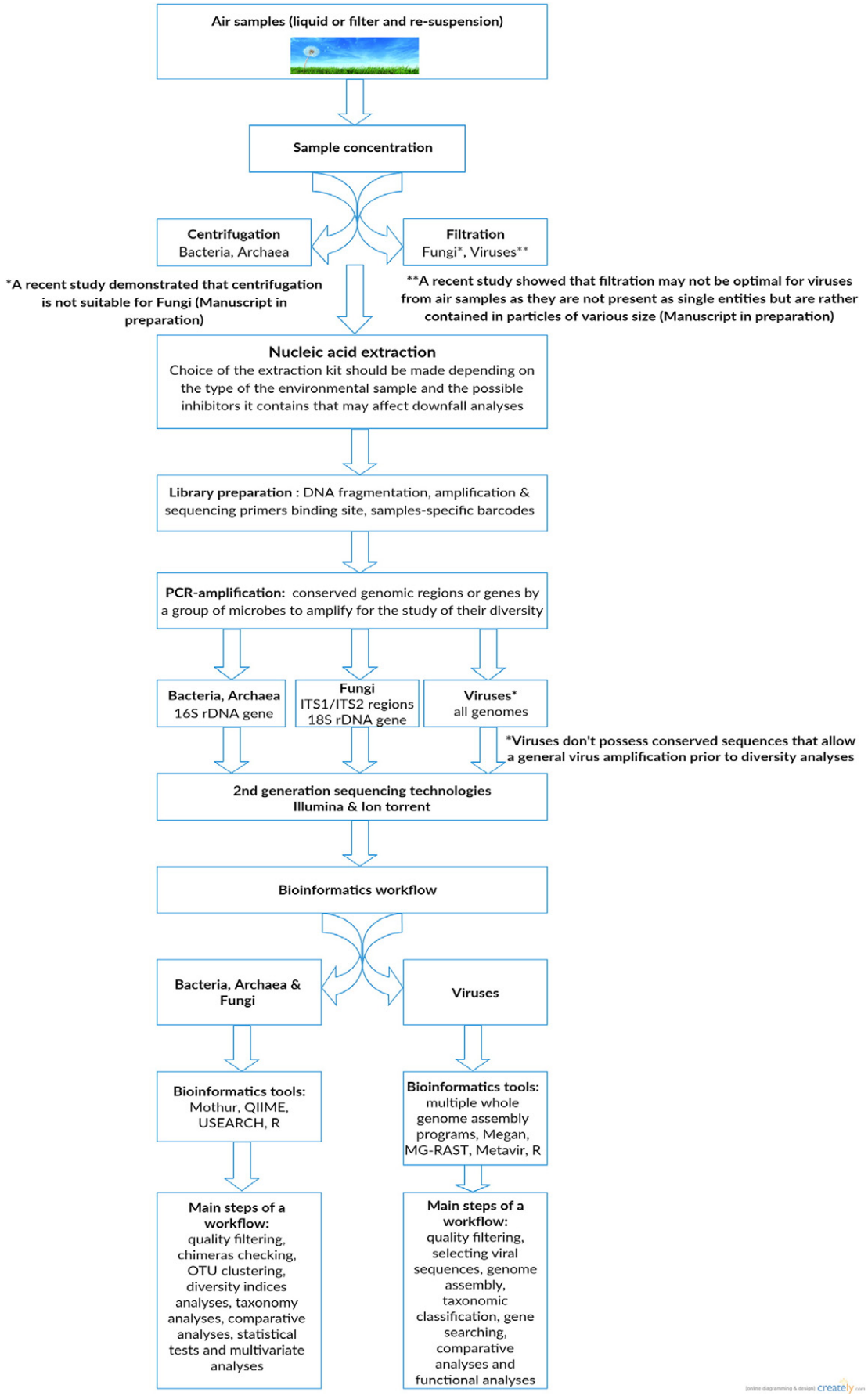


Fig. 1. From the environment to the computer. Experimental protocol of a microbial ecology study based on a sequencing approach.

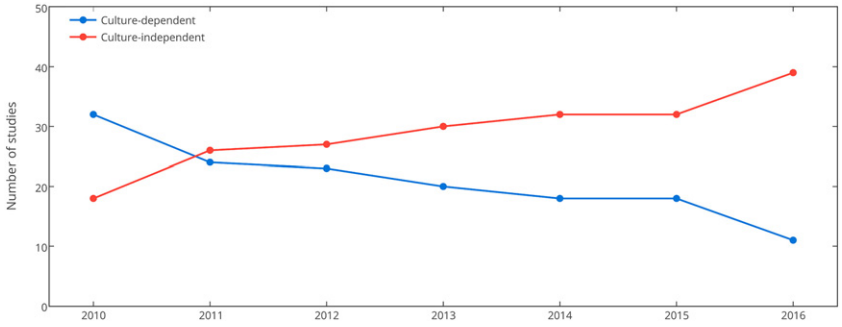


Fig. 2. Distribution of the number of studies using culture-dependent and culture-independent approaches to determine the microbial diversity over the past 7 years on a scale of 50 randomly selected studies.

provide complementary information sometimes, most studies showed a deeper diversity profile when using culture-independent techniques. Table 1 presents a review of some studies in the last 5 years that demonstrated the link between culture-based and molecular approaches with their outcome when determining the microbial diversity in samples from various environments.

3. Bioaerosol sampling

Various aerosol collectors have been developed to capture and analyze airborne particulate matter for airborne pathogen detection and quantification and for biodiversity determination. Air samplers operate in either of the two ways: filtration or gravitational/inertial forces. There are several available air samplers that can be used for subsequent molecular biological analyses. In filtration-based air samplers, aerosols are collected in different types of filters and then removed from the filters by using an elution buffer (Crook, 1995). It is mandatory to remove the particles to be analyzed from the filters to avoid any inhibition from the filters in the following molecular analyses (Després et al., 2007). Among air samplers that operate using gravitational/inertial forces, the most common type used for bioaerosols is the cyclone, where particles from the air are directly associated with the liquid. After concentrating the microorganisms from the liquid, molecular biological method can be applied for microbial diversity analyses (Maron et al., 2005). Electrostatic samplers can also be used for bioaerosols, where particles are collected according to their charges after they are passed through an electrostatic field (Dixkens and Fissan, 1999). After particles are deposited on a medium such as a liquid, they can be analyzed by multiple molecular techniques. One of the advantages of electrostatic sampling is that we can choose the size of the particles to be analyzed. In such cases, dust particle interference presents a major challenge in airborne microorganism sampling. Dust particles, which are largely polydispersed and heterogeneous in nature, may interfere with aerosol collection and detection methods (Moon et al., 2009). The authors of the latter study suggested a new method, *i.e.*, dielectrophoresis, for the separation of biological particles from dust. Environmental factors such as temperature, relative humidity, and ultraviolet radiations from different aerosol sources affect the behavior and fate of the aerosolized particles. In addition, microorganisms' nature, whether pathogenic or not; concentration; relative resistance to stress; and the type of particle they contain must be considered when choosing the bioaerosol sampling method.

Another factor that can affect the choice of air sampling approach is the nucleic acid degradation, especially when studying RNA. When studying metatranscriptomics (the entire transcript material of an environment) or RNA viruses, it becomes important to give particular care to RNA degradation. Indeed, the ribonucleases (often shortened to RNAses) are ubiquitous in nature and are capable of degrading RNA into smaller fragments (Ohtani et al., 1999). Therefore, using a long sampling period can lead to the degradation of nucleic acids, especially RNA. This phenomenon of RNA degradation is exacerbated when filters and dry samplers are used. Turgeon et al. showed that when filters were used for air sampling, the enzymatic activity of influenza A H1N1 neuraminidase decreased compared with that when a wet cyclone sampler was used. They also showed that the neuraminidase enzyme was more stable than influenza RNA genome when filters were used for air sampling (Turgeon et al., 2016). Evidently, a shorter sampling time to avoid the long-term effect of RNAses and the use of a wet sampler should be a priority when looking for the detection and quantification of RNA segments or organisms. The conditions and the mechanisms behind the RNA degradation following aerosol sampling are not well understood. Unfortunately, to the authors' knowledge, there is no aerosol study addressing RNA viruses' stability prior to RT-PCR and the exact half-life of RNA viruses during long-term sampling.

4. Bioaerosol emission patterns and sampling strategies

Some environments and human-related activities can lead to the generation of large numbers of bioaerosols and numerous airborne pathogenic agents. A concrete example is industrial facilities that host confined animal feeding operations (swine barns, dairy farms, poultry, *etc.*), where the bioaerosol sources are numerous and vary with time and space (Nehme et al., 2008; Lecours et al., 2012; Oppliger et al., 2008). These conditions lead to a constant, high bioaerosol generation in confined spaces, and sometimes with limited air change rates and subsequent human exposure. Despite the high bioaerosol generation rates (at some time points), the air still represents a much-diluted medium for microorganisms than other environments. In contrast to solid or liquid matter such as a pile of compost or a lake where particles are enclosed in a finite space, there are no physical boundaries for aerosolized particles as they are more easily dispersed (Brown and Hovmøller, 2002). A high-volume sampling regime is necessary to concentrate bioaerosol content and to increase the probability of discovering rare microorganisms present in this diluted biomass. The volume sampled is linked to two parameters: the flow rate and the time. High-volume sampling can be achieved by a high flow rate of sampling for a relatively short period of time or by a low-to-moderate flow rate of sampling for a longer period of time. Both the scenarios enhance the odds of revealing rare events. An example illustrating the importance of the high-volume sampling for low-concentration pathogen detection was demonstrated in 2012 when Norovirus GII RNA was recovered from air samples in healthcare centers in Eastern Canada by using the high-flow rate sampler Coriolis μ, which was operated at 200 L/min (Bonifait et al., 2015). The nature of the aerosolization source and its environment also affect the type of sampling strategy preferred. For example, a point source, such as a cough or a sneeze, in a closed room requires a low volume of air to be sampled to cover the whole area. Therefore, it will take less time to recover the totality of what was emitted by the point source

Table 1
Studies showing the link between the culture-dependent and culture-independent approaches when determining microbial diversity in various environments.

Sampling environment	Culture *vs.* molecular outcome	References
Bioaerosols (compost)	A review on bioaerosols emitted from composting facilities highlighting the importance of using culture-independent approaches when studying bioaerosols diversity. From the 53 most abundant species, only 11 were detected by culture-dependent methods	Wéry, 2014
Plants (*Cucurbita pepo*)	The number of genera detected with pyrosequencing (Perrott et al., 2013) was higher than the number that could be cultivated (Nehme et al., 2008).	Eevers et al., 2015
Food (sausages)	Sequencing technologies gave more detailed subpopulations at the species level, which were hardly highlighted through culture-dependent approaches.	Fontana et al., 2016
Rhizosphere (*Thymus zygis*)	The number of observed OTUs detected by next-generation sequencing was 3 order of magnitude higher than the cultured bacteria.	Pascual et al., 2016
Soil (mountain)	The isolated genera, which belonged to the phylum *Firmicutes* and *Actinobacteria,* were fast cultivars, whereas few members of *Proteobacteria* were uncultivable even with different cultivation media and conditions.	Venkatachalan et al., 2014
Brewing (rice wine)	Comparative analysis revealed that some species could only be detected using sequencing technique and some species could only be detected by culture-dependent method.	Lv et al., 2015
Soil (petrochemical plant)	Only 8.2% of the bacterial operational taxonomic units were shared between the culture dependent and culture-independent approaches.	Stefani et al., 2015
Food (cheese)	Although few isolates could be identified only by culture dependent method, the culture-independent technique produced a complete inventory of the bacterial species encountered within the Oscypek ecosystem.	Alegría et al., 2012
endotracheal tube	There was little correlation between the results obtained by sequencing and by cultivation. Pyrosequencing identified a much wider variety of bacteria.	Vandecandelaere et al., 2012
Insect	Pyrosequencing recovered 445 microbe operational taxonomic units (OTUs) not detected with traditional techniques.	Kautz et al., 2012
Human skin	When compared to the cultured isolates, culture–independent method showed greater fungal diversity on all skin sites sampled.	Findley et al., 2013
Water	In spite of the expected mismatch between the bacteria targeted by culture dependent and culture-independent methods the results of this study demonstrate the relevance of cultivation approaches to characterize bacterial communities.	Vaz-Moreira et al., 2012
Human feces	The diversity indices (Shannon-Weaver) obtained with culture techniques were lower than those obtained with the molecular techniques, which indicated a superior sensitivity of the molecular technique.	Moles et al., 2013
Breast-milk	Pyrosequencing demonstrated that bacterial diversity in breast milk is by far higher than measured by culture-dependent methods.	Jost et al., 2013

than in a homogenously contaminated open area. In contrast, large environmental sources such as compost piles, wastewater treatment facilities, and swine barns necessitate longer sampling times and sometimes require multiple samples collected over an extended period of time (*e.g.*, one sample per day for a week) to better estimate emissions from the source and evaluate exposure (Wang et al., 2015).

5. Organism peculiarity: what are the main challenges?

Molecular approaches for bioaerosol study were first applied on bacteria (Alvarez et al., 1995; Lange et al., 1997; Orsini et al., 2002). Although other sample content could inhibit molecular reactions, molecular approaches for bacteria and archaea (Nehme et al., 2008; Lecours et al., 2012) are more straightforward, and less fine-tuning is necessary. When analysis of molds and viruses is necessary, several obstacles must be overcome.

5.1. Molds recovered from air in a liquid sample

Because an individual airborne microorganism contains an often undetectable amount of nucleic acids, concentration of microorganisms is necessary before applying molecular techniques to a sample. One of the ways to recover fungi from the air is to use a sampler where the mold spores are suspended in a liquid. Their concentration is then classically found by centrifuging liquid samples and resuspending the pellets in a smaller volume (Vaccari et al., 2006). Centrifugation is a commonly used technique for concentrating bacteria and archaea (Jacobsen and Rasmussen, 1992; Cullison and Jaykus, 2002; Lucore et al., 2000) and can be applied for several types of samples and environmental aerosol samples (Nehme et al., 2008; Lecours et al., 2012; Cayer et al., 2007; Nehmé et al., 2009). However, a recent study demonstrated that centrifugation is not appropriate for mold spores and material recovery (Manuscript in preparation). A comparison between culture-based and qPCR methods both using a centrifugation protocol prior to DNA extraction led to inconsistent results, with the culture-based method showing more concentrated microorganisms in samples. Fungi are particularly difficult to centrifuge as they have many particularities including charges, hydrophobicity, and density, mainly because of their special structure and the fungal protein "hydrophobin," which makes them water-repellent microorganisms. In this study, a new protocol for concentrating mold spores prior to nucleic acid extraction was developed. Samples were filtered using polycarbonate membranes and frozen (flash-frozen) in a buffer solution, after which the frozen filters were pulverized using stainless steel beads and vortexing. Results from using this technique were compared with those from using centrifugation protocols, which showed that the collected DNA from spores by using centrifugation protocols was up to 3 orders of magnitude higher than that collected using the newly developed in-house protocol (Manuscript in preparation). This simple and first step for the recovery of mold material from air samples therefore reveals major problems and leads to underestimation.

5.2. Viruses

Viruses in water samples, *e.g.*, sea and drinking waters, are often concentrated through filtration or ultracentrifugation prior to molecular analyses. Similar to that for bacteria and fungi, this step of concentration is also mandatory for viruses prior to applying molecular approaches to study them in an aerial environmental sample. In contrast to the common thinking, airborne viruses are rarely found as single entities but are rather present as numerous copies attached to particles of various micrometer sizes (Lindsley et al., 2010; Verreault et al., 2008). This statement suggests that the standard universal approaches used to enhance viruses' concentration are not applicable to air samples. Because of the small size of viruses, they are usually isolated and concentrated by filtration or centrifugation. Clearly, concentrating them from an aerosol sample is a challenging task because of their small size as they may be trapped in different and difficult to predict ranges of particle sizes. However, other organisms may need to be treated differently when airborne. Evidently, the stress of the aerosolization may lead viruses to

behave differently compared to other environmental samples. This can be seen in a metagenome analysis of the airborne environment in urban spaces where viruses outnumber bacteria; however, their sequences were under-represented than those of bacteria, probably because the approach used is more suitable for bacteria than viruses (Be et al., 2015).

Another limiting factor in the detection and quantification of viruses, including airborne viruses, is the lack of a universal sequence. Unlike the 16S gene for bacteria and 18S genes/ITS regions for molds, viruses do not share equivalent sequences as no gene or genomic region is homologous across all viruses. Consequently, viral detection can only be achieved by targeting a specific genus or family (Labrie et al., 2013). Another way to detect a broad range of viruses is to use metagenomics (Posada-Cespedes et al., 2016). Thus, a general selection of RNA or DNA viruses can be made by treating nucleic acids with DNases or RNases. It is one of the ways used to target a specific type of viruses.

5.3. Nucleic acid extraction approaches

The nucleic acid extraction process can greatly affect the sequencing analyses by using either PCR or shotgun metagenomic method. Table 2 shows several examples of studies that used molecular approaches, in which the extraction process played a major role in the obtained results. In 2010, Morgan et al. compared different DNA extraction kits and sequencing protocols by using a standardized sample. They concluded that the combination of these two factors considerably affected the observed relative abundances of organisms in the sample (Morgan et al., 2010). Moreover, air samples are fundamentally overdiluted compared to samples from other environments, which may reveal rare events and could lead to potentially larger biases. As the targeted organisms are often in limited concentrations, competition with the background contaminants is a real threat to the authenticity of the results. This is particularly true when working with samples containing a low microbial biomass. Therefore, using different negative controls is crucial in this context to segregate the biological material from the samples from any contaminant in the extraction kit. It remains a challenge to find molecular biology kits that are completely free of residual nucleic acids. However, in a 2014 study, the nucleic acid contamination background was subtracted from data sets, which led to a completely new interpretation of the results (Salter et al., 2014).

The extraction process can lead to different nucleic acid recoveries depending on the organism. For some DNA viruses, qPCR may be performed without DNA purification because the heat shock used to activate the DNA polymerase is sufficient to liberate the DNA from the capsid of all DNA phages tested (Qiu, 2012). However, comparisons of the qPCR yield with and without DNA purification have to be made for all DNA viruses before assuming that this step is necessary. For RNA viruses, nucleic acid extraction is mandatory. The choice of the extraction kit and reagents will also greatly affect the extraction yield. A previous study compared the nucleic acid recoveries of two RNA viruses by using two different extraction methods (Qiagen's QIAamp Viral RNA Mini Kit and Invitrogen's TRIzol®R LS Reagent). RNA extraction was performed on a solution of 1000 plaque forming units per milliliter of pure culture of Phi6 and MS2 phages (Gendron et al., 2010). For both the phages, the first extraction method gave a higher yield than the second one. Additionally, the qPCR quantification for MS2 was two orders of magnitude higher than that for Phi6, even when the same quantity of phages was used. This result may be due to the RNA purification, reverse transcription, and/or qPCR being less efficient for phage Phi6 than MS2. The recovery of fungal DNA is also affected by the extraction process as enzymatic lysis and mechanical agitation of fungal cell walls may lead to different qPCR yields depending on the type of fungi detected (Fredricks et al., 2005). Morgan et al. documented the biases generated by nucleic acid extraction (Morgan et al., 2010). In some cases, it was possible to normalize the output to compensate for less-efficient qPCR detection in the data analysis. These observations highlight the importance of ensuring that samples are prepared using a protocol that is specific to the type of organism we are working with and the need for caution when comparing studies that use different methods.

There are only a few published, peer-reviewed studies of air metagenomes or whole genomes sequenced from a complex environmental sample. A recent paper describing the metagenome found in urban air (Be et al., 2015) observed that the most abundant sequences are from bacterial cells. However, it is known that viruses outnumber bacteria by a few orders of magnitude in the environment (Breitbart and Rohwer, 2005). Viruses are likely underrepresented in the sequences for three reasons: the viral entities in air samples may be difficult to capture, their genetic material could be degraded during the sampling process, and their small genomes make them less likely to be sequenced against millions of other sequences. Although there is an obvious obstacle for concentrating viruses in air samples, the best approach for studying this community is the isolation of virus-like particles by size fractionation and sequencing them using metagenomics. Thus, more effort should be made for increasing viruses' concentration in air samples.

Table 2
Examples of some studies that showcases the possible biases introduced by the extraction approaches applying molecular methods to environmental samples.

Target	Method	Principal findings of biases	References
Bacteria-Eukaryota-Archaea	Metagenomics'	The relative abundances of organisms varied significantly depending on the DNA extraction and sequencing protocol utilized.	Morgan et al., 2010
Bacteria	Metagenomics' & PCR-based sequencing	Bacterial DNA contamination in extraction kits can significantly influence the results of microbiota studies in low microbial biomass.	Salter et al., 2014
Virus (RNA)	Quantitative PCR	Different recovery yields depending on the extraction protocol.	Gendron et al., 2010
Fungi	Quantitative PCR	Different recovery yields depending on the extraction protocol.	Fredricks et al., 2005
Bacteria	PCR-based sequencing	Underestimation of some phyla when applying chemical cell lysis extraction protocol	Zhan and Guo, 2012
Bacteria	PCR-based sequencing	The bacterial distribution varied significantly depending on extraction methods used (bead beating *vs.* silica gel columns)	Starke et al., 2014
Bacteria-Fungi-Archaea	Quantitative PCR & PCR-based sequencing	Major impact of DNA extraction on microbial community structure as none of the ten extraction kits tested resulted in 100% comparable bacterial, archaeal and fungal community composition quantitatively and qualitatively.	Henderson et al., 2013
Viruses (DNA &RNA)	Quantitative PCR	Different recovery yields depending on the level of inhibitors present in the sample and on the extraction kit used	Iker et al., 2013
Bacteria	PCR-based sequencing	Extraction methods were the second-greatest contributing factor to variation in community structure.	Brett et al., 2015
Bacteria	PCR-based sequencing	Contamination can affect interpretation of results in microbiome analysis.	Weiss et al., 2014

6. From the environment to bioaerosols

Even if airborne particles in an environment are often linked to a known source, the proportion of organisms present in that source may not be the same as that in the air. The species of microorganisms found in the source will likely be recovered in air samples, but their proportions can greatly differ. This was first suggested by Parker et al. (Parker et al., 1983), who proposed that respiratory pathogens such as *Mycobacterium* spp. and *Legionella* spp. could be found in higher concentrations in aerosols from the environment. A study comparing biogas microbial content with that of an anaerobic digestor suggested that some microbial phyla are overrepresented in the air than in the source (Moletta et al., 2007). Study of aerosols in compost facilities (Veillette et al., 2016; MBareche et al., 2015) showed that proportions of *Proteobacteria* in compost piles were different from those in the air samples. Some species such as *Methylobacterium* were overrepresented in air samples than in compost. Additionally, *Legionella* was the most abundant airborne species in a carcass compost facility but was not detected in compost samples. This is likely due to its very low concentration relative to those of other species. Conversely, some species such as *Pseudomonas* sp. were present in the compost piles but were not recovered from air samples. These observations confirm that environmental sources cannot be assumed to mirror what is present in the surrounding air. This is a concept known as preferential aerosolization. The hypothesis suggests that some bacteria may be preferentially aerosolized compared to others. A recent *in vitro* study showed that some virulent strains of *Streptococcus suis* were preferentially aerosolized compared to other strains (Gauthier et al., 2016) and that aerosolization of gram-negative (*P. aeruginosa*) and gram-positive (*S. aureus*) bacteria are either more viable or concentrated (PCR) in *in vitro* aerosols (Perrott et al., 2013). Although *in vitro* studies are a proof of concept, further studies that include a larger number of strains and other types of bacteria need to be conducted to better understand the process of preferential aerosolization.

7. Challenges in the bioinformatics pipeline

Currently, microbial data are obtained predominantly using 16S rRNA gene sequencing surveys that provide a characterization of bacterial and archaeal diversity. In studies using 16S rRNA gene sequencing, the choice of primer set depends on a number of factors, including compatibility with previous studies and the specificities of the primers (Soergel et al., 2012). 16S rRNA gene sequence data from different microbial environments present bioinformatical, statistical, and computational challenges. The most widely used bioinformatics tools are QIIME and mothur (Caporaso et al., 2010; Schloss et al., 2009). Both packages are open source and have online tutorials and forums that provide the users a step-by-step analysis of the 16S rRNA gene sequences.

In the NGS PCR analysis, errors are common and are sometimes difficult to detect. Chimeras, which are caused by the incomplete extension of DNA strand during amplification that results in a recombination between the two sequences, can cause biases in diversity results, particularly the alpha metrics (Haas et al., 2011; Ley et al., 2008). Multiple chimera checking softwares are available, and they often have different filtering methods. One should choose the software that fits adequately to the project (Haas et al., 2011; Edgar et al., 2011; Wright et al., 2012; Quince et al., 2011).

Data filtering is often performed at the same stage as demultiplexing, where each sequence is assigned to a sample on the basis of the barcode information. Quality threshold includes quality scores, read length, and the presence of ambiguous base calls. After quality filtering, the microbial groups must be identified through OTU analyses where sequences are clustered together according to the sequence identity. In fact, there are different OTU clustering algorithms, and the use of these algorithms considerably affects the findings and interpretations of data. In *de novo* OTU clustering, sequences are clustered into OTUs by comparing them with the whole dataset, without the use of a reference (Schloss and Handelsman, 2005). In contrast, closed OTU clustering uses a reference sequence database, wherein the sequences that do not match the reference sequence database are discarded. Open-reference OTU clustering is a two-step process that combines both the above algorithms. First, closed-reference OTU clustering is done followed by *de novo* clustering of sequences that fail to match the reference database. Each of the clustering methods has pros and cons, and it should be chosen with a particular attention. For example, if datasets that have sequences from different regions of the 16S rRNA gene need to be combined, the closed reference algorithm should be used, as sequences from different regions of the same 16S rRNA gene would cluster into different *de novo* OTUs. In contrast, using the closed reference algorithm in samples from an undiscovered environment may lead to a high percentage of the sequences being discarded due to a failure to match to a reference database.

Labeling the OTUs is the next step, which is performed using taxonomy assignment algorithm with a reference database. For the 16S rRNA gene, the three main databases are Greengenes, Ribosomal Database Project (RDP), and SILVA (McDonald et al., 2012; Cole et al., 2009; Quast et al., 2013). The microbial diversity analysis is typically described within samples (alpha diversity) and between samples (beta diversity). Multiple scripts for data visualization are available through QIIME, mothur, and R statistic software. A particular attention should be given when choosing the classification and clustering methods depending on the metadata information available about the samples. For example, classification methods can be used to determine which taxa differ between predefined groups of samples. However, clustering methods do not make use of any prior knowledge about the samples. Both methods use between-sample distance metrics, and the specificities of each method used can affect the outcome and the interpretation of the analyses. Therefore, it is mandatory to use several different ways of classification and clustering to ensure that the existence of clusters is not dependent on only one set of parameters. Finally, statistical modeling can be used to explain the variations observed between the different samples. As many covariates can be included and many parameters can be statistically controlled, we strongly recommend the consultation of a statistician or a biostatistics expert during the experimental design and the analyses stages.

Characterization of fungal communities is an area that needs more research. Although the bioinformatics pipeline may seem to be the same for eukaryotic marker genes as that for bacterial marker genes, there are no standardized bioinformatics protocols for 16S rRNA gene analyses. This may be because of the lack of a standard marker gene and a reference database. As mentioned in the introduction, several marker gene options exist. However, the ITS region is generally preferred for obtaining high taxonomic resolution. Because of the amplicon length limitations imposed by the sequencing platforms, only one of the two ITS subregions (ITS1 or ITS2) can be used in a sequencing run. To the best of our knowledge, there are no relevant studies that explored the potential of both the subregions in the characterization of fungal communities in an environment. A strategy that is being explored in a bioaerosol project using multiple environments uses both the subregions to demonstrate their potential and relevance for a global fungal characterization, thereby resulting in a standardized methodology.

The bioinformatics workflow used for the ITS sequences differs greatly from the one used for the 16S rRNA gene in the alignment algorithm, as the ITS region is of different lengths compared to the 16S rRNA gene. Therefore, OTU clustering, reference databases, taxonomy assignment, diversity metrics, data visualization, and statistical modeling should be specific to ITS sequence analyses for the characterization of fungal communities. The UNITE database is often used for ITS sequence-based analyses of fungal sequences (Abarenkov et al., 2010).

However, the 18S rRNA gene can generally be used to analyze eukaryotic communities in the same manner as 16S rRNA genes are used. The SILVA database contains many eukaryotic regions and can

be used for such analyses. However, one should confirm that the region of the 18S gene amplified varies depending on the taxa studied and should be aware that the 18S rRNA gene is not sufficient to characterize the fungal diversity and will likely end up to be of questionable utility.

Characterizing the virome of an environment requires a different approach because, as discussed earlier, no gene or genomic region is homologous across all viruses. Metagenomics is the solution for the absence of amplification primers. In addition, metagenomic data are also used to portray functional potential. The bioinformatics workflow consists mainly of processing the NGS reads by removing the adapter first. Subsequently, low-quality and less-complex sequences are subtracted based on quality scores. Specific quality filtering conditions can be adapted for different downstream analyses (Bokulich et al., 2013). Bacterial and related rRNA reads are removed, and the remaining reads are aligned to a virus database. Different alignment softwares are available. *De novo* assembly algorithms reconstruct original genomes present in the sample by merging short genomic fragments into longer sequences called contigs. There are two main types of *de novo* assembly programs: Overlap/Layout/Consensus assemblers that are more suitable for longer reads and de Bruijn Graph Assemblers that are suitable for shorter reads. Metagenomic reads can be taxonomically classified into similarity- and nonsimilarity-based methods. Similarity-based taxonomic classifications are performed by NCBI BLAST searches where sequences are compared to known genomes. However, a large number of sequences generated from *de novo* assembly are unknown. To classify sequences on the basis of the alignment results, several methods such as MEGAN can be used (Huson et al., 2007). MEGAN uses the lowest common ancestor node of all blast-matching sequences in the phylogenetic tree, which reduces the risk of false-positive matches. In contrast, nonsimilarity-based methods explore the composition of the genomes looking for parameters such as k-mers and are useful to classify sequences that do not have any homologs in reference databases. Many programs are available to perform this composition-based method on viral sequences (McHardy et al., 2007; Mohammed et al., 2011; Wang et al., 2012). Programs such as PHAACS, CatchAll, and UCLUST can be used on the contigs spectra generated by Circonspect to estimate viral diversity (Angly et al., 2005; Bunge et al., 2012; Angly et al., 2006).

Although the bioinformatics workflows described earlier were developed for different environments, many of the issues apply to microbial communities of different habitats. Therefore, they can be applied to bioaerosol samples as well. This review encourages researchers entering the field to use molecular tools to characterize the microbial content of bioaerosols in different environments. This active research area needs more standardized protocols that will help build a bioaerosol sequence catalog, which will in turn allow cross-comparative studies globally and give this field of study a great impetus.

8. Conclusion

Improvements in sequencing technologies led to a decrease in DNA sequencing costs and an increase in sequencing speed. These technologies are widely used in clinical settings but less in experimental settings. The application of these technologies to the study of bioaerosols is a necessary next step. It can be said that the molecular methods can be easily applied to bioaerosol samples. However, many technical details need to be fine-tuned to obtain molecular results that are representative of the sampled environment. Modification of known protocols needs to be considered at every step of an experiment to ensure that the optimum conditions for a specific situation are being used. Depending on the target genes and/or organisms, adjustments may be required starting with the initial sampling process. When examining bioaerosols, other downstream steps such as cell isolation and concentration, nucleic acid isolation, and establishment of databases also have to be retested. Optimization can sometimes be a long and hazardous path, but omitting this critical step will greatly decrease the accuracy of the results.

Acknowledgments

The authors acknowledge Amanda Kate Toperoff for English revision of the manuscript. EB received Institut de recherche Robert-Sauvé en santé et sécurité du travail du Québec (IRSST) scholarship.

References

Abarenkov, K., Henrik Nilsson, R., Larsson, K.H., Alexander, I.J., Eberhardt, U., Erland, S., Høiland, K., Kjøller, R., Larsson, E., Pennanen, T., et al., 2010. The UNITE database for molecular identification of fungi—recent updates and future perspectives. New Phytol. 186, 281–285.

Alegría, A., Szczesny, P., Mayo, B., Bardowski, J., Kowalczyka, M., 2012. Biodiversity in Oscypek, a traditional polish cheese, determined by culture-dependent and -independent approaches. Appl. Environ. Microbiol.:1890–1898 http://dx.doi.org/10.1128/AEM.06081-11.

Alvarez, A.J., Buttner, M.P., Stetzenbach, 1995. PCR for bioaerosol monitoring: sensitivity and environmental interference. Appl. Environ. Microbiol. 61 (103639-3644) (October 1995).

Amann, R., Binder, B.J., Oslon, R.J., Chisholm, S.W., Devereux, R., Stahl, A.D., 1990. Appl. Environ. Microbiol. 56 (6), 1919–1925.

Angly, F., Rodriguez-Brito, B., Bangor, D., McNairnie, P., Breitbart, M., et al., 2005. PHACCS, an online tool for estimating the structure and diversity of uncultured viral communities using metagenomic information. BMC Bioinform. 6, 41.

Angly, F.E., Felts, B., Breitbart, M., Salamon, P., Edwards, R.A., Carlson, C., Chan, A.M., Haynes, M., Kelley, S., Liu, H., Mahaffy, J.M., Mueller, J.E., Nulton, J., Olson, R., Parsons, R., Rayhawk, S., Suttle, C.A., Rohwer, F., 2006. The marine viromes of four oceanic regions. PLoS Biol. 4, e368.

Be, N.A., Thissen, J.B., Fofanov, V.Y., Allen, J.E., Rojas, M., Golovko, G., Fofanov, Y., Koshinsky, H., Jaing, C.J., 2015. Metagenomic analysis of the airborne environment in urban spaces. Microb. Ecol. 69 (2), 346–355.

Bokulich, N.A., Subramanian, S., Faith, J.J., Gevers, D., Gordon, J.I., et al., 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nat. Methods 10, 57–59.

Bonifait, L., Veillette, M., Létourneau, V., Grenier, D., Duchaine, C., 2014. Detection of *Streptococcus suis* in Boaerosols of swine confinement buildings. Appl. Environ. Microbiol. 80 (11), 3296–3304.

Bonifait, L., Charlebois, R., Vimont, A., Turgeon, N., Veillette, M., Longtin, Y., Jean, J. abd Duchaine, C., 2015. Detection and quantification of airborne norovirus during outbreaks in healthcare facilities. Clin. Infect. Dis. 61 (3):299–304. http://dx.doi.org/10.1093/cid/civ321.

Breitbart, M., Rohwer, F., 2005. Here a virus, there a virus, everywhere the same virus? Trends Microbiol. 13 (6), 278–284.

Brett, M.W., Waite D.W., Taylor, M.W. 2015. Evaluating variation in human gut microbiota profiles due to DNA extraction method and inter-subject differences. Front. Microbiol. Volume 6 pages 130. DOI=10.3389/fmicb.2015.00130.

Brown, M., Hovmøller, J.K.M., 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science 297, 537–541.

Bunge, J., Woodard, L., Böhning, D., Foster, J.A., Connolly, S., Allen, H.K., 2012. Estimating population diversity with CatchAll. Bioinformatics 28, 1045–1047.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Peña, A.G., Goodrich, J.K., Gordon, J.I., et al., 2010. QIIME allows analysis of high-throughput community sequencing data. Nat. Methods 7, 335–336.

Cayer, M.-P., Veillette, M., Pageau, P., Hamelin, R., Bergeron, M.-J., Mériaux, A., Cormier, Y., Duchaine, C., 2007. Identification of Mycobacteria in Peat Moss Processing Plants: Application of Molecular Biology Approaches.

Chizhikov, V., Rasooly, A., Chumakov, K., Levy, D.D., 2001. Microarray analysis of microbial virulence factors. Appl. Environ. Microbiol. 67 (7):3258–3263. http://dx.doi.org/10.1128/AEM.67.7.3258–3263.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R.J., Kulam-Syed-Mohideen, A.S., McGarrell, D.M., Marsh, T., Garrity, G.M., Tiedje, J.M., 2009. The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res. 37 (Database issue), D141–D145.

Crook, B., 1995. Non-inertial samplers: biological perspectives. In: Cox, C.S., Wathes, C.M. (Eds.), Bioaerosols Handbook. Lewis Publishers, Boca Raton, pp. 269–283.

Cullison, M.A., Jaykus, L., 2002. Magnetized carbonyl iron and insoluble zirconium hydroxide mixture facilitates bacterial concentration and separation from non-fat dry milk. J. Food Prot. 65 (11), 1806.

Després, V.R., Nowoisky, J.F., Klose, M., Conrad, R., Andreae, M.O., Pöschl, U., 2007 Dec 20. 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 (6), 1127–1141.

Dixkens, J., Fissan, H., 1999 May 1. Development of an electrostatic precipitator for off-line particle analysis. Aerosol Sci. Technol. 30 (5), 438–453.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27, 2194–2200.

Eevers, N., Hawthorne, J.R., White, J.C., Vangronsveld, J., Weyens, N., 2015. Exposure of *Cucurbita pepo* to DDE-contamination alters the endophytic community: a cultivation dependent vs a cultivation independent approach. Environ. Pollut. 209 (2016): 147–154. http://dx.doi.org/10.1016/j.envpol.2015.11.038.

Findley, K., Oh, J., Yang, J., Conlan, S., Deming, C., Meyer, J.A., Schoenfeld, D., Nomicos, E., Park, M., Kong, H.H., Segre, J.A., 2013. Topographic diversity of fungal and bacterial communities in human skin. Nature 498. http://dx.doi.org/10.1038/nature1217.

Fontana, C., Bassi, D., Lòpez, C., Pisacane, V., Otero, M.C., Puglisi, E., Rebecchi, A., Cocconcelli, P.S., Vignolo, G., 2016. Microbial ecology involved in the ripening of naturally fermented llama meat sausages. A focus on lactobacilli diversity. Int. J. Food Microbiol. 236 (2016):17–25. http://dx.doi.org/10.1016/j.ijfoodmicro.2016.07.002.

Fredricks, D.N., Smith, C., Meier, A., 2005. Comparison of six DNA extraction methods for recovery of fungal DNA as assessed by quantitative PCR. J. Clin. Microbiol. 43 (10), 5122–5128.

Gauthier, L.L., Bonifait, L., Veillette, M., Turgeon, N., Perrott, P., Grenier, D., Duchaine, C., 2016. Impact of serotype and sequence type on the preferential aerosolization of *Streptococcus suis*. BMC. Res. Notes 9:273. http://dx.doi.org/10.1186/s13104-016-2073-8.

Gendron, L., Verreault, D., Veillette, M., Moineau, S., Duchaine, C., 2010. Evaluation of filters for the sampling and quantification of RNA phage aerosols. Aerosol Sci. Technol. 44 (10), 893–901.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E., et al., Human Microbiome Consortium, 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res. 21, 494–504.

Hadziavdic, K., Lekang, K., Lanzen, A., Jonassen, I., Thompson, E.M., Troedsson, C., 2014. Characterization of the 18S rRNA Gene for designing universal eukaryote specific primers. PLoS One 9 (2), e87624. http://dx.doi.org/10.1371/journal.pone.0087624.

Heidelberg, J.F., Shahamat, M., Levin, M., Rahman, I., Stelma, G., Grim, C., Colwell, R.R., 1997. Effect of Aerosolization on Culturability and viability of gram-negative bacteria. Appl. Environ. Microbiol. 63 (9), 3585–3588.

Henderson, G., Cox, F., Kittelmann, S., Miri, V.H., Zethof, M., Noel, S.J., et al., 2013. Effect of DNA extraction methods and sampling techniques on the apparent structure of cow and sheep rumen microbial communities. PLoS One 8 (9), e74787. http://dx.doi.org/10.1371/journal.pone.0074787.

Huse, S.M., Dethlefsen, L., Huber, J.A., Welch, D.M., Relman, D.A., Sogin, L.M., 2008. Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. PLoS Genet. 4 (11), e1000255. http://dx.doi.org/10.1371/journal.pgen.1000255.

Huson, D.H., Auch, A.F., Qi, J., Schuster, S.C., 2007. MEGAN analysis of metagenomic data. Genome Res. 17, 377–386.

Iker, B.C., Bright, R.K., Pepper, I.L., Gerba, P.C., Kitajima, M., 2013. Evaluation of commercial kits for the extraction and purification of viral nucleic acids from environmental and fecal samples. J. Virol. Methods 191 (1), 24–30.

Jacobsen, C., Rasmussen, O., 1992. Development and application of a new method to extract bacterial DNA from soil based on a separation of bacteria from soil with cation-exchange resin. Appl. Environ. Microbiol. 58 (8), 2458.

Jost, T., Lacroix, C., Braegger, C., Chassard, C., 2013. Assessment of bacterial diversity in breast milk using culture-dependent and culture-independent approaches. Br. J. Nutr. 110:1253–1262. http://dx.doi.org/10.1017/S0007114513000597.

Kautz, S., Rubin, B.E.R., Russell, J.A., Moreau, C.S., 2012. Surveying the microbiome of ants: comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. Appl. Environ. Microbiol. 79 (2):525–534. http://dx.doi.org/10.1128/AEM.03107-12.

Labrie, S.J., Frois-Moniz, K., Osburne, M.S., Kelly, L., Roggensack, S.E., Sullivan, M.B., Gearin, G., Zeng, Q., Fitzgerald, M., Henn, M.R., Chisholm, S.W., 2013. Genomes of marine cyanopodoviruses reveal multiple origins of diversity. Environ. Microbiol. 15 (5): 1356–1376. http://dx.doi.org/10.1111/1462-2920.12053.

Lange, J.L., Thorne, P.S., Lynch, N., 1997. Application of flow cytometry and fluorescent in situ hybridization for assessment of exposures to airborne bacteria. Appl. Environ. Microbiol. 63 (4), 1557–1563 (April 1997).

Lecours, P.B., Veillette, M., Marsolais, D., Duchaine, C., 2012. Characterization of Bioaerosols from dairy barns: reconstructing the puzzle of occupational respiratory diseases by using molecular approaches. Appl. Environ. Microbiol. 78 (9), 3242–3248 May 2012.

Ley, R.E., Hamady, M., Lozupone, C., Turnbaugh, P.J., Ramey, R.R., Bircher, J.S., Schlegel, M.L., Tucker, T.A., Schrenzel, M.D., Knight, R., Gordon, J.I., 2008. Evolution of mammals and their gut microbes. Science 320, 1647–1651.

Lindsley, W.G., Blachere, F.M., Thewlis, R.E., Vishnu, A., Davis, K.A., Cao, G., Beezhold, D.H., 2010. Measurements of airborne influenza virus in aerosol particles from human coughs. PLoS One 5 (11), e15100. http://dx.doi.org/10.1371/journal.pone.0015100.

Lucore, L.A., Cullison, M.A., Jaykus, L.A., 2000. Immobilization with metal hydroxides as a means to concentrate foodborne bacteria for detection by cultural and molecular methods. Appl. Environ. Microbiol. 66 (5), 1769.

Lv, X.C., Chao, Z., Chen, Z.C., Jia, R.B., Liu, Z.B., Zhang, W., Chen, S.J., Rao, P.F., Ni, L., 2015. Microbial community structure and dynamics during the traditional brewing of Fuzhou Hong Qu glutinous rice wine as determined by culture-dependent and culture-independent techniques. Food Control 57 (2015):216e224. http://dx.doi.org/10.1016/j.foodcont.2015.03.054.

Maron, P.-A., Lejon, D.P.H., Carvalho, E., Bizet, K., Lemanceau, P., Ranjard, L., et al., 2005 Jun. Assessing genetic structure and diversity of airborne bacterial communities by DNA fingerprinting and 16S rDNA clone library. Atmos. Environ. 39 (20), 3687–3695.

Martínez, J.L., 2008. Antibiotics and antibiotic resistance genes in natural environments. Science 32 (5887):365–367. http://dx.doi.org/10.1126/science.1159483.

MBareche, H., Veillette, M., Bonifait, L., Bernard, Y., Marchand, G., Duchaine, C., 2015. Preferential aerosolization of bacteria in the air of different composting plants. 4th International Conference and Exhibition on Occupational Health & Safety August 24–25, 2015 Toronto, Canada.

McDonald, D., Price, M.N., Goodrich, J., Nawrocki, E.P., DeSantis, T.Z., Probst, A., Andersen, G.L., Knight, R., Hugenholtz, P., 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. ISME J. 6, 610–618.

McHardy, A.C., Martín, H.G., Tsirigos, A., Hugenholtz, P., Rigoutsos, I., 2007. Accurate phylogenetic classification of variable-length DNA fragments. Nat. Methods 4, 63–72.

Mohammed, M.H., Ghosh, T.S., Singh, N.K., Mande, S.S., 2011. SPHINX—an algorithm for taxonomic binning of metagenomic sequences. Bioinformatics 27, 22–30.

Moles, L., Gomez, M., Heilig, H., Bustos, G., Fuentes, S., Vos, W., Fernandez, L., Rodriguez, J.M., Jimenez, E., 2013. Bacterial diversity in meconium of preterm neonates and evolution of their fecal microbiota during the first month of life. PLoS One 8 (6), e66986. http://dx.doi.org/10.1371/journal.pone.0066986.

Moletta, M., Delgenes, J., Godon, J., 2007. Differences in the aerosolization behavior of microorganisms as revealed through their transport by biogas. Sci. Total Environ. 379 (1), 75–88.

Moon, H.S., Nam, Y.W., Park, J.C., Jung, H.I., 2009. Dielectrophoretic separation of airborne microbes and dust particles using a microfluidic channel for real-time bioaerosol monitoring. Environ. Sci. Technol. 43 (15):5857–5863. http://dx.doi.org/10.1021/es900078z.

Morgan, J.L., Darling, A.E., Eisen, J.A., 2010. Metagenomic sequencing of an in vitro-simulated microbial community. PLoS One 5 (4), e10209.

Nehme, B., Létourneau, V., Foster, R.J., Veillette, M., Duchaine, C., 2008. Culture-independent approach of the bacterial bioaerosol diversity in the standard swine confinement buildings, and assessment of the seasonal effect. Environ. Microbiol. 10 (3), 665–675.

Nehmé, B., Gilbert, Y., Létourneau, V., Forster, R.J., Veillette, M., Villemur, R., Duchaine, C., 2009. Culture-independent characterization of archaeal biodiversity in swine confinement building bioaerosols. Appl. Environ. Microbiol. 75 (175445-5450) (September 2009).

Ohtani, O., Haruki, M., Morikawa, M., Kanaya, S., 1999. Molecular diversities of RNases H. J. Biosci. Bioeng. 88 (1), 12–19.

Oppliger, A., Charrière, N., Droz, P.-O., Rinsoz, T., 2008. Exposure to bioaerosols in poultry houses at different stages of fattening; use of real-time PCR for airborne bacterial quantification. Ann. Occup. Hyg. 52 (5):405–412 (2008). 10.1093/annhyg/men021.

Orsini, M., Laurenti, P., Bonintí, F., Arzani, D., Lanni, A., Romano-Spica, V., 2002. A molecular typing approach for evaluating bioaerosol exposure in wastewater treatment plant workers. Water Res. 36 (2002), 1375–1378.

Parker, B.C., Ford, M.A., Gruft, H., Falkinham, J.O., 1983. Epidemiology of infection by nontuberculous mycobacteria. IV. Preferential aerosolization of mycobacterium intracellulare from natural waters. Am. Rev. Respir. Dis. 128 (4), 652–656.

Pascual, J., Blanco, S., Garcìa-Lòpez, M., Garcìa-Salamanca, A., Bursakov, A.S., Genilloud, O., Bills, G.F., Ramos, L.J., Dillewijin, P.V., 2016. Assessing bacterial diversity in the rhizosphere of thymus zygis growing in the sierra Nevada National Park (Spain) through culture-dependent and independent approaches. PLoS One http://dx.doi.org/10.1371/journal.pone.0146558.

Perrott, P., Turgeon, N., Veillette, M., Duchaine, C., 2013. Preferential Aerosolisation of Respiratory Pathogens. American Association for Aerosol Research 32nd Annual Conference, September 30–October 3 2013, Porltand, OR, USA.

Posada-Cespedes, S., Seifert, D., Beerenwinkel, N., 2016. Recent advances in inferring viral diversity from high-throughput sequencing data. Viral research. http://dx.doi.org/10.1016/j.virusres.2016.09.016.

Qiu, X., 2012. Heat induced capsid diassembly and DNA release of bacteriophage λ. PLoS One 7 (7), e39793. http://dx.doi.org/10.1371/journal.pone.0039793.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glo¨ckner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41 (Database issue), D590–D596.

Quince, C., Lanzen, A., Davenport, R.J., Turnbaugh, P.J., 2011. Removing noise from pyrosequenced amplicons. BMC Bioinform. 12, 38.

Richardson, P.M., Victor, V., Rubin, Edward M., Rokhsar, Daniel S., Banfield, Jillian F., 2004. Community structure and metabolism through reconstruction of microbial genomes from the environment. Nature 428, 37–43.

Salter, S.J., Cox, M.J., Turek, E.M., Calus, S.T., Cookson, W.O., Moffatt, M.F., Turner, P., Parkhill, J., Loman, N.J., Walker, A.W., 2014. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 12, 87.

Schloss, P.D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Appl. Environ. Microbiol. 71, 1501–1506.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., et al., 2009. Introducing mothur: open-source, platform-independent, communitysupported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Shade, A., Hogan, C.S., Klimowicz, A.K., Linske, M., McManus, P.S., Handelsman, J., 2012. Culturing captures members of the soil rare biosphere. Environ. Microbiol. 14 (9), 2247–2252.

Soergel, D.A., Dey, N., Knight, R., Brenner, S.E., 2012. Selection of primers for optimal taxonomic classification of environmental 16S rRNA gene sequences. ISME J. 6, 1440–1444.

Sogin, M.L., 1991. Early evolution and the origin of eukaryotes. Current opinion in Genetics & Development. Volume1. Issue 4, 457–463.

Starke, C.I., Vahjen, W., Pieper, Zentek, J., 2014. The influence of DNA extraction procedure and primer set on the bacterial community analysis by pyrosequencing of barcoded 16S rRNA gene amplicons. Biochem. Mol. Biol. Int. 2014, 548683 (10 pages). http://dx.doi.org/10.1155/2014/548683.

Stefani, F.O.P., Bell, T.H., Marchand, C., De la Providencia, I.E., El Yassimi, A., St-Arnaud, M., Hijri, M., 2015. Culture-dependent and –independent methods capture different microbial community fractions in hydrocarbon-contaminated soils. PLoS One http://dx.doi.org/10.1371/journal.pone.0128272.

Turgeon, N., Toulouse, M.J., Ho, J., Li, D., Duchaine, C., 2016. Neuraminidase activity as a potential enzymatic marker for rapid detection of airborne viruses. Can. J. Microbiol. 0 (0). http://dx.doi.org/10.1139/cjm-2016-0450.

Vaccari, D.A., Strom, P.F., Alleman, J.E., 2006. Environmental biology for engineers and scientists. Chapter 11: Quantifying Microorganisms and their Activity. Section 11.3: Sampling, Storage and Preparation. Wiley-Interscience, pp. 299–300 (ISBN-13 978-0-471-72239-7).

Vandecandelaere, I., Matthijs, N., Nieuwerburgh, F.V., Deforce, D., Vosters, P., De Bus, L., Nelis, H.J., Depuydt, P., Coenye, T., 2012. Assessment of microbial diversity in biofilms recovered from endotracheal tubes using culture dependent and independent approaches. PLoS One 7 (6), e38401. http://dx.doi.org/10.1371/journal.pone.0038401.

Vaz-Moreira, I., Egas, C., Nunes, O.C., Manaia, C.M., 2012. Bacterial diversity from the source to the tap: a comparative study based on 16S rRNA gene-DGGE and culture-dependent methods. FEMS Microbiol. Ecol. 83 (2013):361–374. http://dx.doi.org/10.1111/1574-6941.12002.

Veillette, M., Bonifait, M., Marchand, G., Duchaine, C., 2016. Bioaerosols exposure in indoor composting installations: classical and molecular approaches and preferential aerosolization. 4th Workplace and Indoor Aerosols Conference 20–22 April 2016 Barcelona, Spain.

Venkatachalan, S., Gowdaman, V., Prabagaran, S.R., 2014. Culturable and culture-independent bacterial diversity and the prevalence of cold-adapted enzymes from the Himalayan Mountain ranges of India and Nepal. Environ. Microbiol. http://dx.doi.org/10.1007/s00248-014-0476-4.

Verreault, D., Moineau, S., Duchaine, C., 2008. Methods for sampling of airborne viruses. Microbiol. Mol. Biol. Rev. 72 (3), 413–444.

Wang, Y., Leung, H.C., Yiu, S.M., Chin, F.Y., 2012. MetaCluster 4.0: a novel binning algorithm for NGS reads and huge number of species. J. Comput. Biol. 19, 241–249.

Wang, C.-H., Chen, B.T., Han, B.-C., Liu, A.C.-Y., Hung, P.-C., Chen, C.-Y., et al., 2015. Field Evaluation of personal sampling methods for multiple bioaerosols. PLoS One 10 (3), e0120308. http://dx.doi.org/10.1371/journal.pone.0120308.

Weiss, S., Amir, A., Hyde, R.E., Metcall, J.L., Song, S.J., Knight, R., 2014. Tracking down the sources of experimental contamination in microbiome studies. Genome Biol. 15, 564.

Wéry, N., 2014. Bioaerosols from composting facilities – a review. Front. Cell. Infect. Microbiol. 4 (42). http://dx.doi.org/10.3389/fcimb.2014.00042.

Wright, E.S., Yilmaz, L.S., Noguera, D.R., 2012. DECIPHER, a searchbased approach to chimera identification for 16S rRNA sequences. Appl. Environ. Microbiol. 78, 717–725.

Zhan, T., Guo, F., 2012. Biases during DNA extraction of activated sludge samples revealed by high throughput sequencing. Appl. Microbiol. Biotechnol. 97, 4607–4616.