Aerobiologia (2020) 36:9–23
https://doi.org/10.1007/s10453-019-09574-2

ORIGINAL PAPER

Relevance of plant and fungal DNA metabarcoding in aerobiology

Elisa Banchi · Alberto Pallavicini · Lucia Muggia

Received: 31 October 2018 / Accepted: 8 March 2019 / Published online: 19 March 2019
© Springer Nature B.V. 2019

Abstract Aerobiology studies biological particles present in the air of both indoor and outdoor environments. In the past few years, the application of molecular techniques has considerably improved the knowledge of organismal diversity in air samples and highlighted its ecological and clinical importance. DNA metabarcoding in particular has been recently applied worldwide to describe the biological composition of airborne bacteria, fungi and plants. Here, we provide a comprehensive overview of those aerobiology studies which, in the past few years, have applied DNA metabarcoding techniques. We focus specifically on researches including plants and fungi and review the multiplicity of experimental plans, sampling procedures, laboratory processing methods, the choice of the DNA barcode(s), the selection of sequencing platforms and bioinformatics pipelines for the comparability of data obtained from multiple origins.

Keywords Air biodiversity · Bioaerosol · High-throughput sequencing · Pollen · Sequencing · Spores

E. Banchi · A. Pallavicini (✉) · L. Muggia (✉)
Department of Life Sciences, University of Trieste, Via Giorgieri 10, 34127 Trieste, Italy
e-mail: pallavic@units.it

L. Muggia
e-mail: lmuggia@units.it

1 Introduction

Aerobiology studies biological particles present in the air of both outdoor and indoor environments (Lacey and West 2006). As other wide and long-established research fields (e.g., ecology, zoology or paleontology), aerobiology has been impacted by the advent and development of molecular techniques that can flank, integrate and improve long-term and more traditional approaches.

Among the newest molecular techniques which have become available, DNA metabarcoding (or targeted amplicon parallel sequencing) has broken through (Taberlet et al. 2012): It uses high-throughput sequencing (HTS) to detect and describe in the most complete way the biological diversity of organismal communities at different levels (Deiner et al. 2017). Its application in the analyses of airborne particles has rapidly increased in the past few years and has proved successful for the understanding of microbial diversity of indoor and outdoor air samples.

Previously published reviews on bioaerosol metagenomics (including DNA metabarcoding) were conducted by Behzad et al. (2015) and Núñez et al. (2016); the latter has, though, focused on urban environments only. The authors highlighted how HTS represents a promising approach for real-time monitoring of bioaerosols and the study of the "air genome." However, the lack of standardized strategies has limited the comparison of data from different

studies. More recently, Yoo et al. (2017) reviewed DNA metabarcoding studies within a broader spectrum of molecular techniques applied to bioaerosol microbial communities and illustrated the need of comprehensive barcode gene databases for the understanding of airborne communities. Microbial communities, however, include both prokaryotes and eukaryotes and are regarded with equal importance either from an ecological/environmental or clinical/medical point of view.

In this review, we consider the contributions in which DNA metabarcoding techniques have been applied to study airborne samples, and specifically focus on those studies including plants and fungi. Here, we also survey the multiplicity of experimental plans, sampling procedures, laboratory processing methods and the choice of the DNA barcodes for the comparability of data among studies. In this way, we aim to (i) provide an overview of the main results obtained so far, (ii) outline challenges and potential solutions in data analyses of airborne samples using metabarcoding and (iii) highlight the importance of the metabarcoding application and its perspectives in studies of airborne particles.

2 Airborne plant and fungal DNA metabarcoding

2.1 Literature survey

Online databases Scopus and Google Scholar were searched for in the period 25–30 September 2018 for retrieving eligible studies to be considered here. Firstly, the terms "air," "airborne," "plants," "pollen," "fungi," "amplicon sequencing," "targeted sequencing," "metabarcoding," "metagenomics," "high-throughput-sequencing" and "microbiome" were searched for in different combinations in "title," "abstract" and "keywords" research fields of the research engines. Secondly, we selected only those studies dealing with HTS-based DNA metabarcoding of airborne plants and fungi, while the studies applying metagenomics shotgun sequencing (i.e., HTS of the total DNA), Sanger sequencing, amplicon cloning and microarrays were excluded, because these were already treated in the reviews by Behzad et al. (2015) and Nuñez et al. (2016). A total of 30 studies were found (four on plants and 26 on fungi; Table 1), which represent only a marginal number in comparison with the huge amount of contributions (more than 30 only in 2018) dealing with community metabarcoding of water and soil samples (Behzad et al. 2015). Nevertheless, the percentage of metabarcoding studies of airborne plants and fungi is constantly increasing, from five in 2016, to eight in 2017 and ten in the first 9 months of 2018 (Table 1). Definitely, this evidences that the study of the aerobiome is still in its infancy.

2.2 Successes versus drawbacks

DNA metabarcoding results of both indoor and outdoor studies are usually consistent with those obtained from the taxonomic identification and abundance assessment of the organismal diversity using microscopy analyses. DNA metabarcoding, though, can provide up to tenfold higher resolution of the airborne biodiversity (Kraaijeveld et al. 2015, for pollen; Banchi et al. 2018a, for fungi). It further offers the possibility to monitor, estimate, compare and taxonomically assign taxa without the need of a priori knowledge of their composition or a specialized expertise at morphological and systematic classification levels. Being DNA metabarcoding a culture-independent approach, it overcomes failures in detecting any organism which may not grow in culture and is dead or dormant (An et al. 2018).

Alternatively, DNA metabarcoding results may be biased by sampling methods, DNA extractions and amplifications, choice of the DNA barcode(s) or lack of reference sequences in public databases. Beside these, DNA metabarcoding is a semi-quantitative method (Degois et al. 2017). These and further aspects are specifically detailed and discussed in the frame of airborne plant and fungal studies in the following paragraphs.

2.3 Sampling methods

Airborne microorganisms are scattered in the atmosphere, thus making their concentration lower than in other environments, such as soil or water (Behzad et al. 2015). A number of sampling approaches are used to collect airborne biological particles, but it is still argued which one would perform best for molecular analyses (Behzad et al. 2015). Indeed, the sampling procedure can overall affect the results, especially concerning the low represented taxa

Table 1 Summary of DNA metabarcoding studies on airborne samples (target) of plants (P), fungi (F), Eukarya (E)

Target	Sampling	DNA barcode	Sequencing platform	Bioinformatic pipeline	Reference database	In/ Out	Country	References
P	Volumetric	trnL	PGM	TSSV	Genbank	Out	Netherlands	Kraaijeveld et al. (2015)
P	Filtration	ITS2	454F, MiSeq	Mothur	Genbank	In	Finland	Korpelainen and Pietiläinen (2017a)
P§	Passive	UPA	MiSeq	QIIME	SILVA	Out	USA	Sherwood et al. (2017)
P#	Volumetric	ITS2, rbcL	MiSeq	Python scripts	Genbank	Out	UK	Brennan et al. (2018)
F	Andersen	ITS	454F	QIIME	RDP	Out	USA	Yamamoto et al. (2012)
F	Passive	ITS1	454F	QIIME	UNITE	In/ Out	USA	Adams et al. (2013)
F	Filtration	ITS	454F	QIIME	Genbank	Out	Israel	Dannemiller et al. (2014)
F	Filtration	LSU	454 J	Mothur	Genbank	Out	South Korea	Oh et al. (2014)
F*	Filtration	ITS1	MiSeq	UPARSE	UNITE	In/ Out	USA	Adams et al. (2015)
F*	Filtration	ITS2	454 J	–	UNITE, EzFungi	In/ Out	South Korea	Shin et al. (2015)
F	Filtration	ITS2	454F	–	Genbank	In	Finland	Korpelainen et al. (2016)
F	Filtration	ITS1	HiSeq	Mothur	Genbank	In	South Korea	Kumari et al. (2016)
F	Filtration	ITS1	MiSeq	QIIME, UPARSE	UNITE	Out	China	Yan et al. (2016)
E	Filtration	18S (V1)	454F	Mothur	SILVA	In/ Out	France	Degois et al. (2017)
F	Filtration	ITS2	454F	Mothur	Genbank	In	Finland	Korpelainen and Pietiläinen (2017b)
F	Filtration	ITS1	MiSeq	–	Genbank	In/ Out	USA	Lemons et al. (2017)
F*	Passive	ITS1	MiSeq	QIIME, UPARSE	UNITE	In	USA	Luongo et al. (2017)
F	Volumetric	ITS1	454 J	QIIME	UNITE, Genbank	Out	NW Europe	Nicolaisen et al. (2017)
F	Passive	ITS2	MiSeq	Mothur	RDP	In	France	Rocchi et al. (2017)
F	Filtration	ITS1	MiSeq	UPARSE	UNITE	In	Honk Hong	Tong et al. (2017)
F	Volumetric	ITS2	454F	–	UNITE	Out	Finland	Abrego et al. (2018)
F	Passive	ITS1, ITS2	MiSeq	Mothur	UNITE, Genbank	Out	France	Aguayo et al. (2018)
F	Filtration	ITS1	MiSeq	QIIME	fungalITSdatabase^	In/ Out	South Korea	An et al. (2018)
F	Volumetric	ITS2	PGM	QIIME	UNITE, Genbank	Out	Italy	Banchi et al. (2018a)
F	Filtration	ITS1	MiSeq	UPARSE	UNITE	In	USA	Coombs et al. (2018)
F*	Filtration	ITS1	MiSeq	Mothur	UNITE	Out	China	Du et al. (2018a)
F*	Filtration	ITS1	MiSeq	Mothur	UNITE	Out	China	Du et al. (2018b)

Table 1 continued

Target	Sampling	DNA barcode	Sequencing platform	Bioinformatic pipeline	Reference database	In/ Out	Country	References
F*	Filtration	ITS1	MiSeq	QIIME	fungalITSdatabase^	In/ Out	Singapore	Luhung et al. (2018)
F	Andersen	ITS1	MiSeq	QIIME	UNITE	Out	South Korea	Woo et al. (2018)
F*	Andersen	ITS	HiSeq	QIIME, UPARSE	RDP	In	China	Yang et al. (2018)

The table reports type of sampling (sampling), DNA barcode sequenced (DNA barcode), sequencing platform, principal bioinformatic pipeline (Bionformatic pipeline), reference database, indoor or outdoor sampling (in/out), country where the study was carried out (country) and the corresponding reference in chronological and then alphabetical order (reference). Sequencing platforms are the following: 454F (454 Life Sciences GS FLX, Roche), 454J (454 Life Sciences GS Junior, Roche), HiSeq (HiSeq, Illumina), MiSeq (MiSeq, Illumina), PGM (Ion Torrent PGM, Thermo Fisher Scientific). Bioinformatic pipeline: QIIME (Quantitative Insights Into Microbial Ecology). Database: RDP (Ribosomal Database Project)

[#]Poaceae; [§]algae; *includes also the analysis of bacteria; ^fungalITSdatabaseID.fasta; °mock community included in the study

(Fahlgren et al. 2011). Fungi (often together with bacteria) and plants are usually sampled by filtration, more rarely by passive collection or using Andersen samplers (Table 1). Aguayo et al. (2018) tested different passive fungal spore traps and recovery protocols and, based on diversity levels, could confirm that the best performances were obtained with Whatman No. 1 filter—coupled with a shaking protocol—and a Whatman No. 3 filter—with a sticky layer coupled with a rubbing spore recovery protocol. Besides the choice of the sampling device, also the air volume represents an important factor. So far no standardized air volume has been proposed to be sampled (Nuñez et al. 2016). However, in Italy for example, national guidelines for the weekly airborne allergenic bulletins, which are based on microscopy identification, are applied (http://www.pollnet.it). Here, the volumetric samplers used (i.e., Hirst type provided with an adhesive tape) have been standardized on a suction flow corresponding to the breath of an average person (10 l/min). However, during DNA analyses we have experienced that the adhesive tapes may contain PCR inhibitors, an issue that deserves further verification.

The choice of a sampling approach should take into account (i) the characteristics of the sampler and (ii) the aim of the study. As both indoor and outdoor fungal concentrations have been estimated in 10^2–10^3 spores m^{-3} of air (Prussin et al. 2014), and daily samples were successfully amplified from volumetric sampling (Banchi et al. 2018a), 14×10^2–10^3 particles are enough for DNA metabarcoding study. On the other hand, samplers based on filtration can reach air flows of > 200 l/min (Núñez et al. 2016). Because both volumetric and filtration samplers have been used to collect plant and fungal airborne samples for DNA metabarcoding (Table 1), it should be the aim of the study (indoor, outdoor, human exposure or environmental application) and the sampling time to determine which sampling procedure is to apply for obtaining the most reliable representation of the airborne organismal diversity.

Guidelines about sampling procedure should help researchers working on the same topic (e.g., workplace bioaerosol exposure) to converge in order to reinforce their results, especially in a context that could affect rules and law of health and work safety.

2.4 DNA extraction

DNA is extracted from airborne samples using mainly proprietary kits (usually PowerSoil DNA isolation Kit, Qiagen, formerly MoBio); to a lesser extent manual, non-commercial DNA extraction methods are performed. Still, no standardized, best-performing DNA extraction method has been proposed yet. While Brooks et al. (2015), working on bacterial DNA metabarcoding, demonstrated that different kits strongly influence the composition and abundance of bacterial mock communities, Bell et al. (2016)

suggested that similar biases may affect also pollen detection. On the other hand, Aguayo et al. (2018) did not find any influence of the DNA extraction kits on airborne fungal samples.

Pollen has a high structural integrity due to the exine layer (Bell et al. 2016), whereas in fungal spores and hyphae chitin confers resistance to the cell walls (Klimek-Ochab et al. 2011). Therefore, the DNA extraction of airborne plant (pollen) and fungal particles (spores and hyphael fragments) should consider also an efficient method to lyse the cells for maximizing the DNA yield. The mechanical cell disruption (i.e., using bead-beaters) seems to be the most effective method for retrieving enough DNA quantity to be analyzed in the most complete way (Behzad et al. 2015). An ineffective cell disruption, indeed, would bias the DNA amount augmenting the amount of extracted DNA for those more fragile and single-nucleate cells than for the more resistant ones, making the first be over-represented if the more resistant are not lysed. It is therefore of key importance to include an efficient, mechanical cell-disruption step prior to DNA extraction.

2.5 DNA barcodes for plants and fungi

A DNA barcode needs to satisfy different requirements: to be enough interspecific variable and conserved in the same species, to have a suitable length for an effective amplification and sequencing and to present highly conserved flanking regions for primer design (Kress and Erickson 2008). The choice of a barcode is nowadays constrained by the sequencing platforms that provide different read lengths which can affect the barcode validity. Nevertheless, even if a standard barcode is selected, it is necessary to progressively verify its universality and effectiveness.

The determination of an universally accepted plant DNA barcode has always been debated, as different markers (used individually or combined) have been proposed and both plastidial and nuclear regions have been applied (Cheng et al. 2016). The markers that have been applied for pollen DNA barcoding are the plastidial rbcL, matK, trnL, trnH-psbA and the nuclear ribosomal DNA (rDNA) ITS2 (Bell et al. 2016), being rbcL, trnL and ITS2 used also in DNA metabarcoding of airborne pollen (Kraaijeveld et al. 2015; Korpelainen and Pietiläinen 2017a; Brennan et al. 2018).

Green algae usually lack the matK locus, with the only exception of Charophytes (Buchheim et al. 2011); therefore, a part of the 23S region, named Universal Plastid Amplicon (UPA, Sherwood and Presting 2007), is also used in DNA barcoding and metabarcoding studies (Sherwood et al. 2017).

Even if the presence of plastids in pollen is still debated (Azhagiri and Maliga 2007; Tang et al. 2009), plastidial markers have been successfully amplified in airborne pollen samples and used in melissopalynology analyses (Galimberti et al. 2014; Richardson et al. 2015). Brennan et al. (2018) found that the ITS2 locus provides a higher taxonomic resolution than rbcL in identifying grass pollen. The ITS region in fact (or only part of it, e.g., the ITS2 subunit) is commonly used in plant molecular systematics because of its potentially high resolution at genus and species level (Cheng et al. 2016). Bell et al. (2016), in a review on pollen DNA barcoding, suggested trnL and ITS2 as the best markers, because combining suitable sequencing length with genetic variability.

ITS has been also selected as the universal, formal marker for fungi (Schoch et al. 2012) due to its high interspecific variability and conserved primer sites. Due to length constraints still intrinsic of many sequencing platforms, only one of the two subunits, either ITS1 or ITS2, is usually analyzed in DNA metabarcoding, though it is still argued which one provides the best taxonomic resolution in fungi (Banchi et al. 2018b). ITS1 is reported as the most commonly used barcode for fungal aerobiological studies (13 publications), followed by ITS2 (6) and more rarely by the complete ITS region (5; Table 1). Aguayo et al. (2018) sequenced and compared the results from the two subunits and found that choosing ITS1 or ITS2 affected the read number (higher when sequencing the ITS2) and the Shannon index (higher when sequencing the ITS1) but did not affect the estimation of species richness of airborne fungi. In taxonomic identification, ITS1 and ITS2 shared 44% of the total and 84% of the 50 most abundant operational taxonomic units (OTUs). Barcode-dependent biases included length of the amplicons, taxonomy and primer mismatch (Aguayo et al. 2018). For instance, as the ITS1 primers used for DNA metabarcoding did not efficiently amplify the genus *Puccinia*, Nicolaisen et al. (2017) implemented also a qPCR analysis for its detection.

Cheng et al. (2016) have recently developed specific plant and universal ITS primers for the amplification of the entire ITS region and the individual ITS subunits. These new primers can also be differently combined for the amplification of both fungi and plants from environmental samples (Banchi et al. unpublished).

Within the 26 studies on airborne fungi here considered (Table 1), only two used rDNA markers other than ITS, such as the nuclear large subunit (LSU; Oh et al. 2014) and the V1 region of the 18S (Degois et al. 2017, focused on the amplification of all Eukarya).

2.6 Sequencing platforms and bioinformatic pipelines

Sequencing platforms and bioinformatics pipelines are heterogeneous among studies and independent from the group of targeted organisms (Table 1). The 454 pyrosequencing platform (Roche) was the most frequently used up to 2013 and was later surpassed by Illumina sequencing (Núñez et al. 2016; Lear et al. 2018), which is currently the preferred platform (Lindahl et al. 2013). This emerges also from the studies here considered; the most used HTS platform is MiSeq (Illumina) followed by 454 Life Sciences GS FLX (Roche) (Table 1). MiSeq (Illumina) has a read length of up to 2 × 250 bases at relatively low cost, while 454 pyrosequencing, although it provides longer reads (up to 800 bases), is more expensive and out of the market since 2016. Ion Torrent (PGM, Thermo Scientific, read length up to 400 bases) has been used only few times (Table 1), but given its short run-time and the decreasing sequencing costs, it is supposed to be progressively employed in the next years. None of the studies so far has used the so-called third-generation sequencing platforms, such those commercialized by Pacific Biosciences or Oxford Nanopore technologies, which produce long reads (up to thousands kb) but still of too low quality. So far, only Tedersoo et al. (2018) report a DNA metabarcoding analysis of fungi and other Eukarya from soil samples using the PacBio RS platform. Although sequencing technologies have intrinsic platform-specific biases (D'Amore et al. 2016; Reuter et al. 2016), the proportion of the most abundant OTUs is highly congruent for Fungi and Eukarya when reads were obtained by MiSeq and 454 platforms (Mahè et al. 2015; Tessler et al. 2017).

The large amount of data that HTS produces require efficient and accurate bioinformatics analyses, and an important step is the standardization of a bioinformatics pipeline (Bell et al. 2016). The most common pipelines used in DNA metabarcoding analyses for airborne plants and fungi are implemented in QIIME (Caporaso et al. 2010) and Mothur (Schloss et al. 2009). Both offer scalable and user-friendly approaches and are open-source and suitable for the analysis of data obtained from the main sequencing platforms. These programs perform cleaning, trimming and reads alignment, OTUs clustering, taxonomic assignment, phylogenetic analyses and α- and β-diversity analyses (Cui et al. 2013). QIIME is implemented in Python, while Mothur is implemented in C++; they also differ in the alignment methods, OTUs taxonomic assignment and clustering algorithm (Plummer and Twin 2015). None of the studies reviewed here have compared different pipelines. On the contrary, such a comparison was done for bacterial 16S data and showed that either the results obtained were comparable and consistent (Plummer and Twin 2015) or they were significantly different (Golob et al. 2017). For this reason, and in agreement with the suggestion of Golob et al. (2017), it should become a common praxis to include analysis details, such as the software version, the OTU-picking strategy, algorithm and reference database, information about taxon classification, and the reference database selected for allowing reproducibility and interpretability of the data.

Still, an *ad hoc* DNA metabarcoding database has not been constructed for airborne plants yet; the DNA sequences used as references are usually the ones deposited in GenBank (Table 1). Nevertheless, the BOLD database (Ratnasingham and Hebert 2007) and the ITS2 database V (Ankenbrand et al. 2015) offer specific plant sections, with different DNA barcodes and ITS2, respectively. An rbcL database for plants has been recently curated by Bell et al. (2017), as this barcode increases taxonomic resolution for some angiosperm species if used in addition to ITS2.

Alternatively, for fungi most of the studies use the UNITE database (Kõljalg et al. 2013), which includes more than 7000 reference sequences. UNITE, even if not complete, is the largest customized and curated public database for fungal ITS sequences (Nilsson

et al. 2015) and has been formatted for both QIIME and Mothur pipelines. According to our experience, the use of UNITE should be encouraged in studies applying fungal DNA metabarcoding. However, the taxonomic assignment can still be affected by synonymization or misclassification of fungal taxa. This may be the case of those species which are recorded with their sexual and asexual forms (i.e., teleomorphs and anamorphs), for which different names still occur in the database (Cui et al. 2013).

2.7 Rarefaction curves

Rarefaction curves are usually calculated to estimate the efficacy of the HTS effort. For each sample, the number of taxa or OTUs is plotted as a function of the reads number; if the curve levels off, the sampling effort (i.e., sequencing depth) is enough for the estimation of the total biodiversity of the sample. In the studies of Banchi et al. (2018a), Oh et al. (2014) and Yamamoto et al. (2012), the sampling in outdoor urban areas did not reach complete saturation. This was evident by the low number of obtained reads (samples contained from 400 to 5000 reads and from 200 to 600 OTUs), indicating that further biodiversity could have been detected. Rocchi et al. (2017), on the other hand, obtained saturated curves in indoor airborne samples with a much higher sampling effort (> 30,000 reads and 600–1300 OTUs per sample). An accurate evaluation of the best sequencing strategies (e.g., number of samples per run, type of sample) should take into account also this aspect, and the analyses of rarefaction curves should be an ever considered step in the assessment of the sequencing effort.

The inflated number of OTUs and of erroneously identified OTUs emerges as a further important aspect in DNA metabarcoding studies. Most OTUs clustering algorithms are based on sequence similarity threshold: Usually, the 97% similarity threshold is applied to segregate species (Nguyen et al. 2016). After clusterization, low represented OTUs (e.g., singletons) are commonly removed, assuming that at least a portion of them represents amplification/sequencing errors (Bálint et al. 2016). However, acting in this way, only dominant species are usually taken into account, biasing the estimation of the real biodiversity (Frøslev et al. 2017). To tackle this issue, different OTUs post-clustering methods have been recently proposed, such as dbotu3 (Olesen et al. 2017) and LULU (Frøslev et al. 2017). LULU excludes artefactual OTUs retaining rare but real OTUs, combining sequence similarity and co-occurrence patterns (Frøslev et al. 2017). This method has been validated using ITS2 DNA metabarcoding sequences of plants (Frøslev et al. 2017), and it represents a promising tool to be considered for standardization and comparison of data in aerobiology.

Further, DNA barcoding is not a quantitative technique, but it only provides the relative abundance of taxa (An et al. 2018). A precise estimation of the concentration of specific airborne taxa can be obtained combining HTS and quantitative PCR (qPCR). However, when comparing HTS- and taxon- specific qPCR-based concentrations, consistent results are found. Dannemiller et al. (2014) correlated HTS and qPCR for five fungal taxa (*Alternaria alternata*, *Cladosporium cladosporioides*, *Epicoccum nigrum*, *Penicillium* spp. and *Aspergillus* spp.) from outdoor airborne samples. The authors showed that the absolute concentration estimated with HTS was strongly associated with the values obtained with qPCR, similarly to that Yamamoto et al. (2012) found for the same fungal genera in outdoor samples as well. Inside buildings provided with wood heating, Rocchi et al. (2017) confirmed by qPCR the predominance of *E. nigrum* detected previously by DNA metabarcoding. The most precise estimation of the abundance of certain taxa is a key when clinically important allergens and pathogens are detected by HTS, for which there is the need to define species- and dose–effect relationships (Rocchi et al. 2017).

2.8 Mock community

Only four of the 30 studies here considered included a positive control for DNA amplification and HTS, namely the mock communities (Table 1). Mock communities are samples composed by an ad hoc DNA mix of known organisms, pooled in pre-defined (usually equimolar) amount. The mock community is processed in the same way as the other samples and allows to estimate the reliability of the results in terms of taxonomic identity and proportion of reads (i.e., relative abundance), preferential amplification, identification of primer biases, false positive sequences and the presence of any contamination (Lear et al. 2018). It is likely that the use of standardized mock

communities could help the whole standardization of laboratory procedures and the comparison of data among different studies. "Home-made" mock communities have been prepared in individual studies for grasses (Kraaijeveld et al. 2015) and fungi (Adams et al. 2015; Rocchi et al. 2017; Aguayo et al. 2018). Bakker (2018) recently developed new standardized fungal mock communities available upon request. Though, airborne bacterial mock communities are already commercialized (Lear et al. 2018), to the best of our knowledge, no mock communities have been established for airborne plant and fungi yet.

3 Plant and fungal detection in indoor/outdoor airborne samples

3.1 Main applications

Indoor and outdoor environments have represented very common settings to perform DNA metabarcoding studies of airborne plants and fungi (Fig. 1).

Studies related to indoor environments have focused on human exposure to pollen and fungal spores, mainly in relation with allergies and health problems (Shin et al. 2015; Korpelainen et al. 2016; Korpelainen and Pietiläinen 2017a, b; Luongo et al. 2017), occupational exposures (Kumari et al. 2016; Degois et al. 2017; Yang et al. 2018), characteristics of the buildings (Lemons et al. 2017; Rocchi et al. 2017; Coombs et al. 2018) and human occupancy (Adams et al. 2015; Luongo et al. 2017; Tong et al. 2017).

Sampling in outdoor environments, alternatively, offers the possibility to identify which particles are present in or transported to a certain area (e.g., urban, rural), and to assess spatial and temporal shifts (e.g., daily or seasonal) of the communities also in response to different environmental conditions or climate changes, such as pollution, dust events, haze (Yamamoto et al. 2012; Dannemiller et al. 2014; Oh et al. 2014; Kraaijeveld et al. 2015; Yan et al. 2016; Nicolaisen et al. 2017; Sherwood et al. 2017; Abrego et al. 2018; Aguayo et al. 2018; Banchi et al. 2018a; Brennan et al. 2018; Du et al. 2018a, b; Woo et al.

Fig. 1 Schematic representation of indoor and outdoor (urban and natural) environments in which the taxonomic composition and the estimation of plant and fungal airborne diversity are assessed by DNA metabarcoding. General applications include human exposure (including occupational exposure) to biological airborne particles, the identification of allergens, invasive alien species (IAS), human and plant pathogens, as well as the spatial–temporal characterization of airborne communities. Indoor studies also deal with the influence of human presence inside buildings and building features (e.g., materials, types of ventilation, mold problems) on pollen and fungal composition. Outdoor studies assess the influence of meteorological events (e.g., precipitations, wind, humidity), pollution, size of particulate matters (PMs) and environmental changes

2018). Outdoor sampling is also applied for the early detection of plant and human pathogenic fungi and their spread and potential infection events in crops, and it is necessary to detect the dispersion of invasive alien plant species (IAS).

Interesting and novel research lines can merge the knowledge from meteorology, bioremediation and applications in atmospheric transformation to uncover relationships among multiple microbial communities, the effect on ecological processes, and the transport and dynamics of biological particles relevant to human health and agriculture (Prussin et al. 2014; Behzad et al. 2015).

3.2 Plants

So far, only four works have studied airborne pollen (Kraaijeveld et al. 2015; Korpelainen and Pietiläinen 2017a, b; Brennan et al. 2018) and airborne algae diversity (Sherwood et al. 2017) using DNA metabarcoding. Korpelainen and Pietiläinen (2017) analyzed indoor pollen in sampling nurseries, university buildings and a farmhouse: The majority of the sequences belonged to wild plants in most samples. Overall, the more common taxa detected were the allergy-inducing *Brassica* sp. and *Betula* sp., the house plant *Saintpaulia* sp. and the garden ornamental *Syringa* sp. Even if a great variation was highlighted, the authors did not find clear patterns among the sites and the different seasons in terms of taxonomic composition and diversity. In outdoor sampling, Kraaijeveld et al. (2015) applied for the first time DNA barcoding to airborne pollen in Netherlands and provided an accurate, qualitative and quantitative taxonomic assessment, highlighting the potential of this technique to increase the efficiency of pollen monitoring. The more abundant taxa detected were *Betula* sp., species of Pinaceae and Poaceae (grasses), *Urtica* sp., *Acer* sp., *Castanea* sp. and *Salix* sp. Comparing the molecular results with pollen counts, Kraaijeveld et al. (2015) found that the resolution between the two was particularly divergent for grass genera (up to tenfold), which are particularly difficult to identify by morphological traits.

DNA metabarcoding was used by Brennan et al. (2018) to detect seasonal changes in the spatial and temporal distribution of Poaceae pollen throughout the grass allergy season in the UK. Grass genera presented peaks of pollen incidence influenced by latitude. Overall, the most abundant genera were *Alopecurus*, *Festuca*, *Lolium*, *Holcus* and *Poa*, which are known to trigger IgE-mediated responses in grass-sensitized patients (Brennan et al. 2018). This study highlighted how socioeconomic advantages may be taken from a better knowledge of the link between taxon-specific exposure of harmful grass pollen and allergic diseases (Brennan et al. 2018). The implementation of better refined airborne pollen profiling is therefore important.

Sherwood et al. (2017) compared airborne algae from windward and leeward sides of a mountain range in an Hawai'ian island, revealing a fine scale distribution of the airborne algal communities. The OTUs found included mostly cyanobacteria (96.4%) and, in lower amount diatoms, cryptophytes, brown algae and green algae. The biodiversity was higher in windward sites than in leeward sites; the two sides shared only a low number of OTUs which remained mostly uncharacterized at deep taxonomic level.

3.3 Indoor fungi

Workers are daily exposed to specific bioaerosols in their workplaces. Occupational bioaerosol fungal exposure has been investigated with DNA metabarcoding in waste sorting plants (Degois et al. 2017), in swine houses (Kumari et al. 2016) and in chicken farms (Yang et al. 2018). In these environments, Ascomycota is the predominant group of fungi: In waste sorting plants airborne fungi were dominated by genus *Penicillium* and *Aspergillus* (Degois et al. 2017), while in chicken houses the genus *Aspergillus* was the most abundant (Yang et al. 2018). Kumari et al. (2016) found that relative humidity, particulate matters (PMs), NH_3, and stocking density were the main drivers of the variation in fungal composition in swine houses: Here, the most abundant taxa at genus level were *Clavaria* and *Fusarium* (Kumari et al. 2016). Such studies could provide a better understanding of the biological risks in workplaces, as biological airborne particles can contribute to the insurgence of occupational diseases (Degois et al. 2017).

As many people spend most of their time indoor, the characterization of airborne microbiome is important also for public health agencies dealing with monitoring and management of indoor air quality (Shin et al. 2015). University buildings and childcare

facilities were the most investigated places (Shin et al. 2015; Korpelainen et al. 2016; Korpelainen and Pietiläinen 2017b; Luongo et al. 2017). In a childcare facilities study, Shin et al. (2015) found differences in bacterial but not in fungal composition between indoor and outdoor samples. The most abundant group was the basidiomycetous Agaricomycetes (∼ 70%). Human skin-associated fungi, such as *Malassezia* species, were more abundant in indoor air samples, and among the most common allergic fungi, only *Alternaria* and *Penicillium* were significantly more abundant outdoor.

The indoor microbiome likely depends on the outdoor biome, occupants' microbiome and behavior, and the building characteristics (Adams et al. 2013). To test whether buildings features influence the indoor airborne fungal communities, Coombs et al. (2018) compared airborne fungal diversity between green and non-green homes in USA, showing that green renovation does not alter the indoor mycobiome. Studies showing that skin test reactivity to fungal and dust mite allergens was increasing in asthmatic and allergic children living in homes with evaporative cooler, prompted Lemons et al. (2017) to evaluate the fungal and bacterial burden inside homes which use traditional air conditioners or evaporative coolers. While the bacterial composition was similar, the fungal diversity differed significantly: The proportion of Ascomycota was lower in air-conditioned homes (∼ 50%) than in homes with evaporative coolers (∼ 90%). The most prevalent species discovered in air-conditioned homes belonged to the genera *Cryptococcus* and *Aspergillus*, while in evaporative cooled homes the most common fungi belonged to more hydrophilic fungal taxa, in particular represented by Pleosporales, and included *Alternaria alternata* and *Phoma* spp. Heating systems also seem to have influence on fungal composition and abundance; Rocchi et al. (2017) found significantly less fungal diversity and a higher presence of *Epicoccum nigrum* in wood-heated rooms than in rooms heated with other systems, such as electricity.

The fungal taxa found indoors are primarily influenced by those coming from outdoors, while the occupants of a building usually account for the increasing abundance of skin- and body-associated fungi detected in the indoor fungal composition (Adams et al. 2015; Luongo et al. 2017; Tong et al. 2017).

3.4 Outdoor fungi

Most of the studies considered urban areas located worldwide, e.g., Asia, USA and Europe (Table 1). Urban areas are of particular and general interest because of the high concentration of people (Revision of World Population Prospects United Nations 2017), which are likely more affected by allergic respiratory symptoms than people from rural areas. This might be a consequence of the increasing atmospheric concentration of pollutants and particulate matter (PM) in the cities (D'Amato 2002). Fungal spores, together with bacteria, pollen, plant and animal debris, can account for up to the 25% of PMs (Jaenicke 2005; Yan et al. 2016).

The association between PMs, diversity and composition of airborne fungi were investigated in a highly polluted area subjected to smog and haze in Beijing (China) (Yan et al. 2016; Du et al. 2018a, b). Here, the fungal community was dominated by Ascomycota, in particular by Pleosporales and Capnodiales. *Cladosporium* and *Alternaria* were the two most abundant genera in all types of PM samples. PMs, together with relative humidity (RH), temperature and season of the year were significantly correlated also with the variation of fungal diversity and community composition.

Oh et al. (2014) and Woo et al. (2018) studied the airborne species composition in Seoul (Korea) during spring and in the period May–November, respectively. Ascomycota were in general predominant during the whole year with the most common taxa represented by *Cladosporium*, *Alternaria* and *Epicoccum*. Only in the middle of spring the basidiomycetes *Perenniporia* and *Trametes* were more abundant.

In an urban locality in USA, the most abundant Ascomycota genera were *Leptosphaerulina*, *Epicoccum* and *Cladosporium*, while Basidiomycota were represented majorly by *Peniophora*, *Exidia* and *Stereum* (Yamamoto et al. 2012).

Particle size influences the distribution of Ascomycota and Basidiomycota, being Ascomycota relative more abundant in particulate with an increasing aerodynamic diameter, and Basidiomycota abundant in smaller particles (Yamamoto et al. 2012; Woo et al. 2018).

In Europe, metabarcoding studies on airborne fungal communities in urban areas were conducted only by Dannemiller et al. (2014), Nicolaisen et al.

(2017), Abrego et al. (2018) and Banchi et al. (2018a) so far. In Israel, the fungi associated with PM10 in spring and fall varied in diversity (higher in spring) and taxonomic composition according to season. In spring, the most abundant genera were *Botryotinia*, *Cryptococcus*, and *Erysiphe*, while in fall *Arthothelium*, *Aspergillus* and *Cladosporium* (Dannemiller et al. 2014). Overall, Ascomycota was the most abundant phylum, with *Cladosporium* as the most abundant genus. *Cladosporium* was recovered as the most abundant genus also by Banchi et al. (2018a) in different urban areas in Italy, followed by *Alternaria*, *Botrytis* and *Periconia*.

Nicolaisen et al. (2017) investigated both urban (Denmark, Netherland and UK) and natural (oilseed rape fields, UK) areas in Northwest Europe in spring and fall, with focus on plant pathogens. Seasonality had a significant effect on species richness (higher in fall), and samples taken above the fields had in general lower richness than those taken in urban areas. Here, the proportion of Ascomycota (mainly Dothideomycetes) prevailed in spring and that of Basidiomycota in autumn, mainly because of the high abundance of Agaricomycetes. Also in this study, the most abundant genus was *Cladosporium* in urban areas, whereas *Entyloma*, a genus including plant pathogenic smut fungi, was very abundant in rural site. In urban sites, located up to 900 km far from each other, but within the same climatic region, distance had no significant effects on observed fungal species richness.

Molecular identification of outdoor fungi was investigated in natural environment also by Abrego et al. (2018) in different geographical areas in Finland. The aim was to generate a fungal species list occurring in the region of interest. Here, fungal communities were richer in Basidiomycota than in Ascomycota. This represents the opposite trend of that is usually found in urban environments, where pollution and the lack of plant debris seem to favor the presence of some Ascomycota species (Abrego et al. 2018). Notwithstanding this, at genus level, the most abundant taxon was the ascomycete *Cladosporium*, followed by *Fusarium*.

4 Conclusion and future perspectives

Despite its importance, air is still a largely unknown environment. Understanding the complexity of its biodiversity and the high variation of airborne organisms represents a challenge for molecular aerobiologists.

Airborne plant and fungal DNA metabarcoding is a powerful, flexible and reliable method for biodiversity estimation in a multiplicity of environmental settings. New inputs for DNA metabarcoding will derive also from the more advanced single molecule sequencing systems using fluorescence (e.g., Pacific Biosciences) or nanopores (Oxford Nanopore). These systems have the advantage of massive parallel sequencing but do not require amplification processes (thus avoiding its biases), offering low costs, high accuracy and longer reads than the HTS approaches (Thompson and Milos 2011). However, to exploit the potentialities of DNA metabarcoding in aerobiology and compare results from different sources, standardized procedures are needed (Núñez et al. 2016). As the application of DNA metabarcoding is at its beginning, an agreement among different research groups should envision the setup of a common strategy and validation of standard protocols. The establishment of a network which, based on the research aims, could propose guidelines at international level (e.g., setup of sampling, bioinformatic pipelines, comparison of the data) is here encouraged. Some ongoing projects, such as POLLERGEN (http://pollergen.bangor.ac.uk/) and POLSNIF (http://www.dnabiomonitoring.nl/air-sampling) for pollen and Global Spore Sampling Project (GSSP; https://www.helsinki.fi/en/researchgroups/spatial-food-web-ecology/research/gssp) for fungi, represent promising, starting points for this purpose. Also, the integration of DNA metabarcoding with other complementary approaches will definitely improve the quality and the success of aerobiological researches.

Acknowledgements The research was funded by the project Finanziamenti di Ateneo per progetti di Ricerca scientifica (FRA2016) of the University of Trieste. We thank Valentina Torboli for designing Fig. 1.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

Abrego, N., Norros, V., Halme, P., Somervuo, P., Ali-Kovero, H., & Ovaskainen, O. (2018). Give me a sample of air and I will tell which species are found from your region: Molecular identification of fungi from airborne spore samples. *Molecular Ecology Resources, 18*(3), 511–524. https://doi.org/10.1111/1755-0998.12755.

Adams, R. I., Bhangar, S., Pasut, W., Arens, E. A., Taylor, J. W., Lindow, S. E., et al. (2015). Chamber bioaerosol study: Outdoor air and human occupants as sources of indoor airborne microbes. *PLoS ONE, 10*(5), 1–18. https://doi.org/10.1371/journal.pone.0128022.

Adams, R. I., Miletto, M., Taylor, J. W., & Bruns, T. D. (2013). Dispersal in microbes: Fungi in indoor air are dominated by outdoor air and show dispersal limitation at short distances. *ISME Journal, 7*(7), 1262–1273. https://doi.org/10.1038/ismej.2013.28.

Aguayo, J., Fourrier-Jeandel, C., Husson, C., & Ioos, R. (2018). Assessment of passive traps combined with high-throughput sequencing to study airborne fungal communities. *Applied and Environmental Microbiology, 84*(11), e02637-17. https://doi.org/10.1128/aem.02637-17.

An, C., Woo, C., & Yamamoto, N. (2018). Introducing DNA-based methods to compare fungal microbiota and concentrations in indoor, outdoor, and personal air. *Aerobiologia, 34*(1), 1–12. https://doi.org/10.1007/s10453-017-9490-6.

Ankenbrand, M. J., Keller, A., Wolf, M., Schultz, J., & Förster, F. (2015). ITS2 database V: Twice as much. *Molecular Biology and Evolution, 32*(11), 3030–3032. https://doi.org/10.1093/molbev/msv174.

Azhagiri, A. K., & Maliga, P. (2007). Exceptional paternal inheritance of plastids in Arabidopsis suggests that low-frequency leakage of plastids via pollen may be universal in plants. *Plant Journal, 52*(5), 817–823. https://doi.org/10.1111/j.1365-313X.2007.03278.x.

Bakker, M. G. (2018). A fungal mock community control for amplicon sequencing experiments. *Molecular Ecology Resources, 18*(3), 541–556. https://doi.org/10.1111/1755-0998.12760.

Bálint, M., Bahram, M., Eren, A. M., Faust, K., Fuhrman, J. A., Lindahl, B., et al. (2016). Millions of reads, thousands of taxa: Microbial community structure and associations analyzed via marker genesa. *FEMS Microbiology Reviews, 40*(5), 686–700. https://doi.org/10.1093/femsre/fuw017.

Banchi, E., Ametrano, C. G., Stanković, D., Verardo, P., Moretti, O., Gabrielli, F., et al. (2018a). DNA metabarcoding uncovers fungal diversity of mixed airborne samples in Italy. *PLoS ONE, 13*(3), e0194489. https://doi.org/10.1371/journal.pone.0194489.

Banchi, E., Stankovic, D., Fernández-Mendoza, F., Gionechetti, F., Pallavicini, A., & Muggia, L. (2018b). ITS2 metabarcoding analysis complements lichen mycobiome diversity data. *Mycological Progress, 17*(9), 1049–1066. https://doi.org/10.1007/s11557-018-1415-4.

Behzad, H., Gojobori, T., & Mineta, K. (2015). Challenges and opportunities of airborne metagenomics. *Genome Biology and Evolution, 7*(5), 1216–1226. https://doi.org/10.1093/gbe/evv064.

Bell, K. L., Burgess, K. S., Okamoto, K. C., Aranda, R., & Brosi, B. J. (2016). Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic Science International: Genetics, 21,* 110–116. https://doi.org/10.1016/j.fsigen.2015.12.010.

Bell, K. L., Loeffler, V. M., & Brosi, B. J. (2017). An rbcL reference library to aid in the identification of plant species mixtures by DNA metabarcoding. *Applications in plant sciences, 5*(3), 1600110. https://doi.org/10.3732/apps.1600110.

Brennan, G. L., Potter, C., de Vere, N., Griffith, G. W., Skjoth, C. A., Osborne, N. J., et al. (2018). Temperate grass allergy season defined by spatio-temporal shifts in airborne pollen communities. *bioRxiv*, 410829.

Brooks, J. P., Edwards, D. J., Harwich, M. D., Rivera, M. C., Fettweis, J. M., Serrano, M. G., et al. (2015). The truth about metagenomics: Quantifying and counteracting bias in 16S rRNA studies Ecological and evolutionary microbiology. *BMC Microbiology, 15*(1), 1–14. https://doi.org/10.1186/s12866-015-0351-6.

Buchheim, M. A., Keller, A., Koetschan, C., Förster, F., Merget, B., & Wolf, M. (2011). Internal transcribed spacer 2 (nu ITS2 rRNA) sequence-structure phylogenetics: Towards an automated reconstruction of the green algal tree of life. *PLoS ONE, 6*(2), e16931. https://doi.org/10.1371/journal.pone.0016931.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods, 7*(5), 335.

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2016). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular Ecology Resources, 16*(1), 138–149. https://doi.org/10.1111/1755-0998.12438.

Coombs, K., Taft, D., Ward, D. V., Green, B. J., Chew, G. L., Shamsaei, B., et al. (2018). Variability of indoor fungal microbiome of green and non-green low-income homes in Cincinnati, Ohio. *Science of the Total Environment, 610–611,* 212–218. https://doi.org/10.1016/j.scitotenv.2017.07.274.

Cui, L., Morris, A., & Ghedin, E. (2013). The human mycobiome in health and disease. *Genome Medicine, 5*(7), 1–12. https://doi.org/10.1186/gm467.

D'Amato, G. (2002). Environmental urban factors (air pollution and allergens) and the rising trends in allergic respiratory diseases. *Allergy, 57*(Suppl 7), 30–33. https://doi.org/10.1034/j.1398-9995.57.s72.5.x.

D'Amore, R., Ijaz, U. Z., Schirmer, M., Kenny, J. G., Gregory, R., Darby, A. C., et al. (2016). A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics, 17*(1), 55. https://doi.org/10.1186/s12864-015-2194-9.

Dannemiller, K. C., Lang-Yona, N., Yamamoto, N., Rudich, Y., & Peccia, J. (2014). Combining real-time PCR and next-generation DNA sequencing to provide quantitative comparisons of fungal aerosol populations. *Atmospheric Environment, 84,* 113–121. https://doi.org/10.1016/j.atmosenv.2013.11.036.

Degois, J., Clerc, F., Simon, X., Bontemps, C., Leblond, P., & Duquenne, P. (2017). First metagenomic survey of the

microbial diversity in bioaerosols emitted in waste sorting plants. *Annals of Work Exposures and Health, 61*(9), 1076–1086. https://doi.org/10.1093/annweh/wxx075.

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., et al. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology, 26*(21), 5872–5895. https://doi.org/10.1111/mec.14350.

Du, P., Du, R., Lu, Z., Ren, W., & Fu, P. (2018a). Variation of bacterial and fungal community structures in PM2.5 collected during the 2014 APEC summit periods. *Aerosol and Air Quality Research, 18*(2), 444–455. https://doi.org/10.4209/aaqr.2017.07.0238.

Du, P., Du, R., Ren, W., Lu, Z., & Fu, P. (2018b). Seasonal variation characteristic of inhalable microbial communities in PM2.5in Beijing city, China. *Science of the Total Environment, 610–611,* 308–315. https://doi.org/10.1016/j.scitotenv.2017.07.097.

Fahlgren, C., Bratbak, G., Sandaa, R. A., Thyrhaug, R., & Zweifel, U. L. (2011). Diversity of airborne bacteria in samples collected using different devices for aerosol collection. *Aerobiologia, 27*(2), 107–120. https://doi.org/10.1007/s10453-010-9181-z.

Frøslev, T. G., Kjøller, R., Bruun, H. H., Ejrnæs, R., Brunbjerg, A. K., Pietroni, C., et al. (2017). Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nature Communications, 8*(1), 1188. https://doi.org/10.1038/s41467-017-01312-x.

Galimberti, A., De Mattia, F., Bruni, I., Scaccabarozzi, D., Sandionigi, A., Barbuto, M., et al. (2014). A DNA barcoding approach to characterize pollen collected by honeybees. *PLoS ONE, 9*(10), e109363. https://doi.org/10.1371/journal.pone.0109363.

Golob, J. L., Margolis, E., Hoffman, N. G., & Fredricks, D. N. (2017). Evaluating the accuracy of amplicon-based microbiome computational pipelines on simulated human gut microbial communities. *BMC Bioinformatics, 18*(1), 1–12. https://doi.org/10.1186/s12859-017-1690-0.

Jaenicke, R. (2005). Abundance of cellular material and proteins in the atmosphere. *Science, 308*(5718), 73. https://doi.org/10.1126/science.1106335.

Klimek-Ochab, M., Brzezińska-Rodak, M., Zymańczyk-Duda, E., Lejczak, B., & Kafarski, P. (2011). Comparative study of fungal cell disruption-scope and limitations of the methods. *Folia Microbiologica, 56*(5), 469–475. https://doi.org/10.1007/s12223-011-0069-2.

Kõljalg, U., Nilsson, R. H., Abarenkov, K., Tedersoo, L., Taylor, A. F. S., & Bahram, M. (2013). Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology, 22,* 5271–5277. https://doi.org/10.1111/mec.12481.

Korpelainen, H., & Pietiläinen, M. (2017). Biodiversity of pollen in indoor air samples as revealed by DNA metabarcoding. *Nordic Journal of Botany, 35*(5), 602–608. https://doi.org/10.1111/njb.01623.

Korpelainen, H., & Pietilainen, M. (2017). Diversity of indoor fungi as revealed by DNA metabarcoding. *Genome, 27*(Latokartanonkaari 5), gen-2015-0191. https://doi.org/10.1139/gen-2015-0191.

Korpelainen, H., Pietiläinen, M., & Huotari, T. (2016). Effective detection of indoor fungi by metabarcoding. *Annals of Microbiology, 66*(1), 495–498. https://doi.org/10.1007/s13213-015-1118-x.

Kraaijeveld, K., de Weger, L. A. L. A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P. S. P. S., et al. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Molecular Ecology Resources, 15*(1), 8–16. https://doi.org/10.1111/1755-0998.12288.

Kress, W. J., & Erickson, D. L. (2008). DNA barcodes: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences, 105*(8), 2761–2762. https://doi.org/10.1073/pnas.0800476105.

Kumari, P., Woo, C., Yamamoto, N., & Choi, H. L. (2016). Variations in abundance, diversity and community composition of airborne fungi in swine houses across seasons. *Scientific Reports, 6*(November), 1–11. https://doi.org/10.1038/srep37929.

Lacey, M. E., & West, J. S. (2006). *The Air Spora.* Berlin: Springer. https://doi.org/10.1007/s13398-014-0173-7.2.

Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H., Buckley, T., et al. (2018). Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *New Zealand Journal of Ecology, 42,* 10–50. https://doi.org/10.20417/nzjecol.42.9.

Lemons, A. R., Hogan, M. B., Gault, R. A., Holland, K., Sobek, E., Olsen-Wilson, K. A., et al. (2017). Microbial rRNA sequencing analysis of evaporative cooler indoor environments located in the Great Basin Desert region of the United States. *Environmental Science: Processes and Impacts, 19*(2), 101–110. https://doi.org/10.1039/c6em00413j.

Lindahl, B. D., Nilsson, R. H., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjøller, R., et al. (2013). Fungal community analysis by high-throughput sequencing of amplified markers—A user's guide. *The New phytologist, 199*(1), 288–299. https://doi.org/10.1111/nph.12243.

Luhung, I., Wu, Y., Xu, S., Yamamoto, N., Chang, V. W. C., & Nazaroff, W. W. (2018). Exploring temporal patterns of bacterial and fungal DNA accumulation on a ventilation system filter for a Singapore university library. *PLoS ONE, 13*(7), 1–19. https://doi.org/10.1371/journal.pone.0200820.

Luongo, J. C., Barberán, A., Hacker-Cary, R., Morgan, E. E., Miller, S. L., & Fierer, N. (2017). Microbial analyses of airborne dust collected from dormitory rooms predict the sex of occupants. *Indoor Air, 27*(2), 338–344. https://doi.org/10.1111/ina.12302.

Mahé, F., Mayor, J., Bunge, J., Chi, J., Siemensmeyer, T., Stoeck, T., et al. (2015). Comparing high-throughput platforms for sequencing the V4 region of SSU-rDNA in environmental microbial eukaryotic diversity surveys. *Journal of Eukaryotic Microbiology, 62*(3), 338–345. https://doi.org/10.1111/jeu.12187.

Nguyen, N., Warnow, T., Pop, M., & White, B. (2016). PERSPECTIVE A perspective on 16S rRNA operational taxonomic unit clustering using sequence similarity (August 2015). *NPJ Biofilms and Microbiomes, 2,* 16004. https://doi.org/10.1038/npjbio.

Nicolaisen, M., West, J. S., Sapkota, R., Canning, G. G. M., Schoen, C., & Justesen, A. F. (2017). Fungal communities including plant pathogens in near surface air are similar

across northwestern Europe. *Frontiers in Microbiology, 8,* 1729. https://doi.org/10.3389/fmicb.2017.01729.

Nilsson, R. H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M., et al. (2015). A comprehensive, automatically updated fungal ITS sequence dataset for reference-based chimera control in environmental sequencing efforts. *Microbes and Environments, 30*(2), 145–150. https://doi.org/10.1264/jsme2.ME14121.

Núñez, A., de Paz, G. A., Rastrojo, A., García, A. M., Alcamí, A., Montserrat Gutiérrez-Bustillo, A., et al. (2016). Monitoring of airborne biological particles in outdoor atmosphere. Part 2: Metagenomics applied to urban environments. *International Microbiology, 19*(2), 69–80. https://doi.org/10.2436/20.1501.01.265.

Oh, S. Y., Fong, J. J., Park, M. S., Chang, L., & Lim, Y. W. (2014). Identifying airborne fungi in Seoul, Korea using metagenomics. *Journal of Microbiology, 52*(6), 465–472. https://doi.org/10.1007/s12275-014-3550-1.

Olesen, S. W., Duvallet, C., & Alm, E. J. (2017). dbOTU3: A new implementation of distribution-based OTU calling. *PLoS ONE, 12*(5), e0176335.

Plummer, E., & Twin, J. (2015). A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. *Journal of Proteomics & Bioinformatics, 8*(12), 283–291. https://doi.org/10.4172/jpb.1000381.

Prussin, A. J., Marr, L. C., & Bibby, K. J. (2014). Challenges of studying viral aerosol metagenomics and communities in comparison with bacterial and fungal aerosols. *FEMS Microbiology Letters, 357*(1), 1–9. https://doi.org/10.1111/1574-6968.12487.

Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The barcode of life data system (http://www.barcodinglife.org). *Molecular Ecology Notes*, *7*(3), 355–364. https://doi.org/10.1111/j.1471-8286.2007.01678.x.

Reuter, J. A., Spacek, D. V., & Snyder, M. P. (2016). High-throughput sequencing technologies. *Molecular Cell, 58*(4), 586–597. https://doi.org/10.1016/j.molcel.2015.05.004.High-Throughput.

Richardson, R. T., Lin, C.-H., Quijia, J. O., Riusech, N. S., Goodell, K., & Johnson, R. M. (2015). Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Applications in Plant Sciences, 3*(11), 1500043. https://doi.org/10.3732/apps.1500043.

Rocchi, S., Valot, B., Reboux, G., & Millon, L. (2017). DNA metabarcoding to assess indoor fungal communities: Electrostatic dust collectors and Illumina sequencing. *Journal of Microbiological Methods, 139*(January), 107–112. https://doi.org/10.1016/j.mimet.2017.05.014.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology, 75*(23), 7537–7541. https://doi.org/10.1128/AEM.01541-09.

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., et al. (2012). Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America, 109*(16), 1–6. https://doi.org/10.1073/pnas.1117018109.

Sherwood, A. R., Dittbern, M. N., Johnston, E. T., & Conklin, K. Y. (2017). A metabarcoding comparison of windward and leeward airborne algal diversity across the Ko‘olau mountain range on the island of O’ahu, Hawai‘i[1]. *Journal of Phycology, 53*(2), 437–445. https://doi.org/10.1111/jpy.12502.

Sherwood, A. R., & Presting, G. G. (2007). Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. *Journal of Phycology, 43*(3), 605–608. https://doi.org/10.1111/j.1529-8817.2007.00341.x.

Shin, S. K., Kim, J., Ha, S. M., Oh, H. S., Chun, J., Sohn, J., et al. (2015). Metagenomic insights into the bioaerosols in the indoor and outdoor environments of childcare facilities. *PLoS ONE, 10*(5), 1–17. https://doi.org/10.1371/journal.pone.0126960.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. (2012). Environmental DNA. *Molecular Ecology, 21*(8), 1789–1793. https://doi.org/10.1111/j.1365-294X.2012.05542.x.

Tang, L. Y., Nagata, N., Matsushima, R., Chen, Y., Yoshioka, Y., & Sakamoto, W. (2009). Visualization of plastids in pollen grains: Involvement of FtsZ1 in pollen plastid division. *Plant and Cell Physiology, 50*(4), 904–908. https://doi.org/10.1093/pcp/pcp042.

Tedersoo, L., Tooming-Klunderud, A., & Anslan, S. (2018). PacBio metabarcoding of Fungi and other eukaryotes: Errors, biases and perspectives. *New Phytologist, 217*(3), 1370–1385. https://doi.org/10.1111/nph.14776.

Tessler, M., Neumann, J. S., Afshinnekoo, E., Pineda, M., Hersch, R., Velho, L. F. M., et al. (2017). Large-scale differences in microbial biodiversity discovery between 16S amplicon and shotgun sequencing. *Scientific Reports, 7*(1), 1–14. https://doi.org/10.1038/s41598-017-06665-3.

Thompson, J. F., & Milos, P. M. (2011). The properties and applications of single-molecule DNA sequencing. *Genome Biology, 12*(2), 217. https://doi.org/10.1186/gb-2011-12-2-217.

Tong, X., Leung, M. H. Y., Wilkins, D., & Lee, P. K. H. (2017). City-scale distribution and dispersal routes of mycobiome in residences. *Microbiome, 5*(1), 1–13. https://doi.org/10.1186/s40168-017-0346-7.

Woo, C., An, C., Xu, S., Yi, S. M., & Yamamoto, N. (2018). Taxonomic diversity of fungi deposited from the atmosphere. *ISME Journal, 12*(8), 2051–2060. https://doi.org/10.1038/s41396-018-0160-7.

Yamamoto, N., Bibby, K., Qian, J., Hospodsky, D., Rismani-Yazdi, H., Nazaroff, W. W., et al. (2012). Particle-size distributions and seasonal diversity of allergenic and pathogenic fungi in outdoor air. *The ISME Journal, 6*(10), 1801–1811. https://doi.org/10.1038/ismej.2012.30.

Yan, D., Zhang, T., Su, J., Zhao, L. L., Wang, H., Fang, X. M., et al. (2016). Diversity and composition of airborne fungal community associated with particulate matters in Beijing during haze and non-haze days. *Frontiers in Microbiology, 7,* 1–12. https://doi.org/10.3389/fmicb.2016.00487.

Yang, W., Guo, M., Liu, G., Yu, G., Wang, P., Wang, H., et al. (2018). Detection and analysis of fine particulate matter and microbial aerosol in chicken houses in Shandong

Province. *China. Poultry Science, 97*(3), 995–1005. https://doi.org/10.3382/ps/pex388.

Yoo, K., Lee, T. K., Choi, E. J., Yang, J., Shukla, S. K., Hwang, S. I., et al. (2017). Molecular approaches for the detection and monitoring of microbial communities in bioaerosols: A review. *Journal of Environmental Sciences (China), 51,* 234–247. https://doi.org/10.1016/j.jes.2016.07.002.