

METHODOLOGY

Open Access



# Performance evaluation of a new custom, multi-component DNA isolation method optimized for use in shotgun metagenomic sequencing-based aerosol microbiome research

Kari Oline Bøifot<sup>1,2</sup>, Jostein Gohli<sup>1</sup>, Line Victoria Moen<sup>1</sup> and Marius Dybwad<sup>1,2\*</sup> 

## Abstract

**Background:** Aerosol microbiome research advances our understanding of bioaerosols, including how airborne microorganisms affect our health and surrounding environment. Traditional microbiological/molecular methods are commonly used to study bioaerosols, but do not allow for generic, unbiased microbiome profiling. Recent studies have adopted shotgun metagenomic sequencing (SMS) to address this issue. However, SMS requires relatively large DNA inputs, which are challenging when studying low biomass air environments, and puts high requirements on air sampling, sample processing and DNA isolation protocols. Previous SMS studies have consequently adopted various mitigation strategies, including long-duration sampling, sample pooling, and whole genome amplification, each associated with some inherent drawbacks/limitations.

**Results:** Here, we demonstrate a new custom, multi-component DNA isolation method optimized for SMS-based aerosol microbiome research. The method achieves improved DNA yields from filter-collected air samples by isolating DNA from the entire filter extract, and ensures a more comprehensive microbiome representation by combining chemical, enzymatic and mechanical lysis. Benchmarking against two state-of-the-art DNA isolation methods was performed with a mock microbial community and real-world air samples. All methods demonstrated similar performance regarding DNA yield and community representation with the mock community. However, with subway samples, the new method obtained drastically improved DNA yields, while SMS revealed that the new method reported higher diversity. The new method involves intermediate filter extract separation into a pellet and supernatant fraction. Using subway samples, we demonstrate that supernatant inclusion results in improved DNA yields. Furthermore, SMS of pellet and supernatant fractions revealed overall similar taxonomic composition but also identified differences that could bias the microbiome profile, emphasizing the importance of processing the entire filter extract.

**Conclusions:** By demonstrating and benchmarking a new DNA isolation method optimized for SMS-based aerosol microbiome research with both a mock microbial community and real-world air samples, this study contributes to improved selection, harmonization, and standardization of DNA isolation methods. Our findings highlight the importance of ensuring end-to-end sample integrity and using methods with well-defined performance characteristics. Taken together, the demonstrated performance characteristics suggest the new method could be used to improve the quality of SMS-based aerosol microbiome research in low biomass air environments.

**Keywords:** Aerosol microbiome, Air sampling, DNA isolation, Shotgun metagenomic sequencing

\* Correspondence: [marius.dybwad@ffi.no](mailto:marius.dybwad@ffi.no)

<sup>1</sup>Norwegian Defence Research Establishment, P.O. Box 25, NO-2027 Kjeller, Norway

<sup>2</sup>Department of Analytics, Environmental & Forensic Sciences, King's College London, 150 Stamford Street, London SE1 9NH, UK



## Background

The study of bioaerosols is an emerging and expanding research discipline [1], with several important study applications, including surveillance of clinically relevant microbes [2–5], air quality monitoring [6–8] and biodefense [9]. Bioaerosol research has traditionally relied on culture methods; however, not all microorganisms grow under standard laboratory conditions, resulting in underrepresentation of the true microbial diversity [10–13]. Although culture methods are still in use, culture-independent methods are now widespread. Due to the low amount of DNA that is typically obtained from air samples, most culture-independent bioaerosol studies to date have used PCR to target either the bacterial 16S rRNA gene [14, 15] or the fungal 18S rRNA gene/internal transcribed spacer (ITS) region, followed by amplicon sequencing [16, 17]. In contrast to the amplicon sequencing approach, shotgun metagenomic sequencing (SMS) allows for generic, unbiased interrogation of microbial diversity in a sample. However, SMS typically requires a higher quality and quantity of DNA for analysis than other molecular methods. SMS has been used to characterize the human microbiome [18] and environmental microbiomes [19, 20], and has recently been implemented in a few aerosol microbiome studies [2, 21, 22].

Although bioaerosols originate from many different sources and are ubiquitous in almost any indoor and outdoor environment, air is still a very low biomass environment compared to e.g. soil or feces [23]. The low biomass makes it challenging to obtain sufficient DNA amounts for downstream analyses, especially in the context of SMS [21]. An important first step in recovering sufficient biomass and a representative sample from air involves the use of well-characterized air samplers that are capable of rapid and efficient biomass collection [24]. Filter-based aerosol collection is a commonly used method, and the use of hand-portable, high-volume filter-based air sampling equipment may improve the spatiotemporal resolution in aerosol microbiome research [24, 25]. The post-sampling processing steps are also important since the filter-collected biomass must be transformed into a representative high quality DNA sample with minimal loss. It is therefore essential to use a well-characterized DNA isolation method that is capable of comprehensive biomass lysis, sufficient inhibitor removal and sample clean-up, and high efficiency recovery of DNA [25]. In short, the main challenges are typically obtaining sufficient DNA amounts and capturing representative samples that reflect the true diversity of the sampled air environment [2, 22, 25, 26].

With recent advancements in sequencing technology, along with the development of improved strategies for air sampling and sample processing, it should be possible to mitigate the low biomass challenge. Mitigation

strategies that have been attempted in the past include long-duration sampling (days to weeks), pooling of multiple air samples, whole genome amplification (WGA) techniques, and modification of commercial DNA isolation kits originally developed for other environmental matrices such as water and soil [2, 21, 27–29]. Increasing the air sampling time is a common strategy to improve the DNA yield, but this approach may not always be practical. For example, in studies where the aim is to address spatiotemporal variability, the need for long-duration air sampling (e.g. days to weeks) exclude the possibility of aerosol microbiome investigations on shorter timescales. Another challenge with increased air sampling time is that long-duration filter collection may compromise the integrity of stress-sensitive microorganisms, e.g. due to desiccation and osmotic shock [27], and thereby cause a potential loss of DNA from organisms that become membrane-compromised, ruptured or lysed during filter extraction or subsequent processing steps prior to DNA isolation. Liquid extraction of aerosol filters often results in sample volumes that are too large to process with most commercial DNA isolation kits. This introduces a need for adopting additional post-extraction filtration or centrifugation steps to reduce the sample volume before DNA isolation, which may result in loss of both intact microorganisms and DNA, and thereby compromise both the DNA yield and sample diversity. Furthermore, long-duration, high-volume air sampling alone does not always translate into successful recovery of sufficient DNA amounts for SMS [2, 21, 28, 29]. This may be due to the use of different downstream sample processing and DNA isolation methods that have not been sufficiently evaluated regarding their specific performance on air samples, and which therefore may deliver suboptimal performance regarding biomass lysis and/or DNA recovery efficiency. Various modifications of existing sample processing and DNA isolation methods have been proposed to improve the DNA yield from filter-collected air samples. Jiang et al. modified the DNeasy PowerSoil Kit (Qiagen; formerly MO BIO PowerSoil DNA Isolation Kit) by replacing the silica spin column with AMPure XP beads, and introduced sample pre-treatment steps and a secondary filtration step [28]. Yooseph et al. introduced a WGA step to generate sufficient DNA amounts from air samples for SMS [21]. King et al. performed liquid extraction of aerosol filters followed by a secondary filtration step and DNA isolation with the DNeasy PowerWater Kit, and precipitated DNA from the original filtrate before combining the two DNA fractions [2]. Dommergue et al., who also used the DNeasy PowerWater Kit, placed the aerosol filters directly in PowerBead tubes, introduced sample pre-treatment steps, and a centrifugation step to maximize lysate recovery from PowerBead tubes [29]. Recovery of

sufficient DNA amounts and preservation of microbial diversity from air samples is essential to ensure reliable results in SMS-based aerosol microbiome research. Several studies on other sample matrices have looked into how DNA yields can be improved and microbial diversity preserved. Tighe et al. found that using a multi-enzyme cocktail (MetaPolyzyme) that targets bacterial and fungal cell wall components resulted in improved DNA yields [30]. Yuan et al. evaluated different DNA isolation methods for human microbiome samples, and found bead beating and enzymatic lysis to be essential for obtaining an accurate representation of microorganisms in a complex mock community [31]. Abusleme et al. found that bead beating may limit the DNA yield, but also that bead beating was necessary to detect all organisms in a complex mock bacterial community [32]. These observations show that biomass lysis based on a combination of chemical, enzymatic and mechanical principles may be useful to minimize microbiome composition (diversity) bias resulting from insufficient biomass lysis during isolation of DNA from complex environmental assemblages.

It is well established that the choice of DNA isolation method should be based on careful consideration of the specific study aims, including type of targeted organisms and environmental matrices [33]. However, substantial uncertainty exists regarding the extent of microbiome composition (diversity) bias that may be introduced by the use of different sample processing and DNA isolation methods, which makes it difficult to reliably compare microbiome results between different studies and environments. Consequently, several attempts have in recent years been made to improve the harmonization and standardization of DNA isolation methods, especially for common sample matrices such as human [31, 34], soil [35], and water [36] samples. Lear et al. recommended DNA isolation kits for different environmental matrices such as soil, plant and animal tissue, and water [37]. The Earth Microbiome Project demonstrated how procedural standardization allows for comparison of microbial diversity in samples from across the globe [35]. Dommergue et al. proposed an air sampling, filter extraction and DNA isolation method where microbial diversity and chemical composition in air can be investigated using existing high-volume particulate matter samplers used for atmospheric pollution monitoring [29]. Nevertheless, despite substantial effort several unresolved issues remain, e.g., the current reliance on long-duration air sampling raises some questions regarding sample integrity and only offers support for low temporal resolution studies since the necessary sampling time may be days or even weeks. Hence, performance benchmarking, harmonization,

and standardization of air sampling, sample processing and DNA isolation methods is a topic that warrants further study, and especially in the context of SMS-based aerosol microbiome research, which is a research field still largely in its infancy.

The aim of this study was to demonstrate a new custom, multi-component DNA isolation method optimized for SMS-based aerosol microbiome research and perform a comprehensive performance benchmarking of the new method. The custom, multi-component DNA isolation method was specifically developed to maximize the DNA yield and ensure comprehensive biomass lysis from low biomass environmental air samples. The DNA isolation method, hereafter referred to as the “MetaSUB method”, was developed for the MetaSUB Consortium ([www.meta-sub.org](http://www.meta-sub.org)) to complement an ongoing global effort to characterize subway and urban environment microbiomes using surface swab samples, by extending the effort to also include air samples. The MetaSUB method was benchmarked against two other state-of-the-art DNA isolation methods: a custom multi-component DNA isolation method developed for use in aerosol microbiome research published by Jiang et al. [28], and the commercial ZymoBIOMICS DNA Microprep Kit commonly used in environmental microbiome studies [38–41]. The performance of the three DNA isolation methods was evaluated using both a mock microbial community and real-world low biomass subway air samples. As part of this study, we also describe an end-to-end high-volume filter-based air sampling, filter processing and DNA isolation method, hereafter referred to as the “end-to-end MetaSUB method”. Since the MetaSUB method, when used as an integrated element of the end-to-end MetaSUB method, involves intermediate separation of the filter extract into a pellet (subjected to additional lysis) and supernatant fraction that is combined before final DNA purification, the relative contribution of the two fractions to the total DNA yield and observed aerosol microbiome profile was also evaluated using subway air samples.

## Methods

### MetaSUB method

The end-to-end MetaSUB method consists of an integrated air sampling, filter processing and DNA isolation scheme (Fig. 1). The method relies on the use of high-volume, battery-operated, hand-portable, electret filter-based air samplers that allow for flexible, user-adjustable sampling time and rapid change of sampling locations, which in turn provides support for high spatiotemporal resolution air (aerosol biomass) sampling campaigns. Following air sampling, the electret microfibrinous filter is

subjected to a liquid filter extraction procedure, after which the entire filter extract is processed to avoid the need for downstream filtration or centrifugation steps to reduce the sample volume prior to DNA isolation, which may compromise the sample integrity regarding both biomass and DNA yield and composition (diversity).

### Bioaerosol collection

Air (aerosol biomass) samples were collected with SASS3100 (Research International, Monroe, WA, USA), a high-volume electret microfibrinous filter-based air sampler. The air sampler was powered by UBI-2590 lithium-ion rechargeable batteries (Ultralife batteries, NY, USA), operated at a flowrate of 300 l of air per minute (LPM), and mounted on a tripod (~1.5 m above ground) with the inlet facing downward (45°) to avoid direct deposition of large particles. After sampling, the electret filters (Ø 44 mm) were stored in 50 ml polypropylene tubes at -80 °C until further processing.

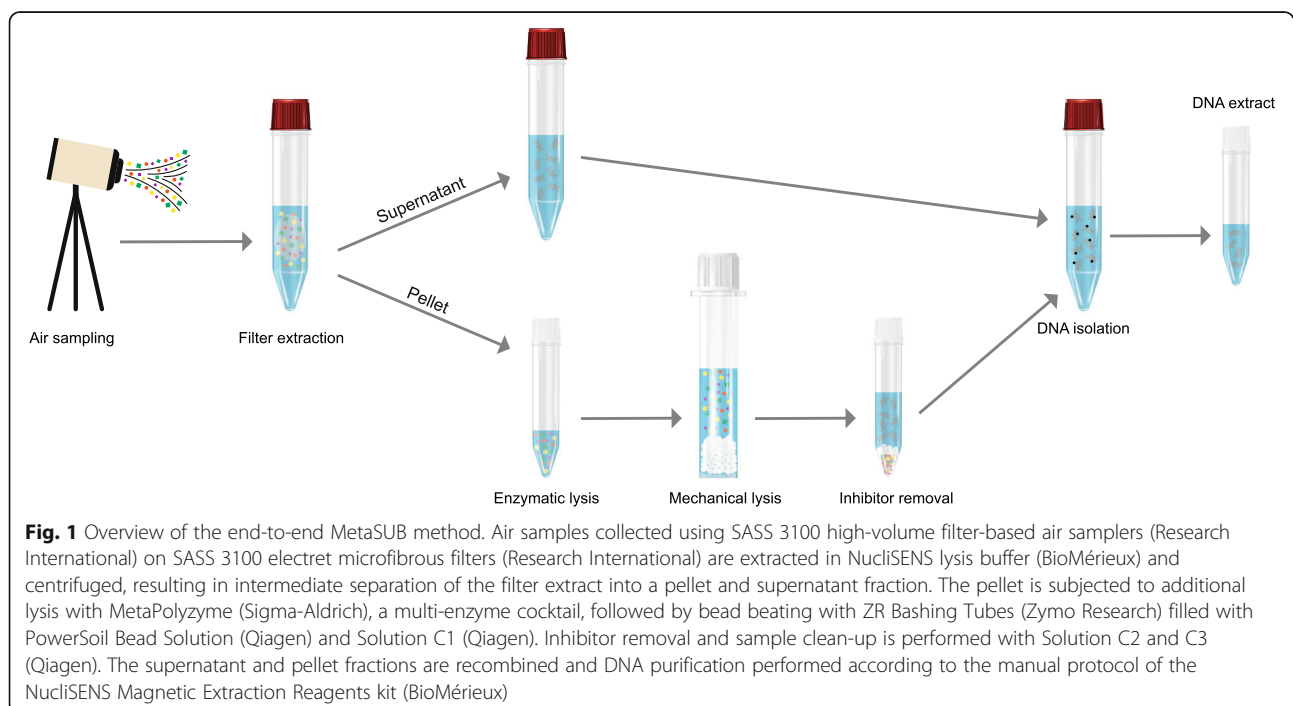
### Filter extraction

Liquid extraction of filter-collected aerosol biomass from the electret filters was performed by removing the filters from their housing and transferring them into 50 ml polypropylene tubes pre-loaded with 10 ml NucliSENS Lysis Buffer (BioMérieux, Marcy-l'Étoile, France). The sample tube was vortexed at maximum speed for 20 s before the filter was transferred into a 10 ml syringe with sterile forceps to extract residual liquid back into the sample tube before discarding the

filter. The sample tube was centrifuged (7000 x g, 30 min) and the supernatant transferred to a new 50 ml polypropylene tube (referred to as filter extract supernatant).

### DNA isolation

The pellet from the sample tube (referred to as filter extract pellet) was transferred to a polypropylene microcentrifuge tube with 1 ml PBS (pH 7.5, Sigma-Aldrich, St. Louis, MO, USA) and centrifuged (17,000 x g, 5 min). The resulting supernatant was carefully removed and combined with the filter extract supernatant. The pellet was dissolved in 150 µl PBS (pH 7.5). MetaPolzyme (Sigma-Aldrich), a multi-enzyme cocktail, was prepared by dissolving the enzyme powder in 1 ml PBS (pH 7.5), and 10 µl MetaPolzyme (5 mg/ml) and 5 µl sodium azide (0.1 M, Sigma-Aldrich) was added to the dissolved pellet sample. Enzymatic digestion was performed at 35 °C for 1 h in a Thermomixer (Eppendorf, Hamburg, Germany) at 1400 rpm. Subsequently, the sample was transferred to ZR BashingBead Lysis Tubes (0.1/0.5 mm beads, Zymo Research, Irvine, CA, USA) pre-filled with 550 µl PowerSoil Bead Solution (Qiagen, Hilden, Germany) and 60 µl Solution C1 (Qiagen). Bead tubes were subjected to bead beating (maximum intensity, 3 min) in a Mini Bead Beater-8 (BioSpec Products, Bartlesville, OK, USA). Bead tubes were centrifuged (13,000 x g, 2 min) and the supernatant treated with Solution C2 and C3 according to the DNeasy PowerSoil protocol





(Qiagen). The resulting supernatant was combined with the original filter extract supernatant before DNA purification. DNA was purified according to the manual protocol of the NucliSENS Magnetic Extraction Reagents kit (BioMérieux) with two modifications; magnetic silica suspension volume was increased to 90  $\mu$ l and incubation time was increased to 20 min. DNA samples were stored at  $-80^{\circ}\text{C}$  until further processing.

#### DNA isolation method described by Jiang et al. (Jiang method)

The custom, multi-component DNA isolation method (protocol steps 13–24) for air samples published by Jiang et al. [28] is based on the DNeasy PowerSoil Kit and AMPure XP magnetic bead separation. Jiang et al. introduced an incubation step in water bath ( $65^{\circ}\text{C}$ ) before bead vortexing, and found that magnetic bead capture recovered more DNA than standard PowerSoil spin columns. The DNA isolation method (protocol steps 13–24) published by Jiang et al. (hereafter referred to as “Jiang”) was used in this study with some minor modifications. Briefly, all samples were pretreated with MetaPolyzyme (as described for the MetaSUB method), before transfer to PowerBead tubes and continuation of DNA isolation according to the Jiang protocol.

#### ZymoBIOMICS DNA microprep kit (Zymobiomics method)

DNA isolation was performed according to the ZymoBIOMICS DNA Microprep Kit (Zymo Research) protocol (hereafter referred to as “Zymobiomics”) with some minor modifications. Briefly, all samples were pretreated with MetaPolyzyme (as described for the MetaSUB method) and bead beating was performed in a Mini Bead Beater-8 (BioSpec Products) for 3 min.

#### Performance evaluation using mock microbial community

The MetaSUB method was compared to the Jiang and Zymobiomics methods using a mock microbial community with a defined quantity and composition. The ZymoBIOMICS Microbial Community Standard (Zymo Research) contains ten microorganisms, eight bacteria (five Gram-positives and three Gram-negatives) and two yeasts. For each sample, the mock community (10  $\mu$ l), corresponding to a theoretical total DNA content of approximately 267 ng, was added to 140  $\mu$ l PBS (pH 7.5) and treated with MetaPolyzyme (as described for the MetaSUB method) before DNA isolation according to the three DNA isolation methods. Total DNA and 16S rRNA gene copy yields were measured for four sample pairs processed with MetaSUB ( $N=4$ ) and Jiang ( $N=4$ ) and six sample pairs processed with MetaSUB ( $N=6$ )

and Zymobiomics ( $N=6$ ). The within-sample differences in total DNA and 16S rRNA gene copy yields were evaluated with one-sample t-tests ( $H_0$ : difference = 0). All statistical analyses were performed in R (version 3.4.3, [www.R-project.org](http://www.R-project.org)). A subset of the mock community samples were subjected to SMS ( $N=12$ ): MetaSUB ( $N=4$ ), Jiang ( $N=4$ ), and Zymobiomics ( $N=4$ ).

#### Performance evaluation using subway air samples

The MetaSUB method was compared to the Jiang and Zymobiomics methods using subway air samples. Only the DNA isolation part of the end-to-end MetaSUB method was evaluated since the air sampling and filter-processing steps were used to collect and process subway air samples to generate equal aliquots of aerosol biomass for paired difference comparisons. An overview of the common sample processing steps and the three evaluated DNA isolation methods is given in Table 1. Air samples were collected for 1 h, corresponding to  $\sim 18\text{ m}^3$  of air sampled (60 min sampling at 300 LPM), during daytime hours at subway stations (Tøyen, Grønland, Stortinget, Nationaltheateret and Majorstuen) in Oslo, Norway, in the period between October 2017 and May 2018. The filter-collected samples were extracted in 10 ml NucliSENS lysis buffer and split into two equal filter extract aliquots. The aliquots were centrifuged ( $7000 \times g$ , 30 min) and only the pellet fractions were used for the comparison of DNA isolation methods. The supernatant fractions were subjected to DNA isolation separately (as described below) and used to investigate the distribution of DNA in the intermediate pellet and supernatant fractions of the MetaSUB method. For the DNA isolation method comparison, 24 air samples were split and the pellets processed with either MetaSUB ( $N=10$ ) and Jiang ( $N=10$ ) or MetaSUB ( $N=14$ ) and Zymobiomics ( $N=14$ ), to enable within-sample comparisons between the MetaSUB method and the two other methods. Since the supernatant fraction was not included in the MetaSUB method for the DNA isolation method comparison, 10 ml of fresh NucliSENS lysis buffer was used. Negative controls (reagents) were included for each DNA isolation method. Total DNA and 16S rRNA gene copy yields were examined and within-sample differences were evaluated with one-sample t-tests ( $H_0$ : difference = 0). All statistical analyses were performed in R (version 3.4.3, [www.R-project.org](http://www.R-project.org)). A subset of the subway air samples ( $N=6$ ) that had been split into two equal aliquots and processed with the three DNA isolation methods were subjected to SMS ( $N=12$ ): MetaSUB ( $N=3$ ) v. Jiang ( $N=3$ ) and MetaSUB ( $N=3$ ) v. Zymobiomics ( $N=3$ ). A negative control

**Table 1** Overview of the three DNA isolation methods evaluated in this work

Method	MetaSUB	Jiang	Zymbiomics
<b>Common processing steps</b> (used to generate equal aerosol biomass aliquots for paired difference comparison)			
Filter extraction (filter-to-liquid)	NucliSENS lysis buffer		
Lysis (enzymatic)	MetaPolyzyme multi-enzyme cocktail		
<b>Method-specific processing steps</b> (used for paired difference comparison on equal aerosol biomass aliquots)			
Lysis (mechanical)	ZR BashingBead Tubes with PowerSoil Bead Solution and Solution C1. Bead beating for 3 min	PowerSoil Bead Tubes with PowerSoil Bead Solution and Solution C1 incubated at 65 °C for 15 min. Bead vortexing for 15 min	ZR BashingBead Tubes with Zymbiomics lysis solution. Bead beating for 3 min
Inhibitor removal and sample clean-up	PowerSoil Solution C2 and C3	PowerSoil Solution C2 and C3	Zymo-Spin IV and Zymo-Spin IV-μHRC Columns
DNA purification	NucliSENS magnetic beads	AMPure XP magnetic beads	Zymo-Spin IC-Z Column

(reagents) for each DNA isolation method was also subjected to SMS ( $N = 3$ ).

#### DNA distribution in intermediate pellet/supernatant fractions of the MetaSUB method

The filter extraction procedure of the MetaSUB method generates two intermediate fractions (pellet and supernatant) that are usually recombined before the final DNA purification (Fig. 1). Differences in total DNA and 16S rRNA gene copy yields between pellet ( $N = 24$ ) and supernatant ( $N = 24$ ) fractions were therefore investigated. DNA was isolated from the supernatant fractions from subway air samples (described above) with the NucliSENS Magnetic Extraction Reagents kit as described for the MetaSUB method. Furthermore, to identify potential differences in DNA composition (diversity) between the pellet and supernatant fractions, DNA isolated with the MetaSUB method from six paired pellet and supernatant fractions ( $N = 12$ ) and one negative control (reagents;  $N = 1$ ) was subjected to SMS.

#### Quantification of total DNA and 16S rRNA gene copies

Total DNA was quantified with Qubit dsDNA HS assays (Life Technologies, Carlsbad, CA, USA) on a Qubit 3.0 Fluorimeter (Life Technologies). Bacterial 16S rRNA gene copies were determined with a 16S rRNA gene qPCR assay performed according to Liu et al. [42] on a LightCycler 480 instrument (Roche Diagnostics, Oslo, Norway). Serial dilutions of *Escherichia coli* DNA (seven 16S rRNA gene copies per genome) were used to generate a standard curve.

#### Shotgun metagenomic sequencing (SMS)

DNA isolated from mock community samples were subjected to SMS (150 bp paired-end) multiplexed on

a MiSeq (~24–30 M paired-end reads, Illumina, San Diego, CA, USA). Library preparation was done with the Nextera DNA Flex kit (Illumina) according to the recommended protocol. DNA isolated from subway air samples were subjected to SMS (150 bp paired-end) multiplexed on one lane (~80–130 M paired-end reads) on a HiSeq 3000 (Illumina). Library preparation was done with the ThruPLEX DNA-Seq kit (Takara Bio, Mountain View, CA, USA) according to the recommended protocol and 18 amplification cycles. Raw sequence reads were demultiplexed, quality trimmed (Trim Galore, v0.4.3;  $\geq Q20$ ,  $\geq 50$  bp) and underwent adapter removal (Cutadapt, v1.16), before analysis on the One Codex platform with default settings [43]. One Codex taxonomic feature tables were imported into R and analyzed in the phyloseq package [44].

All sequence reads not taxonomically assigned to the species level were removed from the 12 mock community samples. Since the aim was to gauge the relative contribution of the ten bacterial and fungal species in the mock community across the three DNA isolation methods, non-target features were binned as “other”. The comparison was made by plotting normalized abundances across all 12 samples.

For the six subway air samples that were split into equal aliquots and processed with the three DNA isolation methods, MetaSUB ( $N = 3$ ) v. Jiang ( $N = 3$ ) and MetaSUB ( $N = 3$ ) v. Zymbiomics ( $N = 3$ ), all taxonomic features not assigned to the genus or species level, along with human reads, were removed. Prevalent features reported in the negative control samples (> 1% of within-sample reads, four in total, accounting for 94.5% of all reads in the negative controls) were stripped from the entire dataset before removing the negative controls. The cleaned samples varied in the number of assigned reads, ranging from 1,160,976 to 5,530,138. After examining the effect of rarefaction

on the  $\alpha$ -diversity measures “Observed”, “Shannon”, and “Simpson” (Additional file 1: Figure S1), all samples were rarefied to the lowest common depth (1160976).

The six paired pellet and supernatant fractions from subway air samples processed with the MetaSUB method underwent the same procedure: removing features not assigned to the genus or species level, along with human reads, and prevalent features in the negative control (12 features, accounting for 99.3% of all reads in the negative control). The effect of rarefaction was evaluated (Additional file 1: Figure S2), and all samples were rarefied to the lowest common depth (453218).

The cleaned SMS datasets were divided into six groups corresponding to the three comparisons (MetaSUB v. Jiang, MetaSUB v. Zymobiomics, and MetaSUB pellet v. supernatant) before summarizing the top phyla, families, genera and species within each group. Taxonomic features with species-level assignment were extracted for analyses of within-sample diversity ( $\alpha$ -diversity: “Observed”, “Shannon”, “Simpson”), where relevant groups were compared by fitting linear models. All features (read counts) were conglomerated to the genus level for analyses of among sample differences ( $\beta$  diversity); Bray Curtis distances were ordinated with PCoA and analyzed with MetaSUB/Jiang, MetaSUB/Zymobiomics, and pellet/supernatant, as predictors in separate PERMANOVA tests. Distance estimation and PERMANOVA was performed with *vegan* (v.2.6.0, <https://github.com/vegandevs/vegan/>). Sample clustering was visualized with PCoA ordination. MegaBLAST analysis of forward reads against the NCBI non-redundant nucleotide database, followed by taxonomic binning using the native lowest common ancestor (LCA) algorithm in MEGAN6 [45], was used to perform a cross-kingdom analysis on the pellet/supernatant samples. Lastly, random forest classification models were performed, using 10,001 trees, with MetaSUB/Jiang, MetaSUB/Zymobiomics, and pellet/supernatant, as response variable and One Codex (species-level) taxonomic features as predictor variables. Separate tests using 501 trees and 1000 permutations were performed to evaluate statistical significance. The random forest models were built using *randomForest* [46].

#### Accession numbers

The sequence data has been deposited in the NCBI Sequence Read Archive under Bioproject ID# PRJNA542423 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA542423>).

## Results

### Performance evaluation using mock microbial community

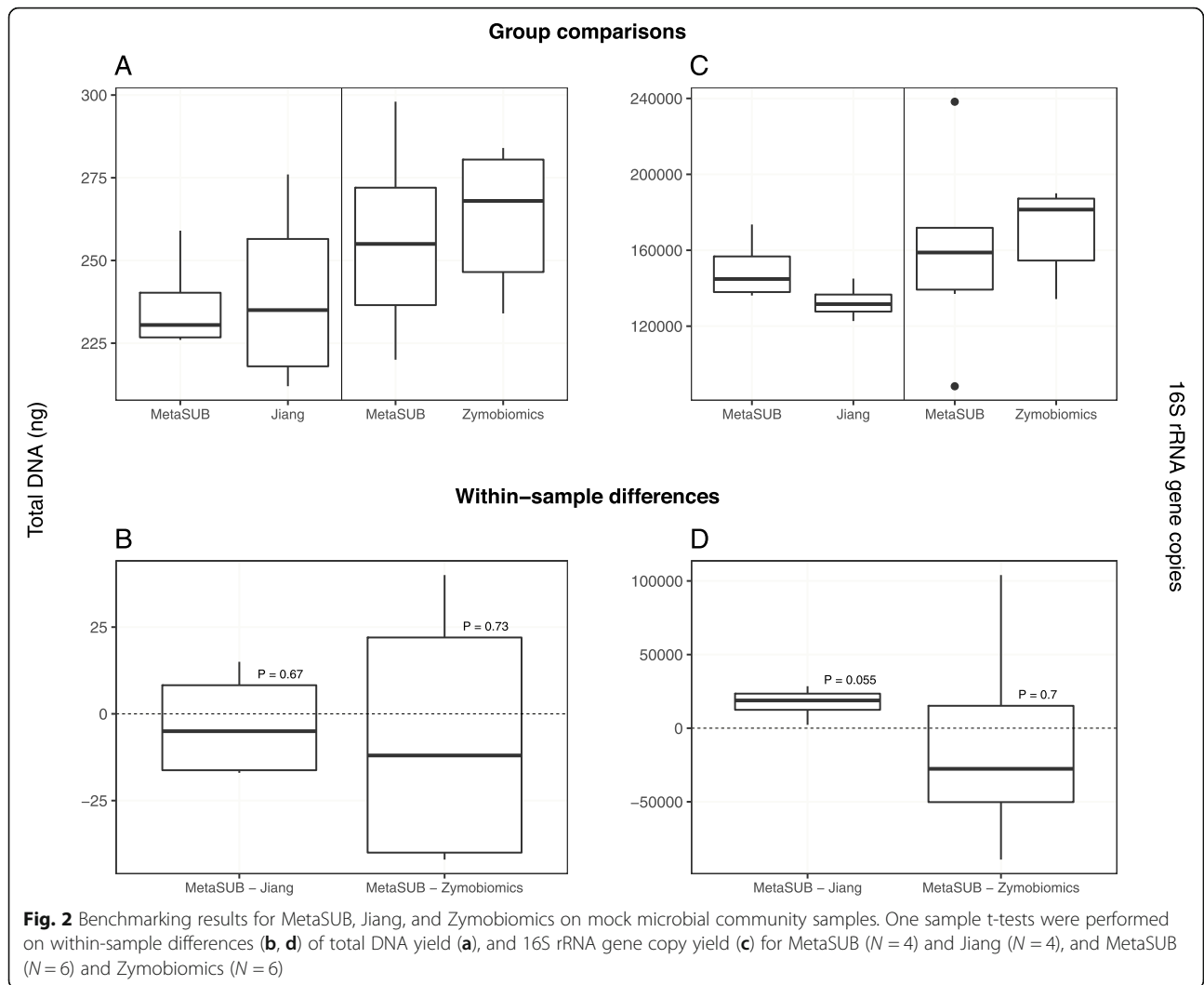
The total DNA and 16S rRNA gene copy yields from mock community samples showed no significant

differences between the MetaSUB method and the other two methods (Fig. 2; Table 2a). However, the MetaSUB method obtained a higher 16S rRNA gene copy yield than Jiang with borderline significance ( $P = 0.055$ ; Fig. 2; Table 2a). The 12 mock community samples that were subjected to SMS showed similar distributions of all ten microbial species in the mock community across the three methods, with MetaSUB and Zymobiomics being nearly identical (Fig. 3).

### Performance evaluation using subway air samples

The total DNA and 16S rRNA gene copy yields from subway air samples showed that the MetaSUB method obtained significantly higher total DNA and 16S rRNA gene copy yields than both Jiang and Zymobiomics (all  $P < 0.001$ ; Fig. 4; Table 2b).

All subway air samples reached saturation with regard to  $\alpha$ -diversity at the lowest common assigned read depth (1,160,976, Additional file 1: Figure S1), which was the depth at which all samples were rarefied to. Taxonomic distributions at the family level were highly similar between the samples processed with MetaSUB and Zymobiomics (Fig. 5). The samples processed with MetaSUB and Jiang were also highly similar, but a skew was observed in the relative abundances for two of the three Jiang samples (Fig. 5). In the MetaSUB v. Zymobiomics comparison, the top ten most abundant phyla were identical between the method pairs, but not identical in their ordering by abundance (Table 3). Of the top ten families, one was uniquely found in the MetaSUB results (*Staphylococcaceae*; lowest abundance) and one in the Zymobiomics results (*Rhodobacteraceae*; second lowest abundance; Table 3). Among the ten top genera, only two were unique for MetaSUB (*Hymenobacter* and *Staphylococcus*) and two for Zymobiomics (*Dietzia* and *Paracoccus*; Table 3). Among the top ten species in each group, only one was unique to MetaSUB (*Chlorogloea* sp. CCALA 695) and one to Zymobiomics (*Lecanicillium* sp. LEC01; Table 3). In the MetaSUB v. Jiang comparison, there were more pronounced differences. The top ten phyla were not identical; *Acidobacteria* was only found in the MetaSUB results and *Planctomycetes* only in the Jiang results (Table 4). The top ten families were identical (but not in ordering); however, Jiang reported a substantially higher relative abundance of the family that was most abundant for both methods (*Micrococcaceae*, MetaSUB: 14% and Jiang: 25.6%; Table 4). Among the ten top genera, two were unique for MetaSUB (*Corynebacterium* and *Hymenobacter*) and two for Jiang (*Dietzia* and *Marmoricola*; Table 4). Here, the most abundant genus in Jiang (*Micrococcus*: 11.7%) was not the most abundant in

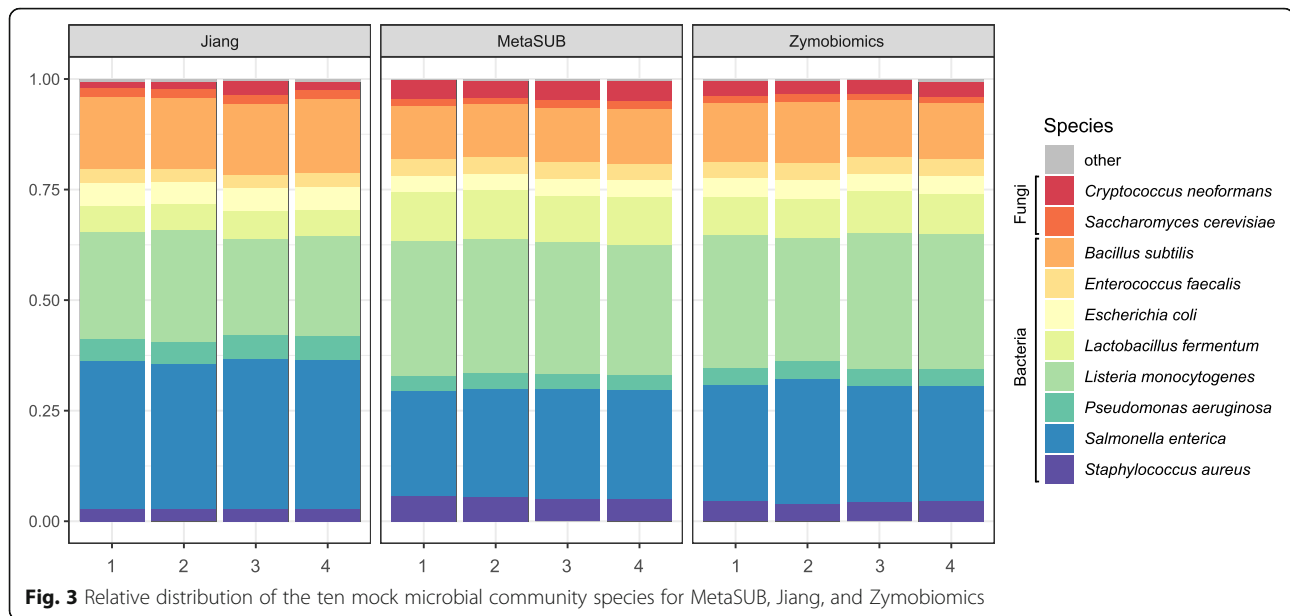


**Table 2** Benchmarking results for MetaSUB, Jiang, and Zymbiomics on mock microbial community and subway air samples

Measure	Within-sample differences	Est	95% CI	T	df	P	
A) Mock microbial community	Total DNA yield (ng)	MetaSUB - Jiang	-3	[-28.5, 22.5]	-0.37	3	0.73
		MetaSUB - Zymbiomics	-7	[-46.5, 32.5]	-0.45	5	0.67
	16S rRNA gene copy yield (copies)	MetaSUB - Jiang	17,107	[- 634, 34,848]	3.07	3	0.055
		MetaSUB - Zymbiomics	-11,452	[- 83,155, 60,251]	-0.41	5	0.70
B) Subway air samples	Total DNA yield (ng)	MetaSUB - Jiang	1.07	[0.77, 1.37]	8.01	9	< 0.001
		MetaSUB - Zymbiomics	1.35	[0.86, 1.85]	5.94	13	< 0.001
	16S rRNA gene copy yield (copies)	MetaSUB - Jiang	5046	[3882, 6211]	9.80	9	< 0.001
		MetaSUB - Zymbiomics	3451	[1741, 5162]	4.36	13	< 0.001

One-sample t-test on within-sample differences ( $H_0$ : difference in within-sample measurements = 0) for different method pairs with mock microbial community (A) and subway air samples (B). Measures from Jiang/Zymbiomics were subtracted from the MetaSUB measures: the estimate (*est*) gives the departure from zero of the resultant values (larger than zero values indicate that MetaSUB had a higher yield than Jiang/Zymbiomics)





MetaSUB (second most abundant; 5.60%) Among the top ten species in each group, only five species were present in both MetaSUB and Jiang results (Table 4).

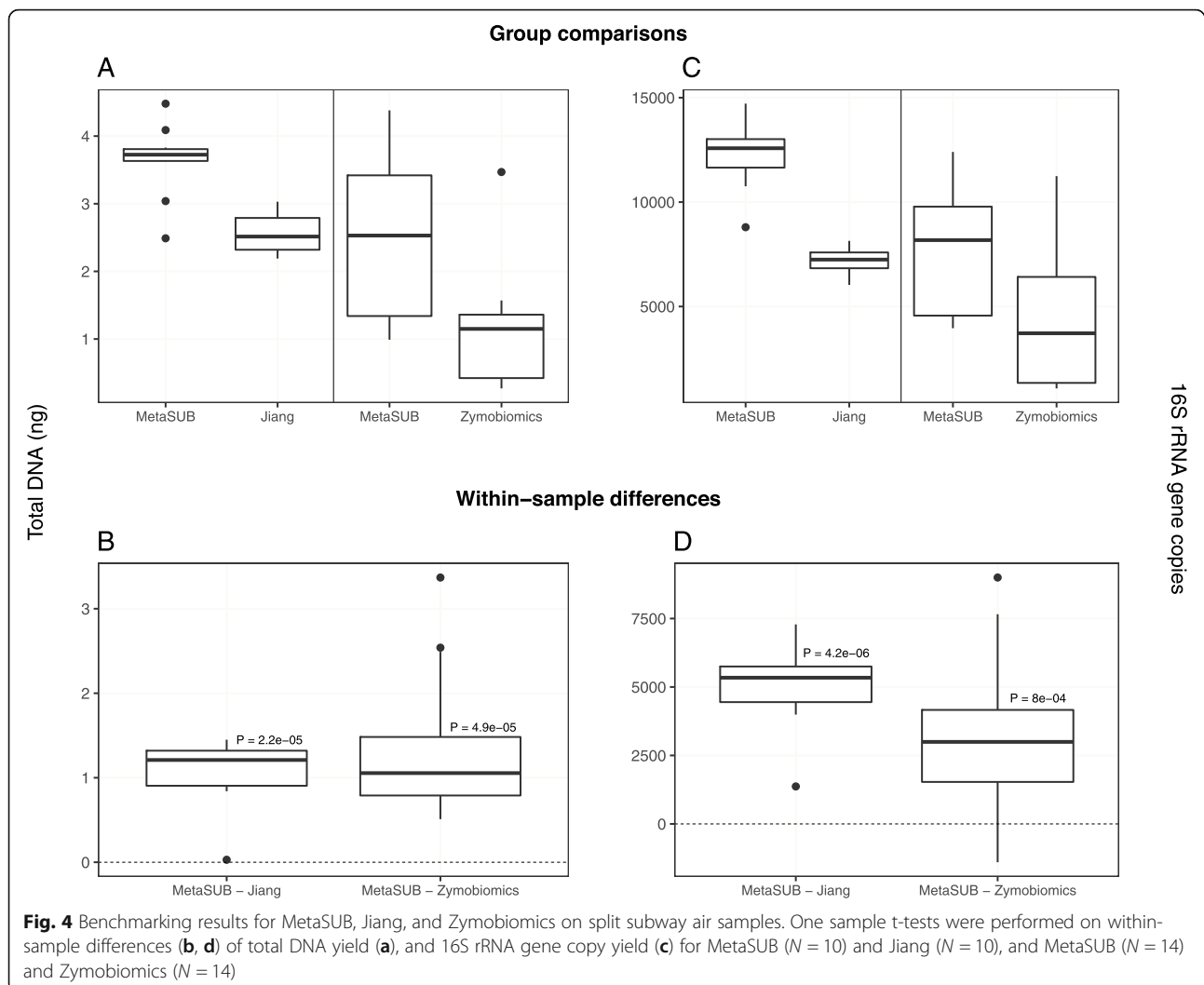
Linear regression of within-sample  $\alpha$ -diversity indices showed that MetaSUB reported significantly higher diversity estimates compared to Zymobiomics (Observed:  $est = 734.3$ ,  $P = 0.01$ ; Shannon:  $est = 0.22$ ,  $P = 0.002$ ; Simpson:  $est = 0.00079$ ,  $P = 0.001$ ; Fig. 6), but no differences were shown between MetaSUB and Jiang  $\alpha$ -diversity estimates (Observed:  $est = 6531$ ; Shannon:  $est = 2.75$ ; Simpson:  $est = 0.028$ ; all  $P > 0.12$ ; Fig. 6). PERMANOVA tests of PCoA ordinated Bray Curtis distances found no significant differences among MetaSUB and Jiang ( $P = 0.1$ ) or MetaSUB and Zymobiomics ( $P = 0.1$ ; Fig. 7).

The random forest classification analysis, where species-level features were scored by their ability to correctly classify the DNA isolation method used, had a perfect out-of-bag error of 0%, and a significant permutation test ( $P < 0.02$ ) for MetaSUB v. Zymobiomics. For MetaSUB v. Jiang, the classification model had an out-of-bag error of 16%, but also here the permutation test was significant ( $P = 0.01$ ). For MetaSUB v. Zymobiomics, the proportions of archaea, bacteria and fungi across the dataset and in the 100 species most important for correctly classifying samples as either MetaSUB or Zymobiomics were highly similar. However, for MetaSUB v. Jiang, 6.0% of all assigned species were fungi, while among the 100 species most important for classification, 20 were fungi. These 20 fungal species all had higher abundances in the MetaSUB results (Additional file 1: Figure S4). The top 30 most important features for both classification models are shown in Fig. 8.

#### DNA distribution in intermediate pellet/supernatant fractions of the MetaSUB method

The distribution of DNA in terms of both amount and composition (diversity) in the intermediate pellet and supernatant fractions of the MetaSUB method was investigated by separately isolating DNA from the two fractions from subway air samples. The results revealed that the supernatant fraction contained  $42\% \pm 6$  of the total DNA yield and  $32\% \pm 12$  of the total 16S rRNA gene copy yield (Additional file 1: Figure S2).

Rarefaction plots of pellet and supernatant samples indicated that  $\alpha$ -diversity indices (particularly Shannon and Simpson) reached saturation before the lowest common assigned read depth (453,218, Additional file 1: Figure S2), which was the depth at which all samples were rarefied to. The taxonomic distributions in pellet and supernatant samples were largely similar (Table 5; Fig. 9). The top ten phyla were identical in the pellet and supernatant group, but not identical in their ordering by abundance (Table 5). Of the top ten families, one was uniquely found in the pellet group (*Rhodobacteraceae*; second lowest abundance) and one only in the supernatant group (*Deinococcaceae*; lowest abundance; Table 5). Among the ten top genera, only one was unique for the pellet group (*Marmoricola*) and one for the supernatant group (*Deinococcus*; Table 5). Among the top ten species in each group, seven species were present in both (Table 5). Linear regression of within-sample  $\alpha$ -diversity indices revealed no significant differences between pellet and supernatant samples (Fig. 10; all  $P > 0.38$ ). A PERMANOVA test of PCoA ordinated Bray Curtis distances found that whether



samples were pellet or supernatant explained 51.7% of the among-sample variance in diversity (Fig. 11;  $P = 0.004$ ).

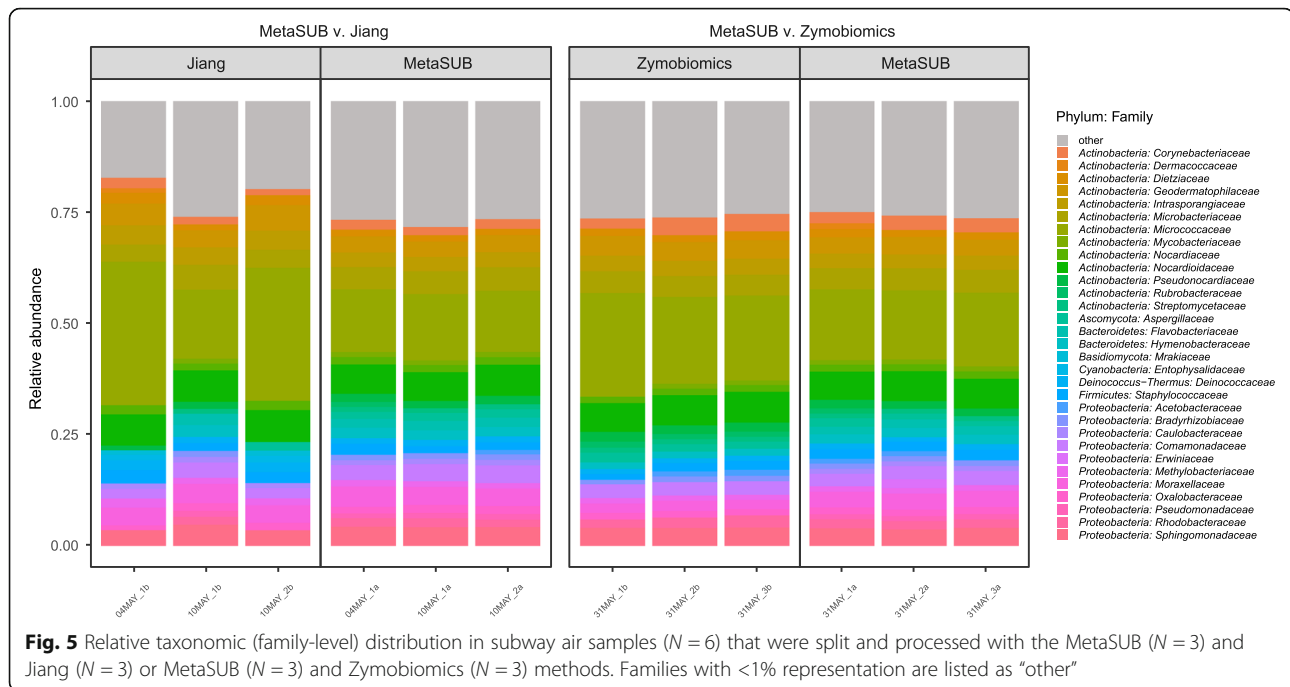
The cross-kingdom analysis revealed substantial differences in the relative representation of almost all examined groups (archaea [domain], bacteria, fungi, plants, human, and other animals) between the pellet and supernatant samples (Fig. 12). While very few reads were assigned to archaea, only pellet samples had any coverage within this group. Pellet samples also had a higher relative number of assigned reads across all sample pairs within bacteria and fungi. The supernatant had a higher relative number of reads assigned as human and other animals, while plants saw similar representation in pellet and supernatant samples.

The random forest classification analysis, where species-level features were scored by their ability to correctly classify the pellet and supernatant groups, had a

perfect out-of-bag error of 0%, and the permutation test was statistically significant ( $P > 0.001$ ). In the entire dataset, 6.0% of the features were assigned as fungi and 0.3% were assigned as archaea, while among the 100 species with the highest variable importance in our classification model, 56 were fungi and two were archaea. Among the top 50 species, 30 were fungi and one archaea. The top 30 most important features are shown in Fig. 13.

## Discussion

Here, we have demonstrated a new custom, multi-component DNA isolation method (“the MetaSUB method”) optimized for SMS-based aerosol microbiome research. By processing the entire filter extract, in combination with thorough chemical, enzymatic and mechanical lysis and DNA purification using magnetic beads, the MetaSUB method drastically



improves the DNA yield from low biomass air samples and reduces the risk of introducing microbiome profile bias. Comprehensive performance benchmarking of the MetaSUB method against two other state-of-the-art DNA isolation methods was done with both a mock microbial community and real-world subway air samples. The benchmarking revealed that the MetaSUB method obtains significantly higher DNA yields from subway air samples than the other two methods, which is an important performance parameter for successful implementation of SMS on low biomass air samples. SMS of subway air samples revealed that the MetaSUB method reported higher diversity than Zymbiomics, and gave better representation of certain fungal species than Jiang. All three DNA isolation methods performed similarly well on mock microbial community samples, both in terms of DNA yield and community representation. As part of this study, we have also described an end-to-end air sampling, filter processing and DNA isolation method (“the end-to-end MetaSUB method”) optimized for SMS-based aerosol microbiome research. The end-to-end MetaSUB method relies on the use of SASS 3100 high-volume electret microfibrinous filter-based air samplers and was shown to be capable of recovering sufficient DNA yields from short-duration subway air samples, which corresponded to  $\sim 9\text{ m}^3$  of air sampled (30 min sampling at 300 LPM) in this study, to facilitate high temporal resolution SMS-based aerosol microbiome investigations.

The performance evaluation of the three DNA isolation methods (MetaSUB, Jiang and Zymbiomics) revealed no significant differences regarding total DNA and 16S rRNA gene copy yields when isolating DNA from mock microbial community samples (Fig. 2). Furthermore, SMS of mock community samples showed that the three methods gave highly similar representation of the ten microbial species present in the mock community (Fig. 3). However, on subway air samples, the MetaSUB method outperformed both Jiang and Zymbiomics regarding total DNA and 16S rRNA gene copy yields (Fig. 4). SMS analyses of subway air samples that had been split and isolated with either MetaSUB and Jiang or MetaSUB and Zymbiomics revealed significantly higher  $\alpha$ -diversity estimates for MetaSUB compared to Zymbiomics (Fig. 6). One of the three samples processed with Jiang showed higher  $\alpha$ -diversity than all three MetaSUB samples, while the other two Jiang samples showed substantially lower diversity estimates (Fig. 6), which rendered the comparison against MetaSUB non-significant for all  $\alpha$ -diversity indices. We have no conclusive explanation for this pattern; however, we observed that the two low-scoring Jiang samples had high duplicate sequence read proportions (62.4 and 71.8%) compared to all other samples (average: 18.6%), and postulate that the variable performance may be related to the recovery of insufficient DNA yields from two of the Jiang samples to allow for reliable SMS. Furthermore, the random forest classification analysis indicates that the Jiang method does not

**Table 3** Abundant microbial taxa in subway air samples (MetaSUB v. Zymbiomics method)

MetaSUB v. Zymbiomics							
Phylum	MetaSub		abundance	Phylum	Zymbiomics		Abundance
	prevalence				prevalence		
	mean	total			mean	total	
<i>Actinobacteria</i>	2.7	9271	51.61%	<i>Actinobacteria</i>	2.7	9301	57.72%
<i>Proteobacteria</i>	2.3	14,269	27.02%	<i>Proteobacteria</i>	2.2	13,421	23.67%
<i>Firmicutes</i>	1.6	3250	4.76%	<i>Ascomycota</i>	2.5	1899	4.91%
<i>Bacteroidetes</i>	2.0	2999	4.74%	<i>Basidiomycota</i>	2.4	622	4.75%
<i>Ascomycota</i>	2.4	1843	4.42%	<i>Firmicutes</i>	1.3	2705	3.30%
<i>Basidiomycota</i>	2.3	606	3.84%	<i>Bacteroidetes</i>	1.5	2257	2.77%
<i>Cyanobacteria</i>	2.3	526	1.35%	<i>Deinococcus-Thermus</i>	2.6	204	1.26%
<i>Deinococcus-Thermus</i>	2.5	196	1.23%	<i>Cyanobacteria</i>	1.9	432	0.66%
<i>Euryarchaeota</i>	1.7	450	0.62%	<i>Euryarchaeota</i>	1.5	393	0.58%
<i>Acidobacteria</i>	2.5	112	0.08%	<i>Acidobacteria</i>	2.6	114	0.08%
Family	prevalence		abundance	Family	prevalence		Abundance
	mean	total			mean	total	
<i>Micrococcaceae</i>	2.8	801	15.72%	<i>Micrococcaceae</i>	2.8	822	20.34%
<i>Nocardiodaceae</i>	2.8	308	6.51%	<i>Nocardiodaceae</i>	2.9	311	6.65%
<i>Microbacteriaceae</i>	2.6	1222	4.89%	<i>Microbacteriaceae</i>	2.6	1235	4.70%
<i>Sphingomonadaceae</i>	2.8	1006	3.88%	<i>Geodermatophilaceae</i>	3.0	150	4.24%
<i>Geodermatophilaceae</i>	3.0	150	3.67%	<i>Sphingomonadaceae</i>	2.8	1013	4.04%
<i>Moraxellaceae</i>	1.9	608	3.54%	<i>Intrasporangiaceae</i>	3.0	211	3.46%
<i>Intrasporangiaceae</i>	3.0	211	3.15%	<i>Corynebacteriaceae</i>	2.6	571	3.35%
<i>Comamonadaceae</i>	2.6	800	2.94%	<i>Comamonadaceae</i>	2.7	838	3.02%
<i>Corynebacteriaceae</i>	2.6	568	2.93%	<i>Rhodobacteraceae</i>	2.6	1484	2.28%
<i>Staphylococcaceae</i>	2.3	401	2.17%	<i>Moraxellaceae</i>	1.6	533	2.14%
Genus	prevalence		abundance	Genus	prevalence		Abundance
	mean	total			mean	total	
<i>Micrococcus</i>	2.9	29	7.27%	<i>Micrococcus</i>	2.9	29	9.34%
<i>Arthrobacter</i>	2.8	399	5.32%	<i>Arthrobacter</i>	2.9	413	6.44%
<i>Nocardioides</i>	2.8	176	3.51%	<i>Nocardioides</i>	2.9	177	3.57%
<i>Sphingomonas</i>	2.9	461	3.09%	<i>Kocuria</i>	2.7	79	3.44%
<i>Corynebacterium</i>	2.6	568	2.93%	<i>Corynebacterium</i>	2.6	571	3.35%
<i>Psychrobacter</i>	2.0	141	2.85%	<i>Sphingomonas</i>	2.9	464	3.23%
<i>Blastococcus</i>	3.0	60	2.44%	<i>Blastococcus</i>	3.0	60	2.92%
<i>Staphylococcus</i>	2.2	308	2.02%	<i>Psychrobacter</i>	2.7	184	1.82%
<i>Kocuria</i>	2.7	79	2.02%	<i>Dietzia</i>	3.0	51	1.70%
<i>Hymenobacter</i>	2.9	98	1.92%	<i>Paracoccus</i>	2.9	167	1.58%
Species	prevalence		abundance	Species	prevalence		Abundance
	mean	total			mean	total	
<i>Micrococcus luteus</i>	3.0	3	1.13%	<i>Micrococcus luteus</i>	3.0	3	1.44%
<i>Arthrobacter sp. H41</i>	3.0	3	1.00%	<i>Arthrobacter sp. H41</i>	3.0	3	1.38%
<i>Rubrobacter aplysinae</i>	3.0	3	0.78%	<i>Rubrobacter aplysinae</i>	3.0	3	0.97%
<i>Arthrobacter sp. Leaf234</i>	3.0	3	0.67%	<i>Arthrobacter sp. Leaf234</i>	3.0	3	0.89%
<i>Marmoricola sp. Leaf446</i>	3.0	3	0.65%	<i>Stereum hirsutum</i>	3.0	3	0.76%



**Table 3** Abundant microbial taxa in subway air samples (MetaSUB v. Zymbiomics method) (Continued)

<i>Chlorogloea</i> sp. CICALA 695	3.0	3	0.63%	<i>Marmoricola</i> sp. Leaf446	3.0	3	0.76%
<i>Deinococcus marmoris</i>	3.0	3	0.60%	<i>Fomitopsis pinicola</i>	3.0	3	0.68%
<i>Stereum hirsutum</i>	3.0	3	0.50%	<i>Blastococcus</i> sp. DSM 44268	3.0	3	0.64%
<i>Blastococcus</i> sp. DSM 44268	3.0	3	0.48%	<i>Deinococcus marmoris</i>	3.0	3	0.57%
<i>Fomitopsis pinicola</i>	3.0	3	0.45%	<i>Lecanicillium</i> sp. LECO1	3.0	3	0.57%

Top ten microbial phyla, families, genera and species in subway air samples ( $N = 3$ ) that were split and processed with the MetaSUB ( $N = 3$ ) and Zymbiomics ( $N = 3$ ) methods

produce the same representation for certain fungal species as the MetaSUB method, since out of the 100 most important species for distinguishing between MetaSUB and Jiang processed samples, 20 were fungal, while across the entire dataset, only 6% of the species were fungal. All of these 20 fungal species had higher representation in MetaSUB samples (Additional file 1: Figure S4).

Our findings highlight the importance of benchmarking DNA isolation methods with both mock communities and real-world samples since the complexity found in the real-world environment is not easily recreated. The observed DNA yield differences among the three methods can probably be attributed to a combination of sub-process efficiency differences, since the methods rely on different combinations of lysis (chemical, enzymatic, and/or mechanical), inhibitor removal and sample clean-up, and DNA purification (magnetic beads and silica spin filters) principles (Table 1). During customization of DNA isolation methods it is therefore important to keep in mind that even subtle procedural differences, including choice of bead solution, intensity and time settings for the bead beating process [47, 48], and different enzyme combinations, may have a large effect on the ultimate biomass lysis efficiency [31]. By replacing the spin columns in the PowerSoil Kit with AMPure XP Beads (magnetic bead purification), Jiang et al. [28] observed a three-fold increase in DNA yield. The multi-component MetaSUB method was developed by adopting and customizing sub-processes from several different DNA isolation methods in an effort to ensure maximized DNA recovery and comprehensive biomass lysis. Note that for the performance benchmarking of DNA isolation methods in this study, only the intermediate pellet fraction of the MetaSUB method was used to facilitate an equal comparison between the three different DNA isolation methods (Fig. 1). The intermediate supernatant fraction would normally also be included in the MetaSUB method and would have constituted approximately 72% of additional DNA, thereby making the DNA yield differences even more pronounced.

Since the filter extraction procedure in the MetaSUB method produces intermediate pellet and supernatant fraction that are combined before DNA purification, we investigated differences in DNA amount and composition (diversity) between the two fractions in an effort to better understand the benefit of including supernatants (i.e., increased DNA yield) and the risk of not including them (i.e., microbiome profile bias). The observed microbial diversity in paired pellet and supernatant samples was highly similar at the phylum (Table 5), family (Table 5; Fig. 9), genus (Table 5) and species (Table 5) levels. Note, however, with direct examination of only the most abundant taxonomic groups in Table 5 and Fig. 9, the similarities do not necessarily extend to groups with low abundance. While we did not find any differences among the pellet and supernatant samples in  $\alpha$ -diversity (Fig. 10), which describes within-sample diversity, there was significant diversity nested among samples, of which the pellet/supernatant grouping explained 51.7% (Fig. 11). The cross-kingdom analyses revealed differences in the taxonomic composition of pellet and supernatant samples (Fig. 12). While human DNA constituted a relatively large proportion of eukaryotic reads, it did not account for all of the difference observed among pellet and supernatant samples within this kingdom; on average, human reads constituted 18% of assigned reads in pellets and 42% in supernatants based on the cross-kingdom analysis (Fig. 12). Human reads reported by One Codex also had a higher relative abundance in supernatants (31 and 67% of assigned reads in pellets and supernatants, respectively). Features assigned as archaea were exclusively observed in pellets; however, caution should be used when interpreting these results, since only eleven features were assigned to this kingdom. The random forest classification model revealed that fungi were particularly important in separating pellet and supernatant samples, especially when accounting for the relatively low representation of fungi across all samples. A recent study by Mbareche et al. has shown that the use of traditional processing methods, e.g., filter extract processing where the supernatant fraction is discarded after a centrifugation step, may lead to an underrepresentation of fungi [49]. In conclusion, concerning the most abundant

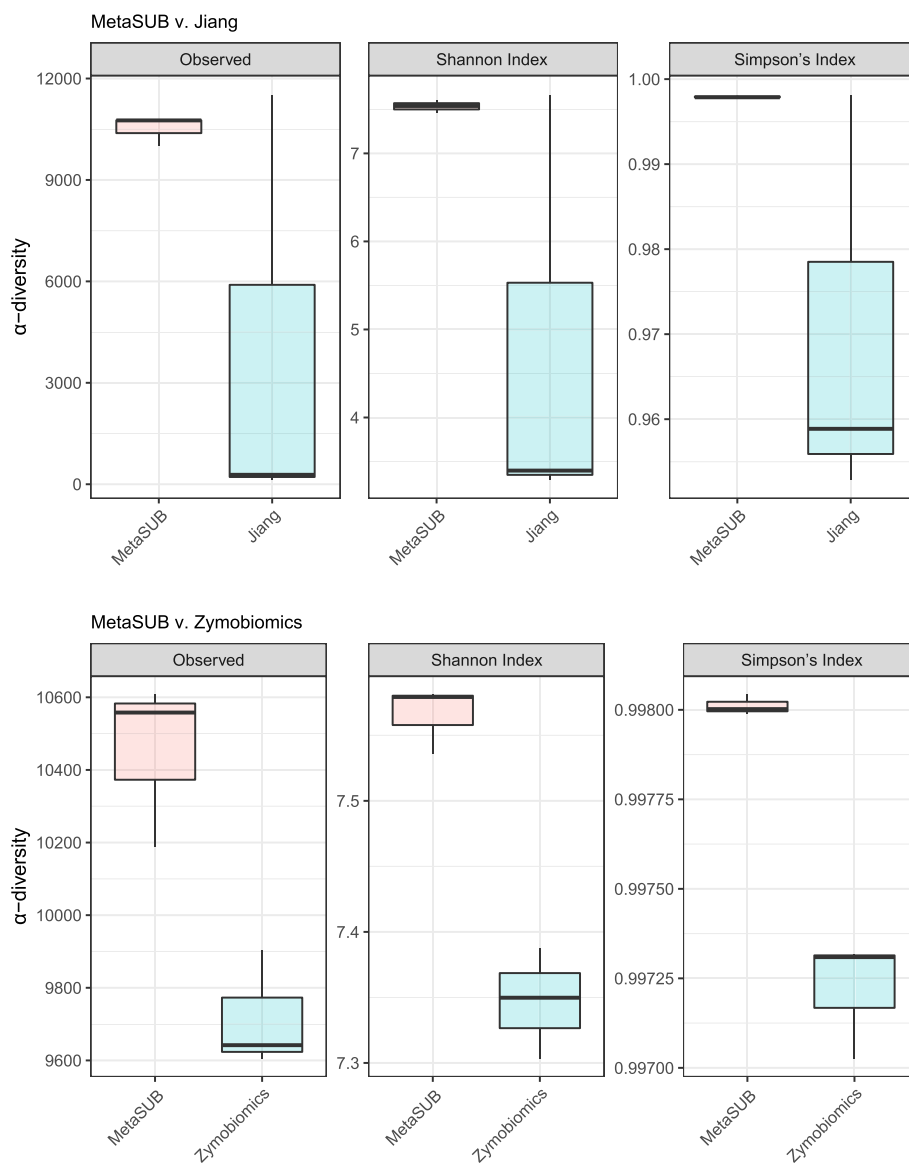
**Table 4** Abundant microbial taxa in subway air samples (MetaSUB v. Jiang method)

MetaSUB v. Jiang							
Phylum	MetaSub		abundance	Phylum	Jiang		
	prevalence				prevalence	Abundance	
	mean	total			mean	total	
<i>Actinobacteria</i>	2.7	9248	48.69%	<i>Actinobacteria</i>	1.1	3625	60.08%
<i>Proteobacteria</i>	2.3	14,379	28.58%	<i>Proteobacteria</i>	1.0	6090	23.40%
<i>Ascomycota</i>	2.4	1843	5.28%	<i>Firmicutes</i>	0.7	1510	4.58%
<i>Bacteroidetes</i>	2.1	3148	5.10%	<i>Bacteroidetes</i>	0.9	1436	3.40%
<i>Firmicutes</i>	1.6	3250	4.52%	<i>Ascomycota</i>	1.1	822	2.78%
<i>Basidiomycota</i>	2.3	598	3.95%	<i>Deinococcus-Thermus</i>	1.0	77	1.94%
<i>Cyanobacteria</i>	2.3	518	1.44%	<i>Basidiomycota</i>	1.1	286	1.91%
<i>Deinococcus-Thermus</i>	2.4	190	1.35%	<i>Cyanobacteria</i>	1.1	243	1.50%
<i>Euryarchaeota</i>	1.9	495	0.63%	<i>Euryarchaeota</i>	0.9	231	0.15%
<i>Acidobacteria</i>	2.6	113	0.08%	<i>Planctomycetes</i>	1.2	52	0.04%
Family	prevalence		abundance	Family	prevalence		
	mean	total			mean	total	
<i>Micrococcaceae</i>	2.7	785	13.96%	<i>Micrococcaceae</i>	1.2	336	25.63%
<i>Nocardioideae</i>	2.8	301	6.58%	<i>Nocardioideae</i>	1.1	123	7.02%
<i>Microbacteriaceae</i>	2.6	1213	5.05%	<i>Geodermatophilaceae</i>	1.3	65	4.77%
<i>Sphingomonadaceae</i>	2.8	1010	4.20%	<i>Microbacteriaceae</i>	1.0	486	4.45%
<i>Moraxellaceae</i>	2.0	655	4.12%	<i>Intrasporangiaceae</i>	1.4	97	4.15%
<i>Comamonadaceae</i>	2.4	750	3.68%	<i>Moraxellaceae</i>	0.9	289	4.05%
<i>Geodermatophilaceae</i>	3.0	149	3.59%	<i>Sphingomonadaceae</i>	1.0	357	3.92%
<i>Intrasporangiaceae</i>	3.0	210	3.23%	<i>Comamonadaceae</i>	1.1	337	2.60%
<i>Hymenobacteraceae</i>	2.9	182	2.17%	<i>Staphylococcaceae</i>	0.9	169	2.39%
<i>Flavobacteriaceae</i>	2.2	1594	2.04%	<i>Dietziaceae</i>	1.1	19	1.99%
Genus	prevalence		abundance	Genus	prevalence		
	mean	total			mean	total	
<i>Arthrobacter</i>	2.7	391	5.63%	<i>Micrococcus</i>	2.1	21	11.69%
<i>Micrococcus</i>	2.8	28	5.60%	<i>Arthrobacter</i>	1.1	153	9.34%
<i>Nocardioides</i>	2.7	167	3.50%	<i>Kocuria</i>	1.2	34	3.53%
<i>Psychrobacter</i>	1.9	131	3.40%	<i>Psychrobacter</i>	0.9	64	3.46%
<i>Sphingomonas</i>	2.9	465	3.38%	<i>Nocardioides</i>	1.0	61	3.38%
<i>Blastococcus</i>	3.0	59	2.30%	<i>Sphingomonas</i>	1.0	155	3.20%
<i>Corynebacterium</i>	2.6	559	2.04%	<i>Blastococcus</i>	1.3	26	3.16%
<i>Hymenobacter</i>	2.9	98	2.02%	<i>Marmoricola</i>	2.0	12	2.78%
<i>Staphylococcus</i>	2.0	289	1.77%	<i>Staphylococcus</i>	0.9	128	2.33%
<i>Kocuria</i>	2.7	79	1.76%	<i>Dietzia</i>	1.1	19	1.99%
Species	prevalence		abundance	Species	prevalence		
	mean	total			mean	total	
<i>Arthrobacter sp. H41</i>	3.0	3	1.14%	<i>Arthrobacter sp. H41</i>	3.0	3	2.36%
<i>Micrococcus luteus</i>	3.0	3	0.88%	<i>Arthrobacter sp. Leaf234</i>	3.0	3	2.15%
<i>Chlorogloea sp. CICALA 695</i>	3.0	3	0.74%	<i>Marmoricola sp. Leaf446</i>	3.0	3	1.71%
<i>Rubrobacter aplysiniae</i>	3.0	3	0.71%	<i>Deinococcus marmoris</i>	3.0	3	1.56%
<i>Arthrobacter sp. Leaf234</i>	3.0	3	0.71%	<i>Blastococcus sp. DSM 44268</i>	3.0	3	1.24%

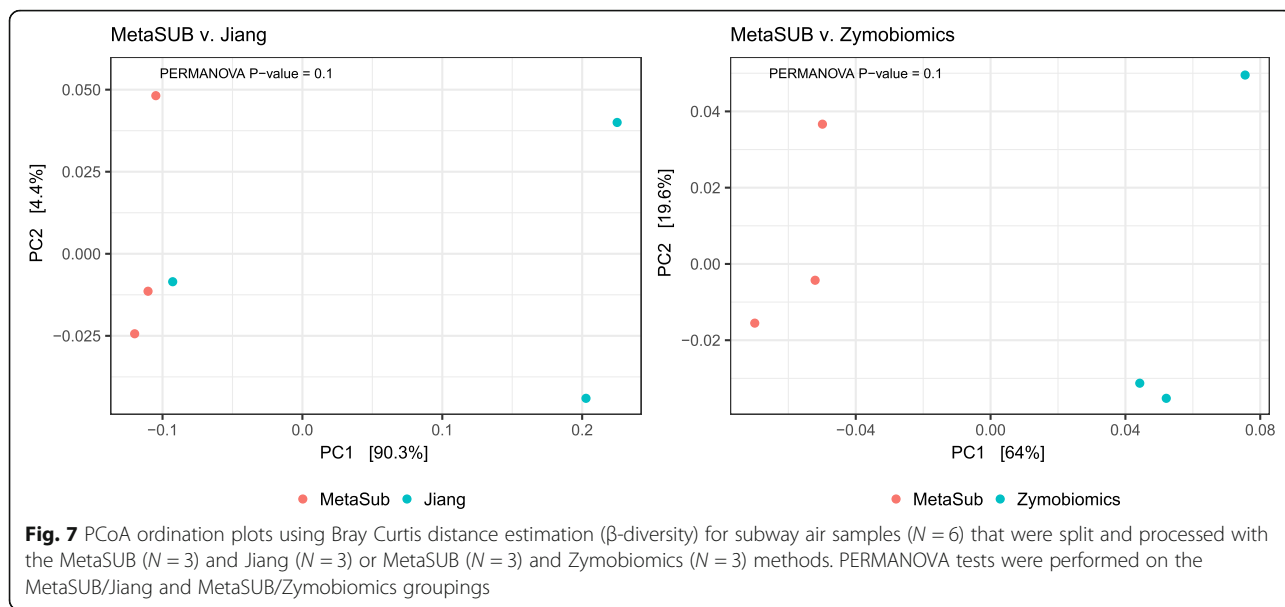
**Table 4** Abundant microbial taxa in subway air samples (MetaSUB v. Jiang method) (Continued)

<i>Aspergillus sp. MA 6041</i>	3.0	3	0.69%	<i>Arthrobacter agilis</i>	3.0	3	1.23%
<i>Marmoricola sp. Leaf446</i>	3.0	3	0.67%	<i>Chlorogloea sp. CCALA 695</i>	3.0	3	1.14%
<i>Deinococcus marmoris</i>	3.0	3	0.66%	<i>Marmoricola scoriae</i>	3.0	3	0.84%
<i>Acidovorax temperans</i>	3.0	3	0.61%	<i>Janibacter sp. Soil728</i>	3.0	3	0.82%
<i>Stereum hirsutum</i>	3.0	3	0.57%	<i>Mrakia frigida</i>	3.0	3	0.77%

Top ten microbial phyla, families, genera and species in subway air samples (N = 3) that were split and processed with the MetaSUB (N = 3) and Jiang (N = 3) methods



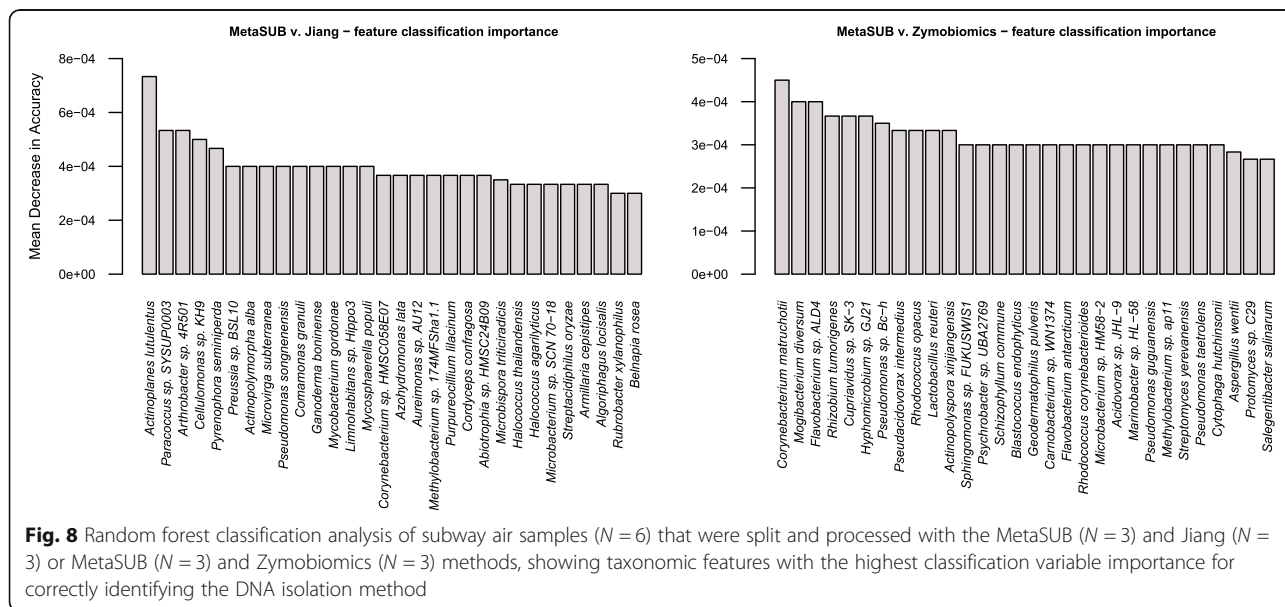
**Fig. 6** Comparison of diversity estimates ( $\alpha$ -diversity) for subway air samples (N = 6) that were split and processed with the MetaSUB (N = 3) and Jiang (N = 3) or MetaSUB (N = 3) and Zymobiomics (N = 3) methods



microbial groups and within-sample diversity estimates, there is little difference between the pellet and supernatant fractions. However, the between-sample diversity analyses show that potentially important diversity may be lost if the entire filter extract is not processed, and that an appreciable amount of this diversity is nested in fungi. In addition, a more general but potentially important reason for processing the entire filter extract in the context of high-volume filter-collected air samples is the variable resistance different types of microorganisms have against sampling-associated stress factors. While stress-resistant microorganisms may be relatively unaffected by sampling-associated stress, stress-sensitive organisms, e.g.

Gram-negative bacteria, may become membrane-impaired, ruptured or even completely lysed due to sampling-associated desiccation during high-volume dry filter collection and subsequent osmotic shock during liquid filter extraction. DNA that becomes liberated from membrane-impaired, ruptured or lysed microorganisms will generally not be recovered by standard centrifugation or filtration processes intended for intact organism capture, and may therefore remain in the supernatant or filtrate fraction.

Taken together, the demonstrated performance of the MetaSUB method, including drastically improved DNA yield from subway air samples and reduced risk





**Table 5** Abundant microbial taxa in pellet and supernatant fractions from subway air samples (MetaSUB method)

Phylum	Supernatant			Phylum	Pellet		
	Prevalence mean	Total	abundance		Prevalence mean	Total	abundance
<i>Actinobacteria</i>	3.5	9479	51.21%	<i>Actinobacteria</i>	4.0	10,653	53.53%
<i>Proteobacteria</i>	3.1	14,453	28.87%	<i>Proteobacteria</i>	3.4	15,894	24.50%
<i>Bacteroidetes</i>	2.8	3139	7.10%	<i>Basidiomycota</i>	3.7	717	5.42%
<i>Firmicutes</i>	2.5	3816	4.23%	<i>Ascomycota</i>	4.2	2109	5.23%
<i>Deinococcus-Thermus</i>	3.2	184	3.06%	<i>Bacteroidetes</i>	2.9	3202	4.44%
<i>Basidiomycota</i>	1.5	290	1.86%	<i>Firmicutes</i>	2.6	3874	2.68%
<i>Ascomycota</i>	2.4	1231	1.62%	<i>Deinococcus-Thermus</i>	3.7	214	2.05%
<i>Cyanobacteria</i>	2.8	431	1.47%	<i>Cyanobacteria</i>	3.6	557	1.03%
<i>Euryarchaeota</i>	1.5	260	0.17%	<i>Euryarchaeota</i>	2.1	367	0.66%
<i>Acidobacteria</i>	3.4	103	0.10%	<i>Acidobacteria</i>	4.4	132	0.11%
Family	Prevalence		abundance	Family	Prevalence		abundance
	mean	Total			mean	Total	
<i>Micrococcaceae</i>	3.6	735	11.27%	<i>Micrococcaceae</i>	3.9	795	14.70%
<i>Nocardiodaceae</i>	3.1	236	9.32%	<i>Nocardiodaceae</i>	3.0	222	8.26%
<i>Microbacteriaceae</i>	3.4	921	5.24%	<i>Geodermatophilaceae</i>	2.9	97	4.65%
<i>Sphingomonadaceae</i>	3.8	1119	5.20%	<i>Microbacteriaceae</i>	3.5	958	4.53%
<i>Geodermatophilaceae</i>	3.7	121	4.60%	<i>Sphingomonadaceae</i>	4.3	1260	4.22%
<i>Moraxellaceae</i>	2.9	794	4.14%	<i>Intrasporangiaceae</i>	3.8	199	4.20%
<i>Hymenobacteraceae</i>	3.5	150	3.93%	<i>Moraxellaceae</i>	2.6	717	2.77%
<i>Intrasporangiaceae</i>	3.7	192	3.81%	<i>Corynebacteriaceae</i>	4.2	881	2.61%
<i>Corynebacteriaceae</i>	3.7	789	3.77%	<i>Rhodobacteraceae</i>	4.3	1698	2.33%
<i>Deinococcaceae</i>	4.0	117	3.03%	<i>Hymenobacteraceae</i>	3.8	163	2.15%
Genus	Prevalence		abundance	Genus	Prevalence		abundance
	mean	Total			mean	Total	
<i>Arthrobacter</i>	3.3	320	6.31%	<i>Arthrobacter</i>	3.7	357	7.74%
<i>Sphingomonas</i>	3.7	512	4.40%	<i>Micrococcus</i>	3.9	39	4.19%
<i>Nocardioides</i>	3.0	127	3.87%	<i>Sphingomonas</i>	4.1	560	3.40%
<i>Hymenobacter</i>	3.4	89	3.85%	<i>Nocardioides</i>	2.9	120	3.14%
<i>Corynebacterium</i>	3.8	767	3.73%	<i>Blastococcus</i>	3.5	39	3.13%
<i>Psychrobacter</i>	2.5	143	3.24%	<i>Corynebacterium</i>	4.2	854	2.59%
<i>Deinococcus</i>	4.0	117	3.03%	<i>Marmoricola</i>	5.8	23	2.56%
<i>Friedmanniella</i>	6.0	18	2.85%	<i>Friedmanniella</i>	6.0	18	2.38%
<i>Blastococcus</i>	4.1	45	2.60%	<i>Psychrobacter</i>	2.3	133	2.37%
<i>Micrococcus</i>	4.6	46	2.31%	<i>Hymenobacter</i>	3.2	84	2.04%
Species	Prevalence		abundance	Species	Prevalence		abundance
	mean	Total			mean	Total	
<i>Deinococcus marmoris</i>	6.0	6	2.07%	<i>Arthrobacter sp. H41</i>	6.0	6	2.07%
<i>Arthrobacter sp. Leaf234</i>	6.0	6	1.91%	<i>Micrococcus luteus</i>	6.0	6	1.96%
<i>Marmoricola sp. Leaf446</i>	6.0	6	1.31%	<i>Rubrobacter aplysinae</i>	6.0	6	1.52%
<i>Friedmanniella flava</i>	6.0	6	1.05%	<i>Arthrobacter sp. Leaf234</i>	6.0	6	1.51%
<i>Friedmanniella sagamiharensis</i>	6.0	6	1.03%	<i>Marmoricola sp. Leaf446</i>	6.0	6	1.42%
<i>Cutibacterium acnes</i>	6.0	6	1.02%	<i>Blastococcus sp. DSM 44268</i>	6.0	6	1.32%

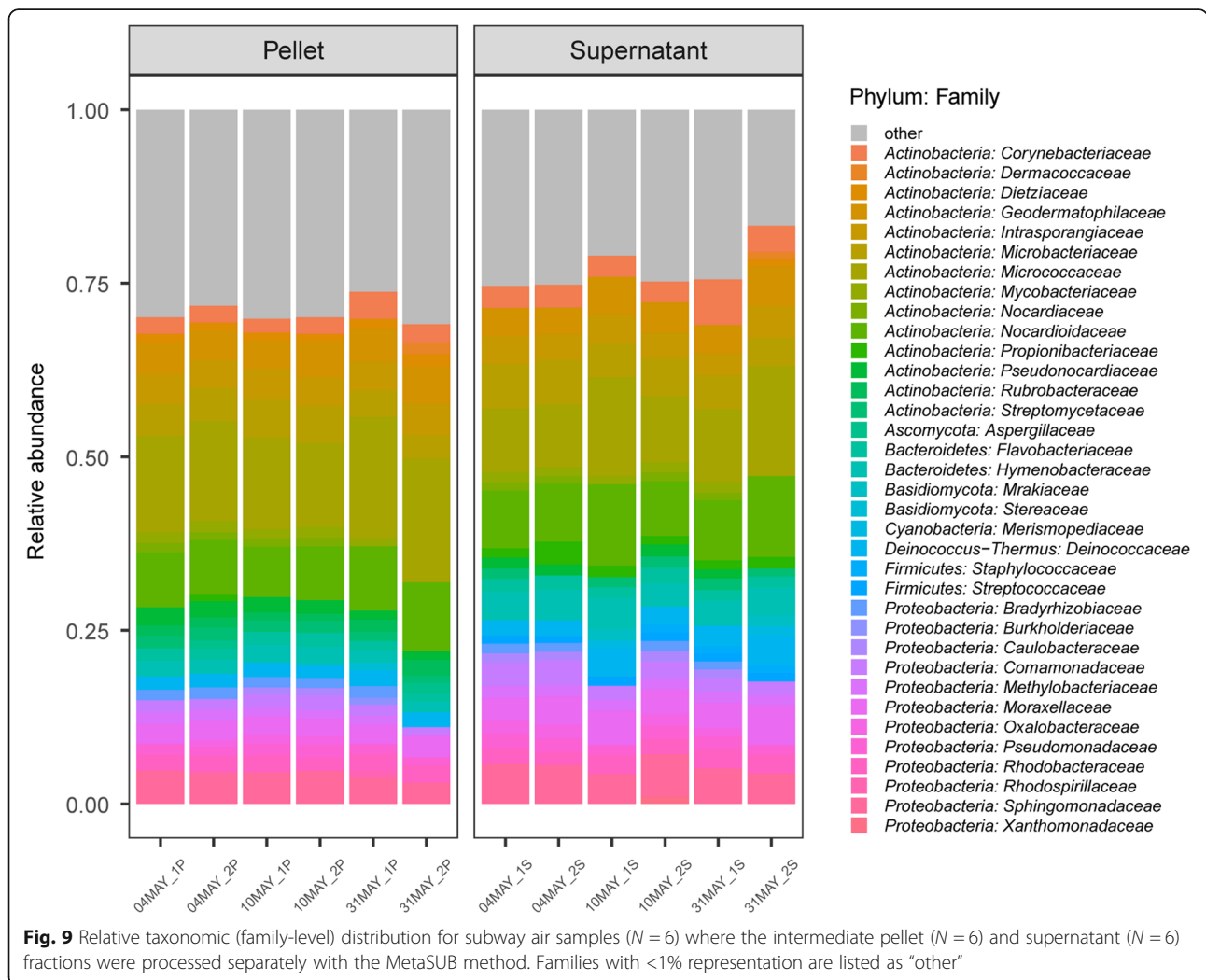
**Table 5** Abundant microbial taxa in pellet and supernatant fractions from subway air samples (MetaSUB method) (Continued)

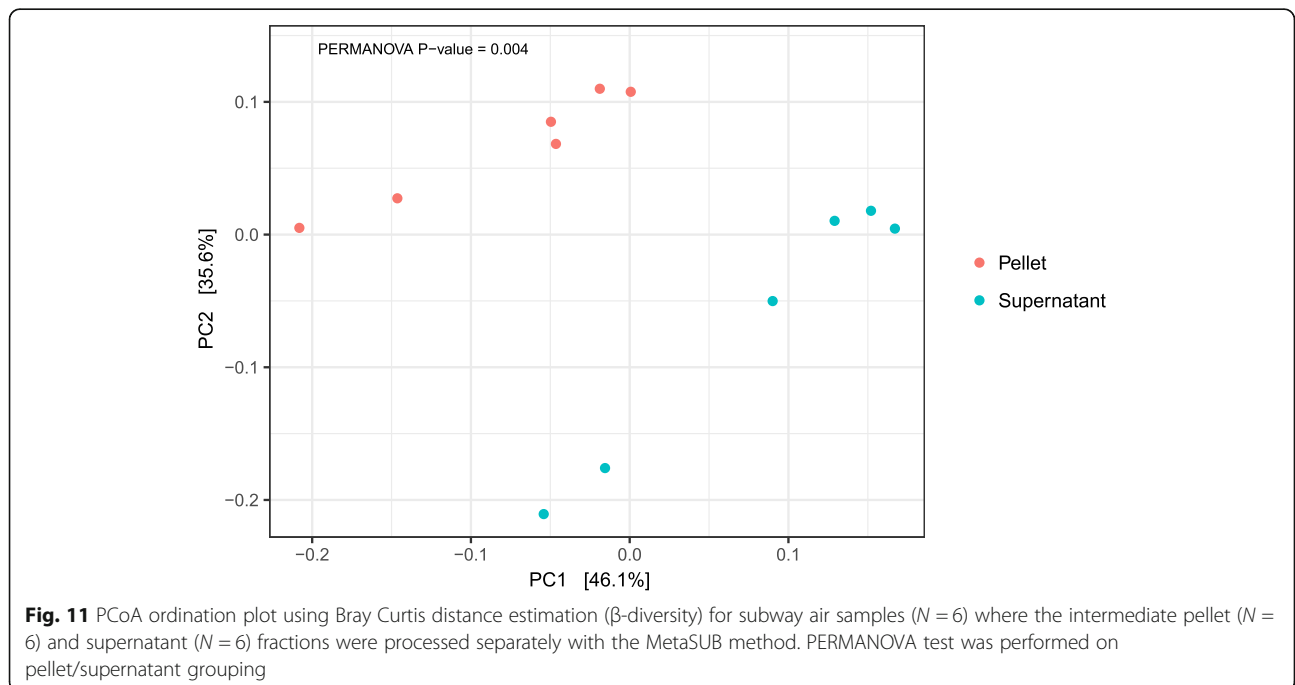
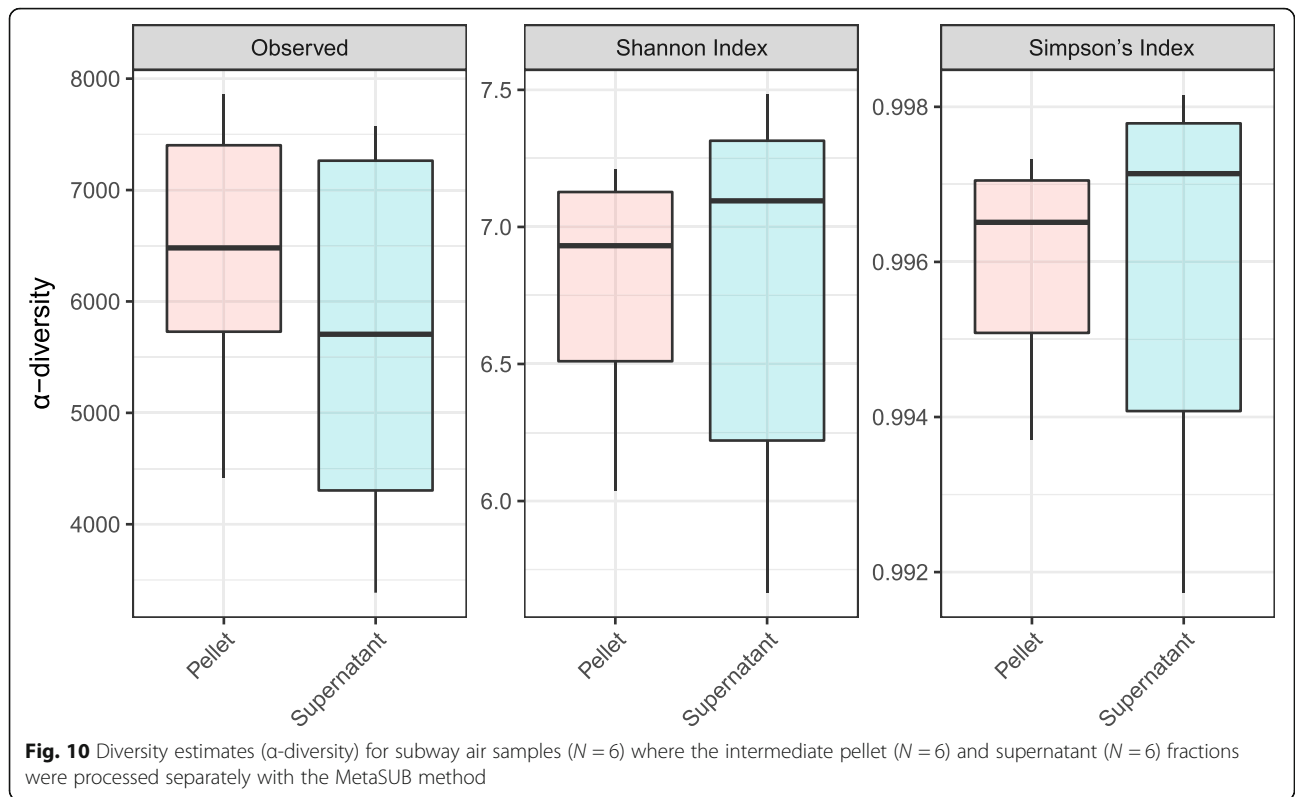
<i>Mrakia frigida</i>	6.0	6	1.02%	<i>Deinococcus marmoris</i>	6.0	6	1.27%
<i>Blastococcus</i> sp. DSM 44268	6.0	6	1.02%	<i>Arthrobacter agilis</i>	6.0	6	0.95%
<i>Micrococcus luteus</i>	6.0	6	1.01%	<i>Friedmanniella flava</i>	6.0	6	0.89%
<i>Arthrobacter</i> sp. H41	6.0	6	0.98%	<i>Stereum hirsutum</i>	6.0	6	0.86%

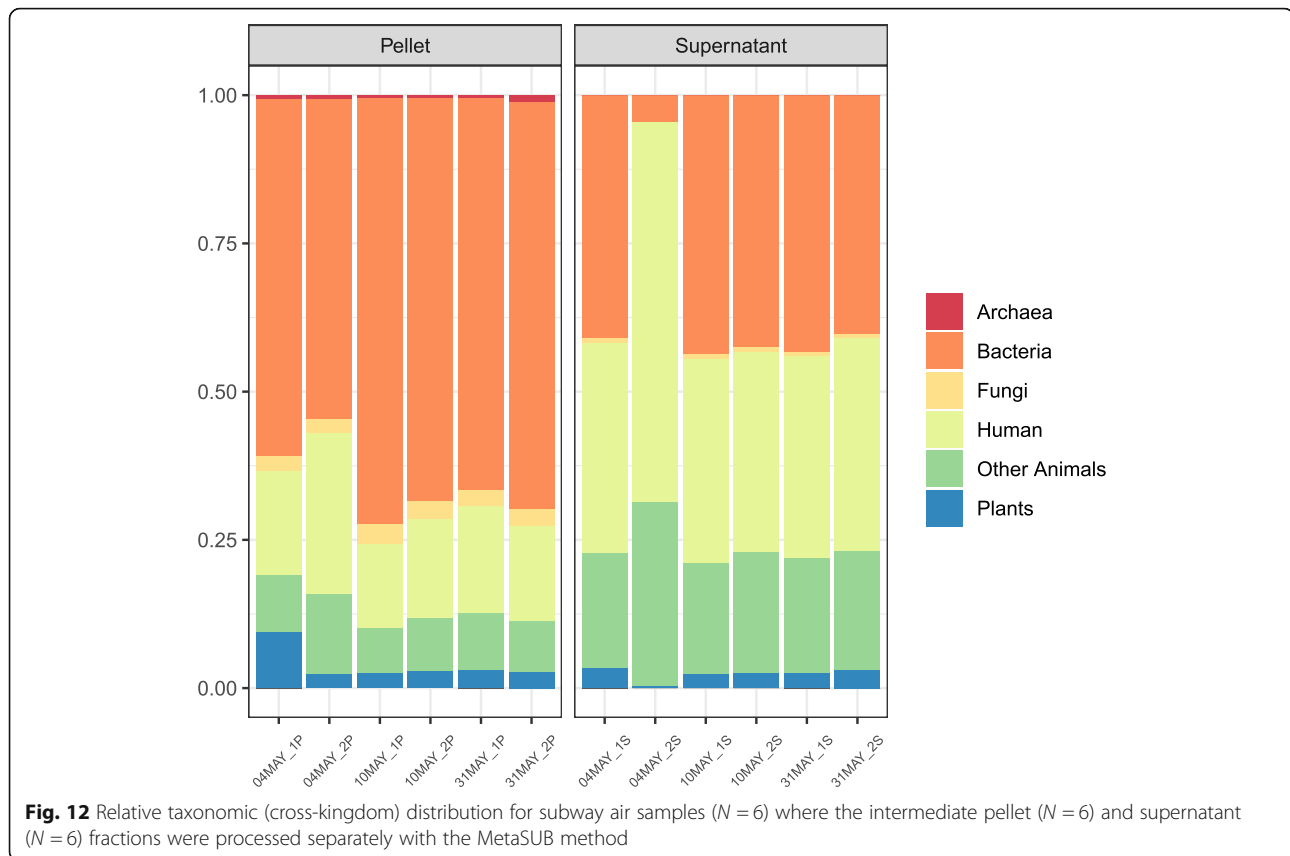
Top ten microbial phyla, families, genera and species in the intermediate pellet (N = 6) and supernatant (N = 6) fractions from subway air samples (N = 6) processed with the MetaSUB method

of microbiome profile bias, highlights the benefit of isolating DNA from the entire filter extract. However, the need for isolating DNA from a relatively large sample volume, a 10 ml filter extract in this work, limits the available selection of out-of-the-box commercial DNA isolation kits and introduces a customization need to ensure reliable performance regarding thorough comprehensive biomass lysis, sufficient inhibitor removal and sample clean-up, and efficient DNA recovery. The custom, multi-component

MetaSUB method is therefore a relatively hands-on (manual), labor-intensive DNA isolation method compared to many out-of-the-box commercial DNA isolation kits. However, an experienced operator can perform the MetaSUB method, including all processing and incubation steps, in approximately 3 hours, while the estimated total processing time for 12 air samples is approximately 4 hours. Furthermore, even without considering the associated benefits of isolating DNA from the entire filter extract, the use of a custom, multi-component DNA







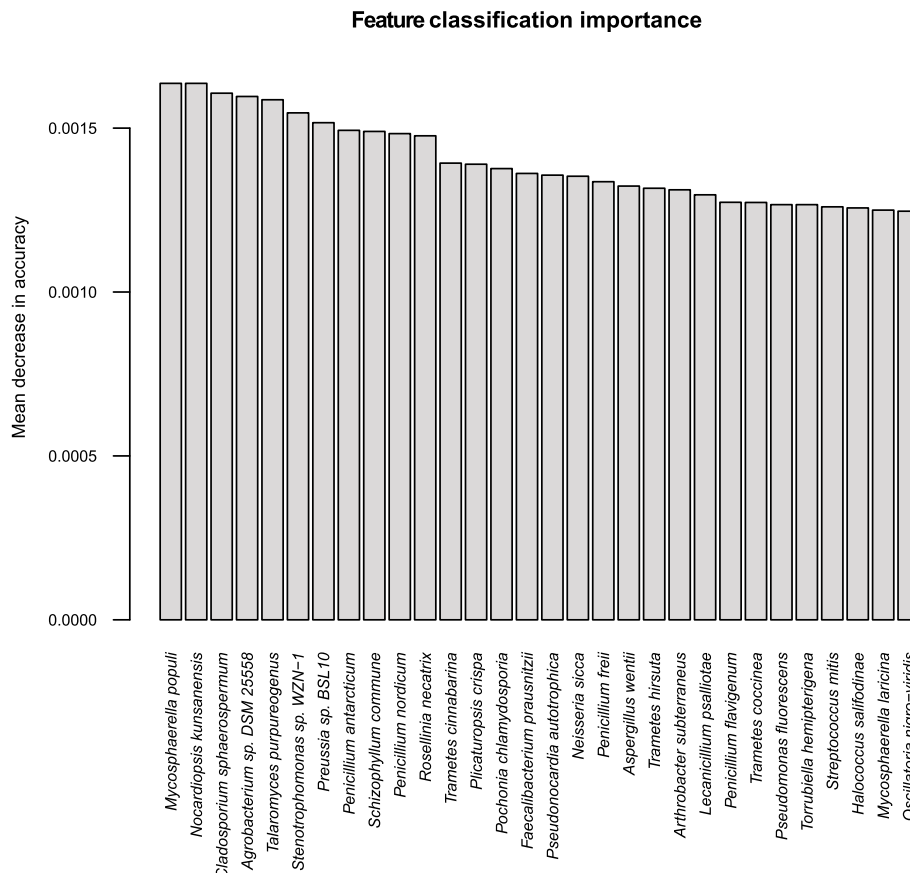
isolation method, including extensively modified commercial DNA isolation kits, appears to be necessary to overcome the unique and inherent challenges associated with SMS-based aerosol microbiome research in complex low biomass air environments [2, 21, 28, 29].

## Conclusions

By demonstrating and benchmarking a new custom, multi-component DNA isolation method (the MetaSUB method) optimized for SMS-based aerosol microbiome research, this study contributes to improved selection, harmonization, and standardization of DNA isolation methods. In the context of SMS-based aerosol microbiome research in low biomass air environments, our findings highlight the importance of ensuring end-to-end sample integrity and using DNA isolation methods with well-defined performance characteristics regarding both DNA yield and community representation. A comprehensive performance benchmarking of the MetaSUB method against two other state-of-the-art DNA isolation methods (Jiang and Zymbiomics) was done with both a mock microbial community and real-world subway air samples. All three DNA isolation

methods performed similarly well on mock community samples, both in terms of DNA yield and community representation. However, the MetaSUB method obtained significantly higher DNA yields than the other two methods from subway air samples, which is an important performance parameter for successful implementation of SMS on low biomass air samples. We also observed significant differences regarding SMS-based community representation across the three methods when applying them to subway air samples. The MetaSUB method reported higher  $\alpha$ -diversity estimates than Zymbiomics, while Jiang appeared to underrepresent certain fungal species. By processing the entire filter extract, in combination with thorough chemical, enzymatic and mechanical biomass lysis, and efficient DNA recovery using magnetic beads, the MetaSUB method may drastically improve the DNA yield from low biomass air samples and reduce the risk of aerosol microbiome profile bias. Taken together, the demonstrated performance characteristics suggest the MetaSUB method could be used to improve the quality of SMS-based aerosol microbiome research in low biomass air environments. Furthermore, the MetaSUB method, when used in combination with the described high-





**Fig. 13** Random forest classification analysis of subway air samples ( $N = 6$ ) where the intermediate pellet ( $N = 6$ ) and supernatant ( $N = 6$ ) fractions were processed separately with the MetaSUB method, showing taxonomic features with the highest classification variable importance for correctly identifying the pellet and supernatant fractions

volume filter-based air sampling, filter processing and DNA isolation scheme (the end-to-end MetaSUB method), could be used to improve the temporal resolution in aerosol microbiome research by reducing the sampling time required to obtain sufficient DNA yields for SMS analysis.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s40793-019-0349-z>.

**Additional file 1: Figure S1.** Rarefaction curves with  $\alpha$ -diversity measures: "Observed", "Shannon", and "Simpson" for subway air samples ( $N = 6$ ) that were split and processed with the MetaSUB ( $N = 3$ ) and Jiang ( $N = 3$ ) or MetaSUB ( $N = 3$ ) and Zymobiomics ( $N = 3$ ) methods. **Figure S2.** Rarefaction curves with  $\alpha$ -diversity measures: "Observed", "Shannon", and "Simpson" for the intermediate pellet ( $N = 6$ ) and supernatant ( $N = 6$ ) fractions from subway air samples ( $N = 6$ ) processed separately with the MetaSUB method. **Figure S3.** Proportion of total DNA and 16S rRNA gene copy yield found in the supernatant fractions, referencing the total yield in the combined pellet and supernatant fractions, from subway air samples ( $N = 24$ ) where the intermediate pellet and supernatant fractions were processed separately with the MetaSUB method. **Figure S4.** The 20 fungal species that were among the top 100 species from the random

forest classification analysis of subway air samples ( $N = 3$ ) that were split and processed with the MetaSUB ( $N = 3$ ) and Jiang ( $N = 3$ ) methods, where Z-score distributions were compared with linear models.

## Acknowledgements

The sequencing service was provided by the Norwegian Sequencing Centre ([www.sequencing.uio.no](http://www.sequencing.uio.no)), a national technology platform hosted by the University of Oslo and supported by the "Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities. We want to thank the Defence Science and Technology Laboratory (Dstl) and MetaSUB international consortium for the long-standing scientific collaboration that has largely inspired the work presented here.

## Authors' contributions

MD conceived, designed and led the study. JG performed the data analysis. LV-M contributed to the experimental work and the data analysis. KO-B performed the experimental work and contributed to the data analysis. All authors contributed to the manuscript writing and approved the final manuscript.

## Funding

The study was funded by the Norwegian Defence Research Establishment FFI.

#### Availability of data and materials

The sequence data has been deposited in the NCBI Sequence Read Archive under Bioproject ID# PRJNA542423 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA542423>).

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 2 June 2019 Accepted: 3 December 2019

Published online: 13 January 2020

#### References

1. Yao M. Bioaerosol: a bridge and opportunity for many scientific research fields. *J Aerosol Sci.* 2018;115:108–12. <https://doi.org/10.1016/j.jaerosci.2017.07.010>.
2. King P, Pham LK, Waltz S, Sphar D, Yamamoto RT, Conrad D, Taplitz R, Torriani F, Forsyth RA. Longitudinal metagenomic analysis of hospital air identifies clinically relevant microbes. *PLoS One.* 2016;11(8):e0160124. <https://doi.org/10.1371/journal.pone.0160124>.
3. Choi JY, Zemke J, Philo SE, Bailey ES, Yondon M, Gray GC. Aerosol sampling in a hospital emergency room setting: a complementary surveillance method for the detection of respiratory viruses. *Front Public Health.* 2018;6:174. <https://doi.org/10.3389/fpubh.2018.00174>.
4. Nguyen TT, Poh MK, Low J, Kalimuddin S, Thoon KC, Ng WC, Anderson BD, Gray GC. Bioaerosol Sampling in Clinical Settings: A Promising, Noninvasive Approach for Detecting Respiratory Viruses. *Open Forum Infect Dis.* 2017;4(1):ofw259. <https://doi.org/10.1093/ofid/ofw259>.
5. Prost K, Kloeze H, Mukhi S, Bozek K, Poljak Z, Mubareka S. Bioaerosol and surface sampling for the surveillance of influenza a virus in swine. *Transbound Emerg Dis.* 2019. <https://doi.org/10.1111/tbed.13139>.
6. Osman S, La Duc MT, Dekas A, Newcombe D, Venkateswaran K. Microbial burden and diversity of commercial airline cabin air during short and long durations of travel. *Int Soc Microb Ecol (ISME J).* 2008;2:482–97. <https://doi.org/10.1038/ismej.2008.11>.
7. Weiss H, Hertzberg VS, Dupont C, Espinoza JL, Levy S, Nelson K, Norris S, Team TFR. The airplane cabin microbiome. *Microb Ecol.* 2019;77(1):87–95. <https://doi.org/10.1007/s00248-018-1191-3>.
8. Zanni S, Lalli F, Foschi E, Bonoli A, Mantecchini L. Indoor air quality real-time monitoring in airport terminal areas: an opportunity for sustainable Management of Micro-Climatic Parameters. *Sensors.* 2018;18(11):3798. <https://doi.org/10.3390/s18113798>.
9. Wagar E. Bioterrorism and the role of the clinical microbiology laboratory. *Clin Microbiol Rev.* 2016;29(1):175–89. <https://doi.org/10.1128/CMR.00033-15>.
10. Amann RI, Ludwig W, Schleifer K-H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Mol Biol Rev.* 1995;59(1):143–69.
11. Radosevich JL, Wilson WJ, Shinn JH, DeSantis TZ, Andersen GL. Development of a high-volume aerosol collection system for the identification of air-borne micro-organisms. *Lett Appl Microbiol.* 2002;34(3):162–7. <https://doi.org/10.1046/j.1472-765x.2002.01048.x>.
12. Toivola M, Alm S, Reponen T, Kolari S, Nevalainen A. Personal exposures and microenvironmental concentrations of particles and bioaerosols. *J Environ Monit.* 2002;4(1):166–74.
13. Eduard W, Heederik D. Methods for quantitative assessment of airborne levels of noninfectious microorganisms in highly contaminated work environments. *Am Ind Hyg Assoc J.* 1998;59(2):113–27. <https://doi.org/10.1080/15428119891010370>.
14. Leung MH, Wilkins D, Li EK, Kong FK, Lee PK. Indoor-air microbiome in an urban subway network: diversity and dynamics. *Appl Environ Microbiol.* 2014;80(21):6760–70. <https://doi.org/10.1128/AEM.02244-14>.
15. Triado-Margarit X, Veillette M, Duchaine C, Talbot M, Amato F, Minguillon MC, Martins V, de Miguel E, Casamayor EO, Moreno T. Bioaerosols in the Barcelona subway system. *Indoor Air.* 2017;27(3):564–75. <https://doi.org/10.1111/ina.12343>.
16. Cáliz J, Triadó-Margarit X, Camarero L, Casamayor EO. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proc Natl Acad Sci.* 2018;115(48):12229–34. <https://doi.org/10.1073/pnas.1812826115>.
17. Hanson B, Zhou Y, Bautista EJ, Urch B, Speck M, Silverman F, Muilenberg M, Phipatanakul W, Weinstock G, Sodergren E, et al. Characterization of the bacterial and fungal microbiome in indoor dust and outdoor air samples: a pilot study. *Environ Sci.* 2016;18(6):713–24. <https://doi.org/10.1039/c5em00639b>.
18. The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature.* 2012;486:207–14. <https://doi.org/10.1038/nature11234>.
19. Afshinnkoo E, Meydan C, Chowdhury S, Jaroudi D, Boyer C, Bernstein N, Maritz JM, Reeves D, Gandara J, Chhangawala S, et al. Geospatial resolution of human and bacterial diversity with City-scale Metagenomics. *Cell Systems.* 2015;1(1):72–87. <https://doi.org/10.1016/j.cels.2015.01.001>.
20. Biller SJ, Berube PM, Dooley K, Williams M, Satinsky BM, Hackl T, Hogle SL, Coe A, Bergauer K, Bouman HA, et al. Marine microbial metagenomes sampled across space and time. *Sci Data.* 2018;5:180176. <https://doi.org/10.1038/sdata.2018.176>.
21. Yooshef S, Andrews-Pfannkoch C, Tenney A, McQuaid J, Williamson S, Thiagarajan M, Bami D, Zeigler-Allen L, Hoffman J, Goll JB, et al. A metagenomic framework for the study of airborne microbial communities. *PLoS One.* 2013;8(12):e81862. <https://doi.org/10.1371/journal.pone.0081862>.
22. Cao C, Jiang W, Wang B, Fang J, Tian G, Jiang J, Zhu TF. Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event. *Environ Sci Technol.* 2014;48(3):1499–507. <https://doi.org/10.1021/es4048472>.
23. Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol.* 2019;27(2):105–17. <https://doi.org/10.1016/j.tim.2018.11.003>.
24. Dybwad M, Skogan G, Blatny JM. Comparative testing and evaluation of nine different air samplers: end-to-end sampling efficiencies as specific performance measurements for bioaerosol applications. *Aerosol Sci Technol.* 2014;48(3):282–95. <https://doi.org/10.1080/02786826.2013.871501>.
25. Behzad H, Gojobori T, Mineta K. Challenges and opportunities of airborne metagenomics. *Genome Biol Evolution.* 2015;7(5):1216–26. <https://doi.org/10.1093/gbe/evw064>.
26. Tringe SG, Hugenholtz P. A renaissance for the pioneering 16S rRNA gene. *Curr Opin Microbiol.* 2008;11(5):442–6. <https://doi.org/10.1016/j.mib.2008.09.011>.
27. Luhung I, Wu Y, Ng CK, Miller D, Cao B, Chang VW. Protocol improvements for Low concentration DNA-based bioaerosol sampling and analysis. *PLoS One.* 2015;10(11):e0141158. <https://doi.org/10.1371/journal.pone.0141158>.
28. Jiang W, Liang P, Wang B, Fang J, Lang J, Tian G, Jiang J, Zhu TF. Optimized DNA extraction and metagenomic sequencing of airborne microbial communities. *Nat Protoc.* 2015;10:768. <https://doi.org/10.1038/nprot.2015.046>.
29. Dommergue A, Amato P, Tignat-Perrier R, Magand O, Thollot A, Joly M, Bouvier L, Sellegri K, Vogel T, Sonke JE, et al. Methods to investigate the global atmospheric microbiome. *Front Microbiol.* 2019;10:00243. <https://doi.org/10.3389/fmicb.2019.00243>.
30. Tighe S, Afshinnkoo E, Rock TM, McGrath K, Alexander N, McIntyre A, Ahsanuddin S, Bezdán D, Green SJ, Joye S, et al. Genomic methods and microbiological technologies for Profiling Novel and Extreme Environments for the extreme microbiome project (XMP). *J Biomol Tech.* 2017;28(1):31–9. <https://doi.org/10.1071/jbt.17-2801-004>.
31. Yuan S, Cohen DB, Ravel J, Abdo Z, Forney LJ. Evaluation of methods for the extraction and purification of DNA from the human microbiome. *PLoS One.* 2012;7(3):e33865. <https://doi.org/10.1371/journal.pone.0033865>.
32. Abusleme L, Hong BY, Dupuy AK, Strausbaugh LD, Diaz PI. Influence of DNA extraction on oral microbial profiles obtained via 16S rRNA gene sequencing. *J Oral Microbiol.* 2014;6:23990. <https://doi.org/10.3402/jom.v6.23990>.
33. Mbareche H, Brisebois E, Veillette M, Duchaine C. Bioaerosol sampling and detection methods based on molecular approaches: no pain no gain. *Sci Total Environ.* 2017;599–600:2095–104. <https://doi.org/10.1016/j.scitotenv.2017.05.076>.
34. The Human Microbiome Project Consortium. A framework for human microbiome research. *Nature.* 2012;486:215–21. <https://doi.org/10.1038/nature11209>.
35. Gilbert JA, Jansson JK, Knight R. Earth Microbiome Project and Global Systems Biology. *mSystems.* 2018;3(3):e00217. <https://doi.org/10.1128/mSystems.00217-17>.
36. Kopf A, Bicap M, Kottmann R, Schnetzer J, Kostadinov I, Lehmann K, Fernandez-Guerra A, Jeanthon C, Rahav E, Ullrich M, et al. The ocean sampling day consortium. *GigaScience.* 2015;4(1):s13742. <https://doi.org/10.1186/s13742-015-0066-5>.

37. Lear G, Dickie I, Banks J, Boyer S, Buckley HL, Buckley TR, Cruickshank R, Dopheide A, Handley KM, Hermans S, et al. Methods for the extraction, storage, amplification and sequencing of DNA from environmental samples. *N Z J Ecol.* 2018;42(1):10–50. <https://doi.org/10.20417/nzjecol.42.9>.
38. Al-Hebshi NN, Baraniya D, Chen T, Hill J, Puri S, Tellez M, Hasan NA, Colwell RR, Ismail A. Metagenome sequencing-based strain-level and functional characterization of supragingival microbiome associated with dental caries in children. *J Oral Microbiol.* 2018;11(1):1557986. <https://doi.org/10.1080/20002297.2018.1557986>.
39. Honeyman AS, Day ML, Spear JR. Regional fresh snowfall microbiology and chemistry are driven by geography in storm-tracked events, Colorado, USA. *PeerJ.* 2018;6:e5961. <https://doi.org/10.7717/peerj.5961>.
40. Zaikova E, Goerlitz DS, Tighe SW, Wagner NY, Bai Y, Hall BL, Bevilacqua JG, Weng MM, Samuels-Fair MD, Johnson SS. Antarctic relic microbial mat Community revealed by Metagenomics and Metatranscriptomics. *Front Ecol Evol.* 2019;7(1):10.3389. <https://doi.org/10.3389/fevo.2019.00001>.
41. Trivedi CB, Lau GE, Grasby SE, Templeton AS, Spear JR. Low-temperature Sulfidic-ice microbial communities, Borup fiord pass, Canadian High Arctic. *Front Microbiol.* 2018;9:01622. <https://doi.org/10.3389/fmicb.2018.01622>.
42. Liu CM, Aziz M, Kachur S, Hsueh P-R, Huang Y-T, Keim P, Price LB. BactQuant: an enhanced broad-coverage bacterial quantitative real-time PCR assay. *BMC Microbiol.* 2012;12:56. <https://doi.org/10.1186/1471-2180-12-56>.
43. Minot SS, Krumm N, Greenfield NB. One Codex: A Sensitive and Accurate Data Platform for Genomic Microbial Identification. *bioRxiv.* 2015;027607. <https://doi.org/10.1101/027607>.
44. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One.* 2013;8(4):e61217.
45. Huson DH, Auch AF, Qi J, Schuster SC. MEGAN analysis of metagenomic data. *Genome Res.* 2007;17(3):377–86.
46. Liaw A, Wiener M. Classification and regression by randomForest. *R News.* 2002;2(3):18–22.
47. Miller DN, Bryant JE, Madsen EL, Ghiorse WC. Evaluation and optimization of DNA extraction and purification procedures for soil and sediment samples. *Appl Environ Microbiol.* 1999;65(11):4715–24.
48. Griffiths LJ, Anyim M, Doffman SR, Wilks M, Millar MR, Agrawal SG. Comparison of DNA extraction methods for *Aspergillus fumigatus* using real-time PCR. *J Med Microbiol.* 2006;55(9):1187–91. <https://doi.org/10.1099/jmm.0.46510-0>.
49. Mbareche H, Veillette M, Teertstra W, Kegel W, Bilodeau GJ, Wösten HAB, Duchaine C. Fungal cells recovery from air samples: a tale of loss and gain. *Appl Environ Microbiol.* 2019;02941. <https://doi.org/10.1128/aem.02941-18>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

