Environmental DNA assessment of airborne plant and fungal seasonal diversity

Elisa Banchi[1,2], Claudio G. Ametrano[1], Enrico Tordoni[1], David Stanković[1,3], Silvia Ongaro[1], ARPA working group[4], Mauro Tretiach[1], Alberto Pallavicini[1,2*], Lucia Muggia[1*]

[1] Department of Life Sciences, University of Trieste, via Giorgieri 10, I-34127 Trieste, Italy

[2] National Institute of Oceanography and Applied Geophysics - OGS, via Piccard 54, I-34151, Trieste, Italy

[3] Marine Biology Station, National Institute of Biology, Fornače 41, SLO-6330 Piran, Slovenia

[4]Regional Agency for Environmental Protection, Italy: Pierluigi Verardo and Francesca Tassan (ARPA Friuli Venezia Giulia), Nadia Trobiani (ARPA Marche), Olga Moretti (ARPA Umbria), Maria Francesca Borney (ARPA Valle d'Aosta), Stefania Lazzarin (ARPA Veneto)

* Corresponding authors:

Lucia Muggia, Department of Life Sciences, University of Trieste, via Giorgieri 10, I-34127 Trieste, Italy; lmuggia@units.it

Alberto Pallavicini, Department of Life Sciences, University of Trieste, via Giorgieri 4, I-34127 Trieste, Italy; pallavic@units.it

Running title: eDNA metabarcoding of airborne plants and fungi

Abstract

Environmental DNA (eDNA) metabarcoding and metagenomics analyses can improve taxonomic resolution in biodiversity studies. Only recently, these techniques have been applied in aerobiology, to target bacteria, fungi and plants in airborne samples. Here, we present a nine-month aerobiological study applying eDNA metabarcoding in which we analyzed simultaneously airborne diversity and variation of fungi and plants across five locations in North and Central Italy. We correlated species composition with the ecological characteristics of the sites and the seasons. The most abundant taxa among all sites and seasons were the fungal genera *Alternaria*, *Cladosporium*, and *Epicoccum* and the plant genera *Brassica*, *Corylus*, *Cupressus* and *Linum*, the latter being much more variable among sites. PERMANOVA and indicator species analyses showed that the plant diversity from air samples is significantly correlated with seasons, while that of fungi varied according to the interaction between seasons and sites. The results consolidate the performance of a new eDNA metabarcoding pipeline for the simultaneous amplification and analysis of airborne plant and fungal particles. They also highlight the promising complementarity of this approach with more traditional biomonitoring frameworks and routine reports of air quality provided by environmental agencies.

Keywords: Aerobiology, High Throughput Sequencing, Italy, PERMANOVA, PLANiTS, Pollen, Spore.

Introduction

Biodiversity assessments are a central focus in multiple fields of research, ranging from estimating the gain and/or loss of species due to ecological factors to the detection of invasive

(alien) species, contaminants in food products or clinical allergens and pathogens. In the past years, molecular techniques, such as environmental DNA (eDNA, Taberlet et al., 2012) metabarcoding (targeted amplicon parallel sequencing) have facilitated increased taxonomic resolution in complex communities (Deiner et al., 2017). The analysis of eDNA has revolutionized biodiversity assessments, allowing the detection of elusive or rare species (Deiner et al., 2017; Ruppert et al., 2019). Although eDNA metabarcoding is still a semi-quantitative method (Degois et al., 2017; Holdaway et al., 2017), it has found broad applications in ecology, mainly concerning conservation, biomonitoring and ecosystem assessments (Thomsen & Willerslev, 2015). However, unique characteristics of sample sites and environmental features, in addition to the targeted species/organismal groups, often require study-specific optimization to ensure an appropriate workflow, starting from sampling design to data analysis. Most studies focus on distinct organismal groups individually, e.g., bacteria, plants, rather than a comprehensive perspective of all microbial diversity (Banchi et al., 2019), despite the fact that different components of communities all provide valuable insight in ecological/environmental or clinical/medical contexts.

The vast majority of ecological studies complemented by eDNA metabarcoding analyses have been performed in aquatic and terrestrial environments (Behzad et al., 2015). Recently, airborne particles from diverse habitats have been surveyed by molecular approaches, as reviewed by Banchi et al. (2019). Traditionally, aerobiology, *i.e.* the study of airborne biological particles (Lacey & West, 2006), relied heavily on microscopy to assess species diversity and abundance. Recent advancements in high-throughput sequencing and eDNA metabarcoding technologies have provided increased taxonomic resolution of airborne biodiversity, including pollen (Kraaijeveld et al., 2015), bacteria (García-Mena et al., 2016) and fungi (Banchi et al., 2018). eDNA metabarcoding is a culture-independent approach, and

allows the detection also of those organisms which may not grow in culture, are dead or dormant (An et al., 2018); and it further offers, therefore, the possibility to monitor, estimate, compare and taxonomically assign taxa without the need for specialized expertise at morphological and systematic classification levels. However, studies still suffer from the lack of standardized sampling methods, several biases that can be generated during sampling, DNA extractions and amplifications, the choice of the target DNA barcode(s), different sequencing platforms, and the lack of reference sequences in public databases (Banchi et al., 2019). In this context, the use of mock communities (samples composed by an *ad hoc* DNA mix of known organisms) as positive controls are crucial 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). The use of mock communities, though, has been reported in about 10% of the studies dealing with eDNA metabarcoding on airborne plants and fungi (Banchi et al., 2019). While bacterial mock communities are commercialized (Lear et al., 2018), few mock communities for fungi have been proposed for standardization (Bakker, 2018; Egan et al., 2018; McTaggart et al., 2019).To the best of our knowledge, there are no standardized mock communities already established for plant and fungi together. Moreover, while studies on bacterial diversity in airborne sample are relatively common, much less has been reported for plants (including higher plants - Streptophytes, and algae - Chlorophytes) and fungi (sometimes co-analyzed with bacteria); even rarer are the studies which report on multiple groups of organisms (Degois et al., 2017 on Eukarya).

To date, few studies have investigated the diversity of airborne pollen (Brennan et al., 2019; Korpelainen & Pietiläinen, 2017; Kraaijeveld et al., 2015) and algae (Sherwood et al., 2017) through eDNA metabarcoding. Among these, only Kraaijeveld et al. (2015) provided

an accurate, qualitative and quantitative taxonomic assessment of pollen diversity in the Netherlands, highlighting the potential of this technique to increase the efficiency of pollen monitoring. Kraaijeveld et al. (2015) stressed the importance of this approach to detect allergenic species with finer resolution, such as for grass taxa, which are particularly challenging when only morphological traits are used for their identification. Brennan et al. (2019) also used eDNA metabarcoding to detect seasonal changes in the spatial and temporal distribution of Poaceae pollen throughout the allergy season in the United Kingdom and highlighted how socio-economic advantages might arise from a better knowledge on taxon-specific exposure of pollen and allergic diseases.

In fungal ecology, eDNA metabarcoding has improved our understanding of the diversity of airborne fungi. Fungi are indeed more difficult to survey than plants for multiple reasons. Fungal communities show high spatial variation - even at fine scales (Kubartova et al., 2012; Nacke et al., 2016). Ephemeral environmental conditions, such as humidity, coupled with the reproductive phenology of fungal species influences their occurrence – being seasonal or even depending on the time of the day (Elbert et al., 2007; Kramer, 1982; Pashley et al., 2012). Fungal spores can rapidly be transported over long distances - likely causing fast homogenization of the local atmospheric fungal communities (Abrego et al., 2018; Norros et al., 2014; Rieux et al., 2014).

Despite the improvements in sequencing technologies and molecular-based taxonomic research in the past decades (Cheng et al., 2016; Nilsson et al., 2015; Schoch et al., 2012), knowledge about the global biodiversity of airborne pollen and fungal particles remains limited. Robust, accurate, and consistent taxonomic classification of samples are challenging in molecular-based biodiversity assessments due to the lack of comprehensive, curated DNA reference libraries. In particular, the taxonomic discrimination to the species level is rarely

accomplished, and clustering approaches based on sequence similarity have no direct taxonomic comparison (Núñez et al., 2017). Finding solutions to these drawbacks would be particularly important concerning the detection of allergens and pathogenic agents.

Recent studies support the use of eDNA metabarcoding as a promising approach to increase the quality and sensitivity of aerobiological monitoring. Banchi et al. (2018) assessed the taxonomic composition and diversity of airborne fungi in mixed airborne samples gathered during two weeks in four sites of North-Eastern and Central Italy and compared the molecular data with those obtained by traditional microscopy. In that study, the number of fungal taxa identified with DNA metabarcoding was ten-times higher than the number of taxa identified by using traditional microscopy analyses.

Here, we present the first aerobiological study in which both plant and fungal diversity are simultaneously assessed from airborne samples during an extended period of time. Our specific research aims were to (*i*) assess the airborne diversity of plants and fungi considering their spatio-temporal variation and (*ii*) consolidate and improve the pipeline of Banchi et al. (2018) to analyze mixed airborne samples of plants and fungi. We investigated airborne fungal and plant diversity and its variation in space and time across five localities distributed in North and Central Italy for nine months using high-throughput sequencing and eDNA metabarcoding. To assess reliability and amplification biases, we established mock communities of plants and fungi. The mock communities were used in a trial with six selected samples and amplified with different primer combinations to evaluate whether different primer pairs affect the sequencing results. Our results provide valuable perspective into the suitability of eDNA metabarcoding for biomonitoring and routine analyses of air quality at a broad scales.

Materials and Methods

Sampling

Airborne biological particles were sampled with a volumetric sampler (VPPS 2010, Lanzoni) mounted with sticky tape (Melinex®). The sampling was performed in five Italian sites (Fig. 1) by the collaborators of the units of the Regional Agency for Environmental Protection (ARPA). Air samplers were placed on the roof of the ARPA buildings, located in urban areas, at about 15-20 m from the ground. The general workflow of the study is reported in Fig. 2. The sampling was performed every second week for nine months starting from March to November 2017, including Spring (weeks 1-7), Summer (weeks 8-13) and Autumn (weeks 14-18). The sampling tape was cut according to the sampling days and rolled to fit individually into 1.5 ml tubes; these were sent periodically to the laboratory of the University of Trieste and were stored at room temperature until being processed all together. Each week-long sampling period consisted of five days, from Tuesday to Sunday (for a total of six tape fragments for each sampling week). Mondays were excluded because on each Monday the sampling tape was replaced, and this was done at different time at each sampling station, thus causing inconsistency in the sampling.

Sampling sites were selected to maximize the geographical and climatic diversity over the study area (two sites were selected in North-Eastern Italy, two in Central Italy and one in North-Western Italy; Fig. 1). Sampling sites include: 1) Friuli Venezia Giulia - FVG (Pordenone, 45°57'09.2"N - 12°40'54.2"E, 4 m a.s.l.), 2) Marche (Ascoli Piceno, 42°52'50.0"N -13°42'27.4"E, 154 m a.s.l.), 3) Umbria (Terni, 42°34'48.0"N -12°37'59.4"E, 130 m a.s.l.), 4) Valle d'Aosta - VdA (Aosta - St. Christophe, 45°44'31.3"N -7°21'29.97" E, 619 m a.s.l.), and 5) Veneto (Vicenza, 45°31'49.6"N - 11°35'24.2"E, 39 m a.s.l.).

Climatic data

In general, VdA records a lower average of annual temperatures (3.6 °C) than the other four sites (FVG records 11.8 °C, Veneto 11.6 °C, Marche 13.6 °C and Umbria 12.9 °C, respectively). The average annual precipitation is rather homogeneous among the five sites, measuring 1.065 mm in FVG, 797 mm in Marche, 808 mm in Umbria, 840 mm in VdA and 845 mm in Veneto (ISTAT 2010). According to Köppen climate classification, the climate conditions in FVG, Marche, Umbria and Veneto are classified as humid-subtropical (Cfa), while VdA is classified as cold continental (Dfc) (Kottek, Grieser, Beck, Rudolf, & Rubel, 2006).

During the sample collection, climatic data were retrieved from regional meteorological stations nearest to the sampling sites (100 m away from the sampling point in FVG and Umbria, 1 km in VdA, 6 km in Marche, and 10 km in Veneto). Meteorological data were aggregated into two records per month to be comparable with the sampling protocol used for the aerobiological samples. The averages of the following climatic variables were selected for the analysis: mean daily temperature (*T_ave*, °C), maximum daily temperature (*T_max*, °C), minimum daily temperature (*T_min*, °C), daily rainfall (*Prec_ave*, mm/day), relative humidity (*Ur_ave*, %) and wind speed (*Wind_ave*, km/h). Climatic data were standardized (zero mean, unit variance) before statistical analyses.

Primer design

The simultaneous amplification of both fungal and plant ITS2 was performed using a forward primer coupled with three reverse primers (Fig. 2; Supplementary Materials Fig. 1S). The reverse complement of the primer ITS-u2 (GAAYCATCGARTCTTTGAACGC; Cheng et al., 2016) was used as forward primer and renamed as ITS-u2_F. This primer is located in the

193 5.8S gene and is the closest to the ITS2 region. The selected three reverse primers were: *i)*
194 ITS-p4 (CCGCTTAKTGATATGCTTAAA; Cheng et al., 2016) to promote the amplification
195 of plant DNA; *ii)* ITS-f4 (CCGCTTATTGATATGCTTAAG), here newly designed, to
196 promote the amplification of fungal DNA [ITS-f4 is a slight modification of ITS-p4 of Cheng
197 et al. (2016) to better complement to ascomycetes and basidiomycetes ITS2 sequences;
198 Supplementary Materials Fig. 1S]; *iii)* ITS4U_R (TCCTCCGCTTAKTGATATGC) to
199 promote the amplification of both plant and fungal DNA at the same time, it was designed
200 starting from the forward primer ITS4F (White et al.,1990).
201
202 *Preparation of mock communities*
203 Two mock communities were assembled with plants and fungal samples (Table 1) to test and
204 ensure the reliability of the laboratory workflow and to evaluate to which extent different
205 amounts of starting material may affect the representativeness of different taxa. Plant species
206 were chosen among those growing in the Botanic Garden of the University of Trieste, and
207 both angiosperms and gymnosperms were included. Pollen samples were collected directly
208 from identified plants in season. We took care to include the pollen of *Corylus avellana*, as it
209 is a well-known allergenic plant. Chlorophyta were chosen among the algal cultures stored at
210 the University of Trieste. Fungal material included both Ascomycota and Basidiomycota and
211 was retrieved either from spores collected in the Botanic Garden or mycelia of fungal cultures
212 available at the laboratory of the University of Trieste. *Cladosporium* and *Alternaria* spores,
213 which are among the most commonly recovered airborne fungi, were passively sampled on
214 agar plates, isolated and cultured by the cooperation partner ARPA Valle d'Aosta, while
215 spores of *Erysiphe necator* were recovered from infected leaves of *Vitis vinifera*. *Erysiphe*
216 *necator* was included because it was detected by microscopy analyses but not highlighted

among the sequenced fungi (Banchi et al. 2018). The addition of *E. necator* to the present mock community aimed at verifying whether this fungus is or not amplifiable and detectable by DNA analyses. The identity of each taxon in the mock communities was confirmed by sequencing the ITS2 region (see methods below).

DNA was extracted from each sample used in the mock communities using the ZR Fungal/Bacterial DNA MicroPrep™ Kit (Zymo Research) and quantified with Qubit™ Fluorimeter (Thermo Fisher Scientific). The ITS2 region was amplified with the forward primer ITSu3 and the reverse primer ITSu4 (Cheng et al., 2016). The PCR reaction mix contained 3 µl DNA template (10–20 ng), 1× Taq Buffer A (Kapa Biosystems), 1 U Taq DNA Polymerase (Kapa Biosystems), 200 µM dNTPs, 400 nM of each primer in a final volume of 50 µl. The PCR amplifications were performed with the following cycling profile: 95 °C for 3 min and 35 cycles at 95 °C for 30 s, 55 °C for 30 s, 72 °C for 1 min followed by a final extension at 72 °C for 1 min. A negative control ('no template control') was used to verify the absence of non-specific amplification products along with the whole amplification and sequencing process. Sanger sequencing of PCR products was performed with an ABI 3130 Genetic Analyzer (Thermo Fisher Scientific). DNA amplifications were then pooled in two mock communities: one in which the samples were mixed in even amount (mock E), the other in which the samples were serially diluted (mock D) into different concentrations (Table 1). The mock D was prepared in this way to include also the DNA of *Alternaria* and *Erysiphe necator*, which were extracted and amplified in a too low amount and had to be left out from mock E. The different concentrations of the fungal and plant taxa were proportionally calculated starting from the amount of the extracted DNA.

DNA extraction of aerobiological samples

DNA extractions from the sampling tape were performed individually for each of the six fragments of the sampling weeks using ZR Fungal/Bacterial DNA MicroPrep™ Kit (Zymo Research) following Banchi et al. (2018). After extraction, the six DNA samples were pooled in equal amount into a single week sample and further processed for DNA metabarcoding analysis. The whole experiment combined a total of 18-week samples for each site; each week was treated further in the analyses as an individual "sample", resulting in a total of 90 samples. DNA extraction was also performed on six not-exposed tapes (treated as blanks) to assess possible contaminations due to samples processing. These extractions were used as additional negative controls during the subsequent PCR amplifications.

Library preparation

Amplicons for HTS sequencing of the airborne samples and the mock communities were obtained with two sequentially PCR amplifications following Banchi et al. (2018): the first, primary PCR, amplifies the target sequence ("I PCR" in Fig. 2); the second, outer PCR, is performed to attach the molecular identifiers (MID) for multiplex sequencing to the PCR products ("II PCR" in Fig. 2). For each sample, the two PCR amplifications were performed independently for either of the three primer pairs; the three reactions were pooled after the outer PCR, to ensure that the three PCR products were successfully amplified.

The reaction mix of the primary PCR contained 2 µl DNA template (~10 ng), 8 µl SSOAdvanced™ SYBR® Green Supermix (Bio-Rad) and 200 nM of forward and reverse primers in a final volume of 15 µl. The PCR amplifications were performed in a CFX 96™ PCR System (Bio-Rad) with the following cycling profile: 98 °C for 30 seconds and 35 cycles at 95 °C for 10 seconds and 60 °C for 20 seconds. The reaction mix of the outer PCR contained 0.5 µl of the first PCR product, 8 µL SSOAdvanced™ SYBR® Green Supermix

(Bio-Rad) and 200 nM of each primer (10 µM) in a final volume of 15 µl. This cocktail was processed for 12 PCR cycles with the previous amplification conditions. Amplicons were checked for their quality and length by agarose gel electrophoresis, and then pooled to obtain 90 samples and two mock communities (E and D).

Comparison of primer pairs

We compared three different approaches on six samples randomly selected among each site, to evaluate whether different primer combinations affect the sequencing results (Fig. 2). The PCR reactions were set as follow: (*i*) primary and outer PCRs were performed as described in previous section ("individual"); (*ii*) the primary PCR was performed with a mix of all the primers ("mixed": 200 nM of the forward primer and 200/3 nM of each reverse primers); (*iii*) primary and outer PCR were performed separately with each primer pair ("primer pair").

High Throughput Sequencing

All the amplicons obtained were run on a 2% agarose gel from which ~400 bp products were gel extracted, quantified with Qubit™ Fluorimeter (Thermo Fisher Scientific) and pooled in equal amount to prepare two libraries, with 58 and 57 samples, respectively. Both libraries were sequenced with an Ion Torrent Personal Genome Machine (PGM, Thermo Fisher Scientific) on two 316™ chips (Thermo Fisher Scientific).

Plant and fungal ITS2 reference sequence databases

We prepared two independent reference databases for the ITS2 of plants and fungi. Sequences were retrieved from NCBI, and the ITS region was extracted with ITSx (Bengtsson-Palme et al., 2010). The sequences underwent identity check to remove misidentified records and were

clustered at 99% identity with cd-hit (Fu et al., 2012) to reduce redundancy and computational effort. Finally, taxonomy for each sequence was added at phylum, class, order, family, genus and species level with entrez_qiime (Baker, 2016). The plant reference database (PLANiTS2; Banchi et al., 2020) consisted of 699,968 sequences retrieved from NCBI, which after the ITS2 selection and trimming were reduced to 313,175 sequences. The fungal reference database consisted of 826,895 sequences retrieved from NCBI, which after the ITS2 selection and trimming were reduced to 470,759 sequences. The clustered libraries, and therefore the references database, consisted of 96,771 plant (PLANiTS2; Banchi et al., 2020) and 89,814 fungal sequences.

Sequence data analysis

The sequence data generated for this study are available at the NCBI short read repository under the accession number PRJNA576572. The obtained sequences were de-multiplexed, trimmed (from primers and adapters), and quality filtered (minimum length 200 bp, minimum average quality score 20) with CLC Genomics Workbench v.12 (Qiagen). Reference-based Operational Taxonomic Units (OTUs) clustering and taxonomic assignment were performed with Microbial Genomics module in CLC Genomics Workbench v.12 (Qiagen). The merged fungi and plants reference libraries were used as references. The parameters were set as follow: 97% similarity (70% for the new OTUs), minimum occurrence two (to remove singletons). Taxonomic assignment were also performed in QIIME2 (Bolyen et al., 2019) with the alignment-based taxonomy consensus method based on vsearch 2.0.3 (Rognes et al., 2016) applying at 97% of identity and the merged fungi and plants libraries as reference.

eDNA metabarcoding analysis allows the detection of invasive/exotic species, which may represent major threats for diverse ecosystems (Pejchar & Mooney, 2009). In our

analyses we search specifically for alien species taking as reference fungal and plant alien species lists reported in the DAISIE European Invasive Alien Species (IAS) Database (http://www.europe-aliens.org/) and the checklist of the Italian alien vascular flora (Galasso et al., 2018).

Statistical analyses

To assess reliability and robustness of the two mock communities, we first performed correlation analysis (Spearman's ρ) of taxa abundances between mock communities at genus level; the same analysis was performed at phylum level to compare the primer comparison approaches.

Before statistical analyses, sequence data of Streptophyta and Fungi were normalized at genus level to 20000 reads, and the abundance of each taxa was log(x+1) transformed. A Permutational Analysis of Variance (PERMANOVA; Anderson, 2001) was computed to investigate the fixed effect of Site (five levels: FVG; Marche, Umbria, VdA, Veneto), Season (three levels: Spring, Summer, Autumn) and their interaction on community composition of plant and fungi, treating the two groups independently. Latent gradients in plants and fungi were assessed through Nonmetric MultiDimensional Scaling (NMDS) using Bray-Curtis similarity and 50 random starts of the initial configuration.

All tests were performed by using Bray-Curtis similarity, 4999 permutations of residuals under a reduced model, and type III sums of squares; when the factor resulted significant, pairwise comparisons were performed using *t* statistic and 4999 permutations. NMDS and PERMANOVA were performed using PRIMER 6 software (Clarke & Gorley, 2006) and the add-on package PERMANOVA + (Anderson et al., 2008), whereas all the further statistical analyses were computed in R 3.5.3 (R Core Team 2019). Species Indicator

analysis (Dufrene & Legendre, 1997) was calculated to identify the 'fidelity' between sets of taxa and the interaction factor between sites and season using 'indicspecies' R package (De Caceres & Legendre, 2009); the significance of the relationship was assessed through permutations (n = 999).

The effect of climatic variables on community ordination was assessed through redundancy analysis (RDA, Legendre & Legendre, 2012) computed with 'vegan' R package (Oksanen et al., 2019). Furthermore, to assess the importance of climatic variables on community composition on plant and fungi, a forward selection by permutation (999) was adopted for each dataset following the double-stopping criterion proposed in Blanchet et al. (2008) and using 'adespatial' R package (Dray et al., 2019).

Results

Sequencing

A total of 9,014,361 raw reads (4,498,406 for the first and 4,515,955 for the second chip, respectively) were generated. 5,524,343 reads passed the quality filter and had an average length of 265 bp. A total of 3,208,724 reads were retained after CLC Genomics Workbench v.12 (Qiagen), OTUs clustering and the filtering of chimaeras and singletons. These reads represent the final dataset used for the taxonomic assignment and the statistical analyses (Table S1).

Mock communities

Equal amount (E) and serial diluted (D) mock communities were run in duplicate on both chips. The replicates were highly consistent, being the Spearman's correlation between the taxonomic composition of the equal amount (E) and the one of the two serially diluted (D)

mocks highly significant ($p < 0.01$ and $\rho = 0.99$ for both). At the genus level, all the taxa present in the mock communities were also detected in the samples; and *Acer campestre*, *Bjerkandera adusta*, *Calocybe graveolens*, *Corylus avellana*, *Erysiphe necator*, *Phanerochaete chrysosporium*, *Polyporus squamosus* and *Tulipa gesneriana* were identified further up to the species level. The proportion of the reads belonging to each taxon in E and D mock communities are shown in Fig. 3A. In the E mock community, containing the same amount of DNA for each species, some taxa were either over- or under-represented (Table 1, Fig. 3A); the same is also observed in the D mock community, after the read numbers were normalized to the initial DNA amount (Table 1, Fig. 3A). The pattern of most of the taxa remains the same in the two approaches, while *Campanula* sp. and *Wisteria* sp. were more influenced by the initial amount of DNA. Fungal samples originating from spores were in general under-represented.

Primer pair comparison

The taxonomic composition at kingdom and phylum level coming from the different primer combinations summing the reads of the six samples is reported in Fig. 3B. As expected, plants (Streptophyta and Chlorophyta) were amplified most effectively with primers ITS4U_R (60%) and ITS-p4 (~58%), and at the least effectively using primer ITS-f4 and the "mixed" approach (32% for both). Fungi were best amplified with the "mixed" approach (65%) and with primer ITS-f4 (61%) and at the least using primer ITS-p4 (33%). The lowest amount of reads which could not be assigned to any fungal or plant group was recovered using the primer ITS4U_R in the "primer approach". The Spearman's correlation among samples was significant for all the comparisons (Table 2).

Spatio-temporal diversity patterns

Due to technical problems of the sampler in Marche, samples from three weeks (M8, M10, M15) were lost; therefore only 87 total samples could be further analyzed. The CLC Workbench (Qiagen) assigned at kingdom level 62% of the reads to Fungi and 24% to Plants, while 14% remained unassigned. The results with QIIME2 are largely in accordance with those obtained with CLC (Supplementary Material Fig. S2), and here we report the results from the CLC Workbench. Overall, the taxonomic composition at genus level was represented by 158 plant genera (153 Streptophyta and 5 Chlorophyta) and 613 fungal genera (349 Ascomycota, 259 Basidiomycota and 5 Mucoromycotina; Supplementary Materials Table S2).

A total of 43 alien plant species were detected in our dataset (Supplementary Materials Table S3) and are represented by invasive (12 species), naturalized (16) and random (15) species. In particular, the most abundant IAS across all sites was *Robinia pseudoacacia*. The established fungal plant pathogens detected were *Cryptostroma corticale*, *Fusarium proliferatum* and *Colletotrichum gloeosporioides* (Supplementary Materials Table S3).

The most abundant plant genera at site level and across all seasons were: *Corylus*, *Brassica* and *Linum* in FVG and VdA (18%, 11% and 7%, 32%, 17% and 8%, respectively); *Brassica* (17%), *Cupressus* (14%) and *Linum* (7%) in Marche; *Corylus* (32%), *Brassica* (12%) and *Daucus* (7%) in Umbria; and *Corylus* (36%), *Brassica* (17%) and the green alga *Coccomyxa* (6%) in Veneto (Supplementary Materials Table S4). Seasonal variation was also observed. Considering all sites, in the spring, *Corylus* was the most abundant (57%), followed by *Cupressus* (7%) and the green alga *Coccomyxa* (6%); in summer, *Brassica* (24%), *Linum* (9%) and *Daucus*(9%); and in autumn, *Brassica* (24%), *Linum* (12%) and *Rubus* (8%) (Supplementary Materials Table S5).

409 The most abundant fungal taxa at site level and across all seasons, were *Cladosporium*
410 (34%, 40%, 19% and 44% respectively), *Alternaria* (23%, 6%, 12% and 15% respectively)
411 and *Epicoccum* (15%, 19%, 8% and 8% respectively) in FVG, Marche, VdA and Veneto,
412 while *Cladosporium* (38%), *Calocybe* (13%) and *Alternaria* (11%) were the most abundant in
413 Umbria (Supplementary Materials Table S4). Seasonal variation was also observed.
414 Considering all sites, in spring *Cladosporium* was the most abundant (57%) followed by
415 *Alternaria* (11%) and *Calocybe* (9%), while in summer and autumn, *Cladosporium* (34%
416 each), *Alternaria* (14% and 17% respectively) and *Epicoccum* (15% and 10% respectively)
417 dominated (Supplementary Materials Table S5).
418 Taking into account both sites and seasons for plants, *Corylus* was the most abundant
419 genus in spring for all sites but Marche, in which *Cupressus* was prevailing (Fig. 4A).
420 Summer and autumn share the same most abundant taxa across all the sites, i.e. *Brassica*
421 followed by *Linum*, *Cucurmis* and *Daucus* (Fig. 4A). The composition of fungi is more
422 homogenous, being *Cladosporium*, *Alternaria* and *Epicoccum* the most abundant taxa among
423 all seasons and sites (Fig. 4B).
424 The PERMANOVA outcome (Table 3) and the NMDS ordination (Supplementary
425 Materials Fig. S3) for both plants and fungi confirmed a significant effect of seasonality for
426 plants; post hoc test showed that all the pairwise comparisons were highly significant ($p <$
427 0.001, see Supplementary Materials Table S6). In contrast, fungal community composition
428 varied according to the interaction between Season and Site (Table 3, Supplementary
429 Materials Table S6). This output was further corroborated by the indicator species analysis,
430 which highlighted that fungal taxa were strictly associated with the interaction between
431 Season and Site while plants were not (Supplementary Materials Table S7). Interestingly, the

pollen of alien species (e.g. *Ambrosia artemisiifolia*, *Amaranthus* spp., *Ailanthus altissima*, *Robinia pseudacacia*) showed the highest value of the association index.

RDA (Fig. 5) were highly significant for both plants and fungi ($F_{(9,77)} = 2.94$, $p < 0.001$ and ($F_{(5,81)} = 1.85$, $p < 0.01$, respectively). The climatic variables retained after the variable selection procedure plus the Site factor are showed in Fig. 5. Notably, only *Ur_ave* was significant for both plants and fungi ($p < 0.001$). In addition, *Temp_min* ($p < 0.05$), *Temp_ave* ($p < 0.05$), *Temp_max* ($p < 0.001$), *Wind_ave* ($p < 0.05$) and *Ur_ave* ($p < 0.001$) were found to be significant for plants (Fig. 5A).

Discussion

Simultaneous amplification of plant and fungal DNA from aerobiological samples

eDNA metabarcoding studies have usually focused on single organismal groups, e.g., bacteria, fungi, metazoa, etc., or single species, e.g., rare or threatened species, invasive species etc., in a multiplicity of biotic and abiotic systems (Banchi et al., 2019; Deiner et al., 2017; Taberlet et al., 2012). In aerobiology, eDNA metabarcoding studies have either focused on plants, fungi and/or bacteria in indoor and outdoor environments; however, only a few report on the simultaneous analyses of either two of these three organismal groups (Banchi et al., 2019, and references therein). Aerobiological samples are indeed complexes of taxa belonging to different kingdoms, including pollen, fungal spores, plant and fungal fragments, bacteria, algae, and other biodiversity. Members of the airborne communities are present in different relative abundances, and the detection and characterization of the complex communities remains challenging.

In this study, the combination of eDNA metabarcoding and HTS technology was successfully applied to detect and characterize the airborne diversity of plants and fungi

simultaneously in a nine-month long survey. This was achieved by using multiple combinations of primers to detect both organismal groups. Primers were selected or newly designed to comprehensively capture the taxonomic diversity of the targeted communities. The performance of different primer combinations has been screened for different organisms and sequencing platforms implemented commonly used (e.g. Banos et al., 2018; Bylemans et al., 2018; Riit et al., 2016). Primer specificity is known to represent a key factor in capturing the full range of organismal diversity in mixed samples, especially if the selected barcode is the same for different groups. The ITS region is a common DNA barcode for species identification, designated as a universal barcode for fungi (Nilsson et al., 2009; Schoch et al. 2012), and often used also for plants (Johnson et al., 2019; Kress, 2017; Tremblay et al., 2018), as in this study. As the ITS marker meets the characteristics of a DNA barcode in many cases (Kress & Erickson, 2008), it is suitable for designing primers with a wide range of coverage for different groups of taxa (Cheng et al., 2016; Ruppert et al., 2019). Here, we tested three different primer combinations – using commonly used primers and others that were slighted modified to amplify a broader range of organisms. We also used three different PCR approaches to prepare the libraries for sequencing. The primer pairs were either used individually in single PCR reactions or mixed in the same PCR reaction to produce at the end five different types of libraries which were sequenced. The approach which performed best, capturing the greatest diversity of airborne plants and fungi, and was therefore applied for all samples, was the “individual” one, in which each sample was firstly amplified using all three primer combinations individually, and subsequently the three PCR products were pooled to prepare the library. In this “individual” type of amplification, both plants and fungi were amplified starting from an equal amount of DNA. However, among plants, Chlorophyta could be better amplified and sequenced than in the other two amplification approaches “mixed and

480 "primer pairs", respectively. The "mixed" approach seemed to favor the amplification of
481 fungi, hinting to a possible higher affinity of primers (i.e. higher annealing efficiency) to
482 fungal DNA in mixed samples. Also, in this "mixed" approach, the proportion of sequenced
483 fungal taxa almost corresponded to that obtained using only the fungal-specific primer pair
484 ITS-u2F/ITS-f4. However, if primer pairs are used individually, the pair ITS-u2_F/ITS4U-R
485 is the one which performs best. It co-amplified both plants and fungi of all targeted groups
486 (i.e. Streptophyta, Chlorophyta, Ascomycota and Basidiomycotya), and generated the lowest
487 amount of sequences which remained unassigned. However, the selection of optimal fungal
488 ITS primers still remains unresolved. Li et al. (2019) recently demonstrated that three primers
489 sets could not reach a consensus on fungal community compositions or diversities, and that
490 primer selection determined the observed diversity and composition in soil fungal
491 communities.
492
493 *The importance of mock community in eDNA metabarcoding*
494 A standardized mock community represents an invaluable control in amplicon sequencing
495 studies (Yeh et al., 2018). Though this type of control has been largely neglected at the
496 beginning of the HTS technology, mock communities have been shown to be essential to
497 estimate the reliability of the sequencing results and potential biases related to the eDNA
498 metabarcoding technique (Lear et al., 2018; Rocchi et al., 2017). Increasingly, researchers
499 now include specific, often *ad hoc*-created, mock communities to improve or compare
500 different PCR conditions or to evaluate error rates in the final datasets (Bakker, 2018). Mock
501 communities can be prepared according to different experimental designs, either assembling
502 the DNA of the various taxa in even or in staggered abundances (simulating the different
503 abundances of species as naturally present in many microbial communities), or selecting

different genetic markers, or they can be used to process the sequenced reads with different pipelines and compare the outputs (Bakker, 2018). Recently three mock communities have been proposed for fungi by Bakker (2018). Others have been developed for specific analyses in environmental and medical contexts, such as for arbuscular mycorrhizal fungi (Egan et al., 2018) and fungi in the respiratory tract (McTaggart et al., 2019), but to our knowledge none has been established for mixed samples of plant and fungi together so far.

The two mock communities assembled for this study were comprised of plants and fungi using even and staggered amounts of DNA, including taxa that were expected to be present in the analyzed airborne samples. In particular, these two types of mock communities mirrored the analyzed airborne samples, in which fungal spores or pollen of some taxa can be much more abundant than others. Our mock communities showed high level of reproducibility among the two different sequencing runs and a high level of taxonomic resolution: all the taxa could be identified to the genus level and the half of the taxa up to their species level. Though the preparation of fungal mock community from non-axenic tissues or cultures can be criticized to be a source of contaminations (Bakker, 2018), all the reads that we detected belonged to the community without off-target species, confirming the robustness of the pipelines followed for both the laboratory work and data analysis.

The interpretation of eDNA metabarcoding as a quantitative technique is still debated (Banchi et al., 2019; Lamb et al., 2018 and references therein), because the proportion of reads does not reliably represent the abundance of a certain taxon in the community. Nevertheless, a consensus has been found in considering it a semi-quantitative and not just a presence/absence approach (Amend et al., 2010; Rocchi et al., 2017). This statement is also evidenced by the analyses of our mock communities. Indeed, the sequencing results of the "even" mock community show how some taxa were significantly overrepresented and other

underrepresented. This disproportion is, though, less accentuated in the staggered mock community, where the originally different amounts of DNA influenced the proportion of reads, being the most abundant taxa those for which the highest number of reads was recovered. However, we here have not taken into account the quantity of DNA each species has, which defines the number of gene copies of the targeted marker (ITS2; Lofgren et al., 2019; Bradshaw et al. 2020) and may impair the results of the PCR amplifications (Schoch et al. 2012). It is likely that future researches will succeed in correlating the amount of the target gene by quantitative PCR (qPCR).

Airborne plant and fungal diversity relates to spatial and temporal variables

Our analyses recovered reads from 158 plant (153 Streptophyta and 5 Chlorophyta) and 613 fungal genera (349 Ascomycota, 259 Basidiomycota and 5 Mucoromycotina) across the five Italian localities in a nine-month long sampling. These results captured nearly 10 times the diversity obtained by the daily traditional morphological analyses of pollen and spores that are reported by the regional ARPA agencies (www.pollnet.it). Furthermore, these results can potentially be used in assessing the patterns of occurrence and spread of plants and fungi according to climatic conditions. A better understanding of these patterns may be of particular interest especially when allergenic taxa are detected, allowing for the knowledge of their dispersion patterns and the recommendations of useful guidelines needed for future prevention actions (Sicar et al., 2018). Simulations of pollen and spore dispersion in Europe have been based on results of local and short-term studies and have focused on only a few species (Sofiev et al., 2013; Veriankanité et al., 2010; Zhang et al., 2014). The variation of fungal communities across space and time was somehow quite expected; in fact, fungal spores usually have longer atmospheric residence times with respect to pollen grains, being

552 transported up to thousands of kilometers away from their source (Mayol et al., 2017).
553 *Cladosporium*, *Alternaria*, *Epicoccum* and *Calocybe* are considered the most abundant genera
554 (listed in the order of their abundance), even though there is a significant difference mainly
555 for the less abundant taxa, such as *Sporormiella* and *Oidium*. For these, the more local
556 occurrence may suggest the presence of certain microniches specifically occupied by them. In
557 contrast, plant community showed only temporal variation and localities seem not to play a
558 significant role. *Corylus* was the most abundant genus recovered in spring for all sites except
559 for Marche (for which *Cupressus* was prevailing), whereas in summer and autumn the highest
560 abundance was detected for *Brassica* followed by *Linum*, *Cucurmis* and *Daucus* in all the
561 sampling stations.

562 Contrasting hypotheses have been formulated to explain airborne particle composition
563 (i.e. local *vs* external sources). While the external sources hypothesis predicts that propagules
564 can be transported miles away from their origin under suitable weather conditions (Diamialis
565 et al., 2017; Mayol et al., 2017), local source hypothesis foresees that airborne particles
566 originate close to the sampling station (Skjøth et al. 2012; Oteros et al., 2015; Rojo et al.,
567 2015). Our findings tend to support the local source hypothesis (see Table 3). Indeed, if
568 propagules were transported far from their origin, we would expect a higher degree of
569 similarity among sites and seasons, which is in conflict with the observed pattern of strong
570 variation among site and seasons, suggesting potential differences in air mass circulation
571 among locations - see for instance Innocente et al. (2017) for bacteria. Furthermore, in the
572 present study the sites differ for their climate and orographic characteristics. The effect of
573 climatic variables on airborne samples has been reported (Favero-Longo et al., 2014;
574 Fröhlich-Nowoisky et al., 2016). Our findings are in line with previous metabarcoding and
575 pyrosequencing studies, which have shown remarkable seasonal variation for fungal genera

such as *Alternaria*, *Cladosporium* and *Blumeria*, likely reflecting the different life-styles, phenological differences and substrate preferences of the taxa (Fröhlich-Nowoisky et al., 2009; Yamamoto et al., 2012; Nicolaise et al., 2017). However, Nicolaise et al. (2017) highlighted that airborne fungal composition did not show any distinctive clustering based on sampling site, confirming that their location, within the climatic region studied, was not an important driver of the diversity of airborne fungal spores. Similarly to the study of Nicolaise et al. (2017) we also analyzed roof top samples, thus taking into consideration air masses with a likely greater mixing than that of air at lower elevations. Relative humidity is confirmed to be one of the main factors influencing fungal community (Sadyś et al., 2015), whereas pollen samples are mainly influenced by temperature and wind speed (Grinn-Gofroń et al., 2018). Nonetheless, it is known that certain fungi sporulate either during humid and dry periods, though, local weather events can distort the general pattern (Grinn-Gofroń et al., 2018). For example, in our survey we recovered *Epicoccum*, *Pithomyces* and *Peniophora* being most abundant in the drier summer period. For these reasons, extended areas and time frames should, therefore, be considered in future surveys in order to obtain more reliable information since local climatic phenomena may considerably affect the spreading of pollen and fungal spores. Sampling strategies could be improved further, i.e. by including sampling at additional 'sub-stations' in the same area for statistical support. This is, however, constrained by the availability of multiple sampling devices of the same type, and should consider a normalization of the microclimate variable among each sub-station.

Our study highlights the recurrent presence of several pathogenic, allergenic and invasive alien species (43 species detected) in airborne samples, as well as that of more inconspicuous and neglected taxa, such as lichens (25 species of lichenized fungi). In

599 particular, concerning lichens, long distance dispersal through air may corroborate the wide
600 distribution of certain species (e.g. Otálora et al., 2010; Wert et al., 2011; Leavitt et al., 2018).
601
602 *eDNA metabarcoding for aerobiology*
603 The analysis of eDNA using metabarcoding approaches provides novel perspectives into
604 complex communities from either aquatic, terrestrial or aerial ecosystems. eDNA
605 metabarcoding offers a faster, more direct and accurate method to capture a broader range of
606 biodiversity than traditional morphology-based surveys (Johnson et al., 2019; Ruppert et al.,
607 2019). Though there was a boom of eDNA metabarcoding studies applied to soil and water
608 samples, in contrast extremely few ones have considered aerobiological specimens to date.
609 Only recently, studies are focusing on the proper methodology to make the analysis of
610 aerobiological samples straightforward by testing different sampling methodologies, DNA
611 extraction methods and primer combinations (Johnson et al., 2019). Indeed, different eDNA
612 extraction methods may lead to varying amounts of extracted DNA, with consequently
613 different amplification rate and concentration of PCR products, phenomena which deserve
614 further study and optimization according to the aim of the researches (Johnson et al., 2019).
615 The application of airborne eDNA metabarcoding needs, therefore, to expand and to establish
616 in aerobiology to augment the accuracy of data in biomonitoring studies, detection of invasive
617 and allergenic species (either plants and fungi) or modelling of species dispersion.
618 To the best of our knowledge, this study is one of the few which applies eDNA
619 metabarcoding analyses to airborne samples and the first which takes into account a nine-
620 month-long survey on a wide geographic area to analyze simultaneously airborne plants and
621 fungi. The protocols performed during the laboratory portion of the study were revised and
622 improved from previous pilot study considering only airborne fungal eDNA (Banchi et al.

623 2018), extending the protocols to the simultaneous analyses of airborne plants and fungi. In
624 addition to methodological developments, the implementation of eDNA metabarcoding
625 enables the detection of higher biodiversity and provides insight into the temporal and spatial
626 distribution of this diversity. DNA metabarcoding is proposed as a suitable and powerful tool
627 for detecting rare, pathogenic and invasive species of both fungi and plants, and shows its
628 potentials to complement routine analyses of air biomonitoring. We anticipate that our study
629 will serve as a springboard for future investigations of biodiversity monitoring and monitoring
630 of invasive alien species both for facing human health and agricultural prevention issues.
631

632 **Acknowledgments**

633 The research was funded by the project Finanziamenti di Ateneo per progetti di Ricerca
634 scientifica (FRA2016) assigned to LM by the University of Trieste. We thank Fiorella Florian
635 and Fabrizia Gionechetti (University of Trieste, Italy) for technical help in the laboratory,
636 Elisa Peressotti for suggestions about *E. necator* manipulation, and Silvia Vezzulli and
637 Massimo Pindo (Edmund Mach Foundation, San Michele all'Adige, Italy) for *Erysiphe*
638 *necator* spores and for DNA sequencing of some of the mock community samples,
639 respectively. We also thank ARPA Friuli Venezia Giulia, Marche, Umbria, Valle d'Aosta and
640 Veneto for providing meteorological data. The Grainger Bioinformatics Center, Science and
641 Education, Field Museum of Natural History (Chicago, U.S.A.) is thanked for the offered
642 genome sequencing facilities. Steven D. Leavitt is thanked for revising the English style of
643 the text.
644
645

646 **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,* 511–524. doi:10.1111/1755-0998.12755

Amend, A. S., Seifert, K. A., & Bruns, T. D. (2010). Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology, 19,* 5555–5565. doi:10.1111/j.1365-294X.2010.04898.x

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,* 1–12. doi:10.1007/s10453-017-9490-6

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology, 26,* 32–46. http://doi.org/10.1111/j.1442-9993.2001.01070.pp.x

Anderson, M. J., Gorley, R. N., & Clarke, K. R. (2008). PERMANOVA+ for PRIMER: guide to software and statistical methods. Primer-E, Ltd, Plymouth, UK.

Baker, C. (2016). Bakerccm/entrez_qiime: Entrez_qiime v2.0.

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

Banchi, E., Ametrano, C. G., Stanković, D., Verardo, P., Moretti, O., Gabrielli, F., …, & Muggia, L. (2018). DNA metabarcoding uncovers fungal diversity of mixed airborne samples in Italy. *PLoS ONE, 13,* e0194489. http://doi.10.1371/journal.pone.0194489

Banchi, E., Pallavicini, A., & Muggia, L. (2019). Relevance of plant and fungal DNA metabarcoding in aerobiology. *Aerobiologia*, http://doi.org/10.1007/s10453-019-09574-2

Banos, S., Lentendu, G., Kopf, A., Wubet, T., Gloeckner, F. O., & Reich, M. (2018). A comprehensive fungi-specific 18S rRNA gene sequence primer toolkit suited for diverse research issues and sequencing platforms. *BMC Microbiology, 18,* 190. http://doi.10.1186/s12866-018-1331-4

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

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., ... & Amend, A. (2013). Improved software detection and extraction of ITS1 and ITS 2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods in ecology and evolution, 4,* 914–919. http://doi.10.1111/2041-210X.12073

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al-Ghalith, G. A., ..., & Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature biotechnology, 37,* 852–857. https://doi.org/10.1038/s41587-019-0209-9

Bradshaw, M., Grewe, F., Thomas, A., Harrison, C. H., Lindgren, H., Muggia, L., ... & Leavitt, S. D. (2020). Characterizing the ribosomal tandem repeat and its utility as a DNA barcode in lichen-forming fungi. *BMC Evolutionary Biology, 20,* 2. https://doi.org/10.1186/s12862-019-1571-4

Brennan, G. L., Potter, C., de Vere, N., Griffith, G. W., Skjoth, C. A., Osborne, N. J., …., & Petch, G. M. (2019). Temperate grass allergy season defined by spatio-temporal shifts in airborne pollen communities. *Nature Ecology and Evolution, 3,* 750–754. http://doi. 10.1038/s41559-019-0849-7

Bylemans, J., Gleeson, D. M., Hardy, C. M., & Furlan, E. (2018). Toward an ecoregion scale evaluation of eDNA metabarcoding primers: a case study for the freshwater fish biodiversity of the Murray–Darling Basin (Australia). *Ecology and Evolution, 8,* 8697–8712. http://doi.10.1002/ece3.4387.

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,* 138–149. http://doi.10.1111/1755-0998.12438

Clarke, K. R., & Gorley, R. N. (2006). PRIMER V6: User manual/tutorial. Plymouth: Primer-E Ltd.

De Caceres, M., & Legendre, P. (2009) Associations between species and groups of sites: indices and statistical inference. *Ecology, 90,* 3566– 3574.

Damialis, A., Kaimakamis, E., Konoglou, M., Akritidis, I., Traidl-Hoffmann, C., & Gioulekas, D. (2017). Estimating the abundance of airborne pollen and fungal spores at variable elevations using an aircraft: how high can they fly?. *Scientific reports*, *7*, 1-11. doi:10.1038/srep44535

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,* 1076–1086. http://doi.10.1093/annweh/wxx075

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., ..., & Bernatchez, L. (2017). Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Molecular Ecology, 26,* 5872–5895. http://doi.10.1111/mec.14350

Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guénard, G., Jombart, T., ... & Wagner, H. H. (2019). adespatial: Multivariate multiscale spatial analysis. R package version 0.3–3.

Dufrene, M. & Legendre, P. (1997). Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs, 67,* 345–366. http://doi.org/10.2307/2963459

Egan, C.P., Rummel, A., Kokkoris, V., Klironomos, J., Lekberg, Y., & Hart, M. (2018). Using mock communities of arbuscular mycorrhizal fungi to evaluate fidelity associated with Illumina sequencing. *Fungal Ecology, 33,* 52–64. http://doi.org/10.1016/j.funeco.2018.01.004

Elbert, W., Taylor, P. E., Andreae, M. O., & Poschl, U. (2007). Contribution of fungi to primary biogenic aerosols in the atmosphere: wet and dry discharged spores, carbohydrates, and inorganic ions. *Atmospheric Chemistry and Physics, 7,* 4569–4588. http://doi.org/10.5194/acp-7-4569-2007

Favero-Longo, S. E., Sandrone, S., Matteucci, E., Appolonia, L., & Piervittori, R. (2014). Spores of lichen-forming fungi in the mycoaerosol and their relationships with climate factors. *Science ofTotal Environment, 466–467,* 26–33. http://doi.10.1016/j.scitotenv.2013.06.057

Fröhlich-Nowoisky, J., Kampf, C. J., Weber, B., Huffman, J. A., PoÈhlker, C., Andreae, M. O., …. & Pöschl, U. (2016). Bioaerosols in the earth system: climate, health, and ecosystem interactions. *Atmospheric Research, 182,* 346–376. http://doi.org/10.1016/j.atmosres.2016.07.018

Fröhlich-Nowoisky, J., Pickersgill, D. A., Despres, V. R., & Poschl, U. (2009). High diversity of fungi in air particulate matter. *Proc. Natl. Acad. Sci. U.S.A. 106,* 12814–12819. doi: 10.1073/pnas.0811003106

Fu, L., Niu, B., Zhu, Z., Wu, S., &Li, W. (2012). CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics, 28,* 3150–3152. http://doi.10.1093/bioinformatics/bts565

Galasso, G., Conti, F., Peruzzi, L., Ardenghi, N. M. G., Banfi, E., Celesti-Grapow, L., ...& Bartolucci, F. (2018). An updated checklist of the vascular flora alien to Italy, *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology, 152,* 556–592. http://doi.10.1080/11263504.2018.1441197

García-Mena, J., Murugesan, S., Pérez-Muñoz, A. A., García-Espitia, M., Maya, O., Jacinto-Montiel, M., ... & Núñez-Cardona, M. T. (2016). Airborne bacterial diversity from the low atmosphere of greater Mexico City. *Microbial Ecology, 72,* 70–84. http://doi.10.1007/s00248-016-0747-3

Grinn-Gofroń, A., Bosiacka, B., Bednarz, A., & Wolski, T. (2018). A comparative study of hourly and daily relationships between selected meteorological parameters and airborne fungal spore composition. *Aerobiologia, 34,* 45–54. http://doi.org/10.1007/s10453-017-9493-3

Holdaway, R. J., Wood, J. R., Dickie, I. A., Orwin, K. H., Bellingham, P. J., Richardson, S. J., ... & Buckley T. R. (2017). Using DNA metabarcoding to assess New Zealand's terrestrial biodiversity. *New Zealand Journal of Ecology, 41,* 251–262. http://doi.10.20417/nzjecol.41.28

Johnson, M. D., Cox, R. D., & Barnes, M. A. (2019). Analyzing airborne environmental DNA: A comparison of extraction methods, primer type, and trap type on the ability to detect airborne eDNA from terrestrial plant communities. *Environmental DNA, 1,* 176–185. http://doi. 10.1002/edn3.19

Innocente, E., Squizzato, S., Visin, F., Facca, C., Rampazzo, G., Bertolini, V., ..., & Bestetti, G. (2017). Influence of seasonality, air mass origin and particulate matter chemical composition on

752 airborne bacterial community structure in the Po Valley, Italy. *Science of The Total*
753 *Environment, 593*, 677-687. https://doi.org/10.1016/j.scitotenv.2017.03.199
754 Korpelainen, H., & Pietiläinen, M. (2017). Diversity of indoor fungi as revealed by DNA
755 metabarcoding. *Genome, 60,* 55–64. http://doi.10.1139/gen-2015-0191
756 Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger
757 climate classification updated. *Meteorologische Zeitschrift, 15,* 259–263.
758 http://doi.10.1127/0941-2948/2006/0130
759 Kraaijeveld, K., de Weger, L. A., Ventayol, G. M., Buermans, H., Frank, J., Hiemstra, P. S., & den
760 Dunnen, J.T. (2015). Efficient and sensitive identification and quantification of airborne pollen
761 using next-generation DNA sequencing. *Molecular Ecology Resources, 15,* 8–16.
762 http://doi.10.1111/1755-0998.12288
763 Kramer, C. L. (1982). Production, release and dispersal of basidiospores. In: Frankland, J.C., Hedger,
764 J.N., & Swift M. J. (Eds.), Decomposer Basidiomycetes: their biology and ecology, pp 33–49,
765 Cambridge University Press.
766 Kress, J. W. (2017). Plant DNA barcodes: Applications today and in the future. *Journal of Systematic*
767 *and Evolution, 55,* 291–307. http://doi.org/10.1111/jse.12254
768 Kress, W. J., & Erickson D. L. (2008). DNA barcodes: Genes, genomics, and bioinformatics.
769 *Proceeding of the Natural Academy of Science U S A, 105,* 2761–2762.
770 Kubartova, A., Ottosson, E., Dahlberg, A., & Stenlid, J. (2012). Patterns of fungal communities
771 among and within decaying logs, revealed by 454 sequencing. *Molecular Ecology, 21,* 4514–
772 4532. http://doi.10.1111/j.1365-294X.2012.05723.x
773 Lacey, M. E., & West, J. S. (2006). *The Air Spora.* http://doi.10.1007/s13398-014-0173-7.2
774 Lamb, P.D., Hunter, E., Pinnegar, J.K., Creer,S., Davies, R.G., Taylor, M. (2018). How quantitative is
775 metabarcoding: a meta-analytical approach. *Molecular Ecology, 28,* 420–430.
776 http://doi.10.1111/mec.14920
777 Lear, G., Dickie, I., Banks, J., Boyer, S., Buckley, H., Buckley, T., ..., & Holdaway, R. J. (2018).
778 Methods for the extraction, storage, amplification and sequencing of DNA from environmental
779 samples. *New Zealand Journal of Ecology, 42,* 10. http://doi.10.20417/nzjecol.42.9
780 Legendre, P., &Legendre,L. (2012). *Numerical Ecology*, 3rd edition. Elsevier, Amsterdam.
781 Li, S., Deng, Y., Wang, Z., Zhang, Z., Kong, X., Zhou, W., …, & Qu, Y. (2019) Exploring the
782 accuracy of amplicon-based internal transcribed spacer (ITS) markers for fungal community.
783 *Molecular Ecology Resources* (in press). https://doi.org/10.1111/1755-0998.13097
784 Lofgren, L. A., Uehling, J. K., Branco, S., Bruns, T. D., Martin, F., & Kennedy, P. G. (2019).
785 Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and
786 ecological lifestyles. *Molecular Ecology, 28,* 721–730. https://doi.org/10.1111/mec.14995
787 Mayol, E., Arrieta, J. M., Jiménez, M. A., Martínez-Asensio, A., Garcias-Bonet, N., Dachs, J., ... &
788 Duarte, C. M. (2017). Long-range transport of airborne microbes over the global tropical and
789 subtropical ocean. *Nature Communication, 8,* 201. http://doi.10.1038/s41467-017-00110-9
790 McTaggart, L. R., Copeland, J. K., Surendra, A., Wang, P. W., Husain S., Coburn, B., ... &, Kus, J.
791 W. (2019). Mycobiome sequencing and analysis applied to fungal community profiling of the
792 lower respiratory tract during fungal pathogenesis. *Frontiers in Microbiology, 10,* 512.
793 doi.org/10.3389/fmicb.2019.00512
794 Nacke, H., Goldmann K., Schöning, I., Pfeiffer, B., Kaiser, K., ..., & Wubet, T. (2016). Fine spatial
795 scale variation of soil microbial communities under European beech and Norway spruce.
796 *Frontiers in Microbiology, 7,* 2067. doi.org/10.3389/fmicb.2016.02067
797 Nicolaise M., West, J. S., Sapkota, R., Canning, G. G. M., Schoen, C., & Justesen, A. F. (2017)
798 Fungal communities including plant pathogens in near surface air are similar across
799 Northwestern Europe. *Frontiers in Microbiology, 8,* 1729
800 Nilsson, R. H., Ryberg, M., Abarenkov, K., Sjökvist, E., & Kristiansson, E. (2009). The ITS region as
801 a target for characterization of fungal communities using emerging sequencing technologies.
802 *FEMS Microbiology Letters, 296,* 97–101.
803 Nilsson, R. H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M., ..., &
804 Abarenkov, K. (2015). A comprehensive, automatically updated fungal ITS sequence dataset for

reference-based chimera control in environmental sequencing efforts. *Microbes and Environments, 30,* 145–150. http://doi.10.1264/jsme2.ME14121

Norros, V., Rannik, Ü, Hussein, T., Petäjä, T., Vesala, T., & Ovaskainen, O. (2014). Do small spores disperse further than large spores? *Ecology, 95,* 1612–1621. http://doi.10.1890/13-0877.1

Núñez, M.A., Chiuffo, M.C., Torres, A., Paul, T., Dimarco, R.D., Raal, P., ..., & Richardson, D.M. (2017). Ecology and management of invasive Pinaceae around the world: progress and challenges. *Biological Invasions, 19,* 3099–3120. http://doi.org/10.1007/s1053

Oksanen, J., Kindt, R., Minchin, P.R., Simpson, G.L., Blanchet, F.G., Legendre, P., ..., & Wagner, H. H. (2019). Vegan: Community Ecology Package. R package version 2.0-2.

Otálora, M. A. G., Martínez, I., Aragón, G., & Molina, M. C. (2010). Phylogeography and divergence date estimates of a lichen species complex with a disjunct distribution pattern. *American Journal of Botany, 97,* 216–223. doi: 10.3732/ajb. 0900064

Oteros, J., Garcia-Mozo, H., Alcazar, P., et al. (2015) A new method for determining the sources of airborne particles. *Journal of Environmental Management, 155,* 212–218.

Pashley, C. H, Fairs, A., Free, R. C., Wardlaw, A. J. (2012). DNA analysis of outdoor air reveals a high degree of fungal diversity, temporal variability, and genera not seen by spore morphology. *Fungal Biology, 116,* 214–24. doi:10.1016/j.funbio.2011.11.004

Pejchar, L., & Mooney, H. (2009). Invasive species, ecosystem services and human well-being. *Trends in Ecology & Evolution, 24,* 497–504. doi:10.1016/j.tree.2009.03.016

Rieux, A., Soubeyrand, S., Bonnot, F., Klein E. K., Ngando, J. E., Mehl, A., ..., & de Lapeyre de Bellaire, L. (2014). Long-distance wind-dispersal of spores in a fungal plant pathogen: estimation of anisotropic dispersal kernels from an extensive field experiment. *PLoS ONE, 9,* e103225. http://doi.org/10.1371/journal.pone.0103225

Riit, T., Tedersoo L., Drenkhan, R., Runno-Paurson, E., Kokko, H., & Anslan, S. (2016). Oomycete-specific ITS primers for identification and metabarcoding. *MycoKeys, 14,* 17–30. doi: 10.3897/mycokeys.14.9244

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,* 107–112. http://doi.org/10.1016/j.mimet.2017.05.014

Rognes, T., Flouri, T., Nichols, B., Quince, C., &Mahé, F. (2016). VSEARCH: a versatile open source tool for metagenomics. *PeerJ, 4,* e2584. doi:10.7717/peerj.2584

Rojo, J., Lr, B., Rapp, A., Fernandez-Gonzales, F., & Perez-Badia, R. (2015). Effect of land use and wind direction on the contribution of local sources to airborne pollen. *Science of The Total Environment, 538,* 672–682. doi: 10.1016/j.scitotenv.2015.08.074

Ruppert, K. M. Richard, J. K., & Rahman, Md. S. (2019). Past, present and future perspectives of environmental DNA (eDNA) barcoding: a systematic review in methods, monitoring, and application of global eDNA. *Global Ecology and Conservation, 17,* e00547. doi: 10.1016/j.gecco.2019.e00547

Sadyś, M., Strzelczak, A., Grinn-Gofroń, A., & Kennedy, R. (2015). Application of redundancy analysis for aerobiological data. *International Journal of Biometeorology, 59,* 25–36. doi: 10.1007/s00484-014-0818-4

Schoch, C. L., Seifert, K. A., Huhndorf, S., Robert, V., Spouge, J. L., Levesque, C. A., ..., Chen, W, &Fungal Barcoding Consortium. (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,* 1–6. doi: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‘i1. *Journal of Phycology, 53,* 437–445. doi:10.1111/jpy.12502

Sicard, M., Jorba, O., Izquierdo, R., Alarcón, M., De Linares, C., & Belmonte, J. (2019). Modelling of airborne pollen dispersion in the atmosphere in the Catalonia region, Spain: model description, emission scheme and evaluation of model performance for the case of *Pinus*. *Remote Sensing of Clouds and the Atmosphere XXIV, 11152,* 111520O.

Skjøth, C. A., Sommer, J., Frederiksen, L. & Gosewinkel Karlson, U. (2012). Crop harvest in Denmark and Central Europe contributes to the local load of airborne *Alternaria* spore concentrations in Copenhagen. *Atmospheric Chemistry and Physics,12,* 11107–11123. doi:10.5194/acp-12-11107-2012

Sofiev, M., Siljamo, P., Ranta, H., Linkosalo, T., Jaeger, S., Rasmussen, A., ..., & Kukkonen, J. (2013). A numerical model of birch pollen emission and dispersion in the atmosphere. Description of the emission module. *International Journal of Biometeorology, 57,* 45–58. doi: 10.1007/s00484-012-0532-z

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

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. *Biological Conservation, 183,* 4–18. doi.org/10.1016/j.biocon.2014.11.019

Tremblay, É. D., Duceppe, M.-O., Bérubé, J. A., Kimoto, T., Lemieux, C., & Bilodeau, G. J. (2018). Screening for exotic forest pathogens to increase survey capacity using metagenomics. *Phytopathology, 108,* 1509–1521. http://doi.org/10.1094/PHYTO-02-18-0028-R

Veriankanité, L., Siljamo, P., Sofiev, M., Šauliene, I., & Kukkonen, J. (2010). Modeling analysis of source regions of long-range transported birch pollen that influences allergenic seasons in Lithuania. *Aerobiologia, 26,* 47–62.

Werth, S. (2011). Biogeography and phylogeography of lichen fungi and their photobionts. In Fontaneto, D. (ed) Biogeography of Microscopic Organisms: Is Everything Small Everywhere?, Cambridge, Cambridge University Press, 191–208. doi: 10.1017/CBO9780511974878.011

White, T. J., Bruns, T., Lee, S. J. W. T., & Taylor, J. W. (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., &White, T.J. (Eds.). PCR protocols: a guide to methods and applications. pp. 315–322, New York, Academic Press.

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. *ISME J. 6,* 1801–1811. doi: 10.1038/ismej.2012.30

Yeh, Y-C,. Needham, D. M., Sieradzki, E. T., & Fuhrman, J. A. (2018). Taxon disappearance from microbiome analysis reinforces the value of mock communities as a standard in every sequencing run. *mSystems, 3,* e00023-18

Zhang, R., Duhl, T., Salam, M.T., House, J.M., Flagan, R. C., Avol, E.L., ..., & Van Reken, T.M. (2014). Development of a regional scale pollen emission and transport modeling framework for investigating the impact of climate change on allergic airway disease. *Biogeosciences, 11,* 1461–1478. doi:10.5194/bg-11-1461-2014

Data Accessibility

The sequence data are available at the NCBI short read repository under the accession number PRJNA576572.

902 **Author Contributions**

903 LM, AP designed the research; EB, CGA, LM, performed the research; DS, AP, MT
904 contributed analytical tools; EB, CGA, DS, ET, SO, LM, AP, analyzed the data; ARPA
905 working group contributed in the sampling; EB, LM, CGA, ET, AP wrote the paper.
906
907
908 **Captions to figures and Tables**
909 **Figure 1.** Geographical location of the sampling sites in North and Central Italy. Region and
910 city names are reported (FVG: Friuli Venezia Giulia; VdA: Valle d'Aosta). The map is
911 retrieved and modified from http://www.d-maps.com.
912
913 **Figure 2.** Workflow of the analyses of airborne plant and fungal particles. The three different
914 PCR approaches implemented and compared to evaluate whether the three primer
915 combinations affect sequencing results are indicated by "individual", "mixed" and "primer
916 pair". I PCR: primary PCR; II PCR: outer PCR to attach the molecular identifiers (MID) for
917 multiplex sequencing. The three different PCR amplifications (indicated as "a, b, c") were
918 performed using one forward (F) primer coupled with three reverse (R) primers.
919
920 **Figure 3**. Abundances of taxa reported with the percentage values of reads for (**A**) the mock
921 communities and (**B**) the six samples amplified with the three PCR approaches. **A.**
922 Taxonomic composition of the equal amount (E) and serially diluted (D) mock communities,
923 and for the equals standardized (ES) and diluted standardized (DS), which represent that it
924 should have proportionally been according to the initial DNA amount. **B.** Taxonomic
925 composition at phylum level resulted by summing the reads of the six samples obtained with

926 the three PCR approaches "individual", "mixed" and "primer pair", the latter indicated with
927 the names of the three different reverse primers used (ITS-f4, ITS-p4 and ITS4U_R).
928
929 **Figure 4**. Taxonomic composition at genus level in the three seasons for the five sampling
930 sites for plants (**A**) and fungi (**B**). Abundances of taxa are reported with the percentage values
931 of reads. The most abundant genera are showed (>20% for plants, >10% for fungi), while the
932 less abundant are grouped under “Other”.
933
934 **Figure 5.** Redundancy analyses (RDA) based on (A) plants and (B) fungi. Black dots
935 represent sampling sites, red vectors species, and blue vectors climatic variables. Both RDAs
936 were calculated using log(x+1) transformed species abundance data. The climatic variables
937 displayed here are those retained after forward selection procedure.
938
939 **Table 1**. Plant and fungal taxa used to construct the mock communities. For each sample,
940 kingdom, phylum, species and original material from which the DNA was extracted is
941 reported.
942
943 **Table 2.** Spearman’s correlation among taxa abundances at phylum level in the primer pair
944 comparison. Spearman’s ρ for each comparison is reported. ** $p< 0.01$; * $p< 0.05$.
945
946 **Table 3.** PERMANOVA output calculated for plant and fungi datasets based on Bray-Curtis
947 similarity and 4999 permutations of residuals under a reduced model.