

# Optimized DNA extraction and metagenomic sequencing of airborne microbial communities

Wenjun Jiang<sup>1</sup>, Peng Liang<sup>1</sup>, Buyang Wang<sup>2</sup>, Jianhuo Fang<sup>3</sup>, Jidong Lang<sup>3</sup>, Geng Tian<sup>3</sup>, Jingkun Jiang<sup>2</sup> & Ting F Zhu<sup>1</sup>

<sup>1</sup>School of Life Sciences, Center for Synthetic and Systems Biology, Ministry of Education Key Laboratory of Bioinformatics, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, Tsinghua University, Beijing, China. <sup>2</sup>State Key Joint Laboratory of Environment Simulation and Pollution Control, School of Environment, State Environmental Protection Key Laboratory of Sources and Control of Air Pollution Complex, Tsinghua University, Beijing, China. <sup>3</sup>Center of Biomedical Analysis, School of Life Sciences, Tsinghua University, Beijing, China. Correspondence should be addressed to T.F.Z. (tzhu@biomed.tsinghua.edu.cn).

Published online 23 April 2015; doi:10.1038/nprot.2015.046

Metagenomic sequencing has been widely used for the study of microbial communities from various environments such as soil, ocean, sediment and fresh water. Nonetheless, metagenomic sequencing of microbial communities in the air remains technically challenging, partly owing to the limited mass of collectable atmospheric particulate matter and the low biological content it contains. Here we present an optimized protocol for extracting up to tens of nanograms of airborne microbial genomic DNA from collected particulate matter. With an improved sequencing library preparation protocol, this quantity is sufficient for downstream applications, such as metagenomic sequencing for sampling various genes from the airborne microbial community. The described protocol takes ~12 h of bench time over 2–3 d, and it can be performed with standard molecular biology equipment in the laboratory. A modified version of this protocol may also be used for genomic DNA extraction from other environmental samples of limited mass or low biological content.

## INTRODUCTION

Metagenomic sequencing is a cultivation-independent method that extracts and analyzes genetic materials directly from environmental samples<sup>1–7</sup>. It has greatly expanded our knowledge of life's diversity and accelerated efforts to fill the uncharted chapters of life in various environments such as ocean, soil, sediment, and fresh water, as well as in extreme environments<sup>2,3,5–7</sup>. Nevertheless, despite being widely used to study nearly all the microbial communities from the surface of the Earth, direct metagenomic sequencing of the microbial communities in the air has remained a challenge. This partly resulted from the limited availability of collectable airborne particulate matter (typically in the range of tens to hundreds of milligrams), the low biological content of the samples and incompatibility of traditional DNA extraction with airborne particulate matter collection methods (e.g., the use of quartz filters)<sup>8</sup>. In the past, the majority of the studies on airborne microorganisms were based on 16S rRNA gene-specific amplification and sequencing, which, although contributing tremendously to our understanding of the identity of the airborne microorganisms in various environments<sup>9–13</sup>, has also limited our ability to sample other key microbial genes of the community.

We have developed and applied an optimized protocol for extracting limited quantities of airborne microbial genomic DNA for direct metagenomic sequencing<sup>8</sup>. In this protocol, we have integrated several key modifications into the steps of airborne particulate matter collection, sample pretreatment (Fig. 1), DNA extraction and sequencing library preparation. We show that such modifications and optimization have greatly improved the DNA extraction efficiency of airborne particulate matter samples. By using this protocol, one can extract a few nanograms to tens of nanograms of genomic DNA from tens of milligrams of collected particulate matter, which is sufficient for downstream library preparation and high-throughput sequencing. As an example, we have recently applied this protocol to the study of the airborne microbial communities in the city of Beijing, where we extracted

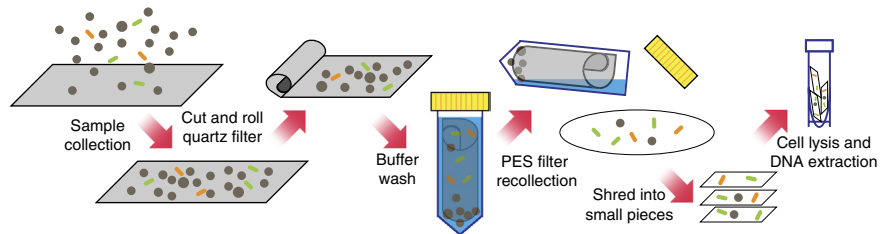
the DNA from airborne microbes for metagenomic sequencing, characterized the diverse microbial communities and identified more than 1,300 species of bacteria, fungi and viruses<sup>8</sup>. The availability of a vast number of metagenomic sequences has enabled us to identify inhalable pathogens and allergens at the species level, as well as to study the unique gene sets of the airborne microbial community.

Various methods, including a number of commercial kits, have been developed and applied for the isolation of DNA from soil, water and other environmental samples<sup>7,14,15</sup>, including single cell-based sequencing methods<sup>16</sup>. Our protocol has integrated these existing DNA extraction methods used for treating soil samples with airborne particulate matter collection, sample pretreatment and sequencing library preparation. The apparent advantage of our protocol is that the DNA extraction requires a very small mass of input sample (typically in the range of tens to hundreds of milligrams). This allows one to analyze and obtain many more sequences than the previous studies that typically used 16S rRNA gene or whole-genome amplification to compensate for the low yield of genomic DNA. In conjunction with an optimized DNA library preparation protocol, the extracted genomic DNA can be used directly for high-throughput sequencing, revealing much more information on the microbial community and the genes it comprises. An additional advantage of this optimized protocol is that with high-throughput sequencing one obtains much more sequence information than from the 16S rRNA genes for identifying bacteria, which typically are limited to the genus or family levels; instead, one can pinpoint the specific species of bacteria, fungi and viruses that might be harmful to human health<sup>8</sup>.

This protocol may be adapted to extract environmental genomic DNA from low-biomass aquatic samples that are collectable on filter membranes (e.g., arctic glacier samples and others alike). In addition to being a research tool, this protocol might be adapted for applications in practical areas such as the detection of airborne



**Figure 1** | Workflow for particulate matter sample pretreatment. After sample collection (Steps 1–5), the sample pretreatment workflow includes cutting and rolling the quartz filter containing the collected particulate matter (Step 8), buffer wash (Step 9), PES filter recollection (Steps 10–11) and shredding the PES filter into small pieces (Step 12) before the subsequent cell lysis and DNA extraction steps (Steps 13–25).



pathogens, evaluation of indoor air quality, and the detection of airborne pathogens responsible for hospital-acquired infections<sup>17</sup>. In this protocol, we used a centrifugation-based method for the separation of the collected particles from quartz filters for achieving better extraction efficiency, a step that might introduce bias in microbial recovery. In addition, so far we have been unable to extract sufficient RNA from the atmospheric particulate matter for high-throughput sequencing, thereby missing the sequences of the RNA viruses. Adaptations and modifications of this protocol need to be developed to extract RNA from airborne microbial communities for the study of airborne RNA viruses such as severe acute respiratory syndrome (SARS) and influenza viruses<sup>18,19</sup>.

Here we describe our protocol for extracting airborne microbial genomic DNA for metagenomic sequencing from collected particulate matter samples. The protocol outlines the steps of particulate matter collection, sample pretreatment for the separation of the collected particles from quartz filter (Fig. 1), DNA extraction from the particles, quantification and quality assessment of extracted genomic DNA, and preparation of the sequencing library.

### Experimental design

**Particulate matter collection (Steps 1–5).** To choose the location for collecting particulate matter samples, one should consider the target airborne microbial community to be studied. For collecting atmospheric particulate matter samples, locations without major interference sources nearby are recommended. For example, in our previous study<sup>8</sup>, we were interested in the atmospheric airborne microbial community, and hence the PM<sub>2.5</sub> and PM<sub>10</sub> (particulate matter with nominal mean aerodynamic diameters of  $\leq 2.5$  and  $\leq 10$   $\mu\text{m}$ , respectively) samples were collected from the rooftop of a two-story building on Tsinghua University campus, a site that has been used to monitor atmospheric air pollution in Beijing since 1999 (refs. 20–22). Particulate matter can be collected on 20.32 cm  $\times$  25.4 cm Tissuquartz filters that have 99.9% typical aerosol retention efficiency. Sample collection can be conducted by high-volume air samplers with an average air flow rate of 1.13 m<sup>3</sup>/min (Fig. 2a), the inlets of which have been shown to be effective with wind speeds of up to 24 km/h (ref. 23). This results in  $\sim 1,559$  m<sup>3</sup> of air flow-through per sampling day (typically 23 h, allowing 1 h for changing filters), which is important to ensure a sufficient mass of input sample (typically in the range of tens to hundreds of milligrams) to be collected. In our previous study<sup>8</sup>, we used a total of three high-volume air samplers: two of which were equipped with PM<sub>2.5</sub> fractionating inlets and one with a PM<sub>10</sub> fractionating inlet (Fig. 2b,c). Weigh each filter before and after sampling (Fig. 2d), and all samples can be used immediately or stored at  $-80$  °C until the next step.

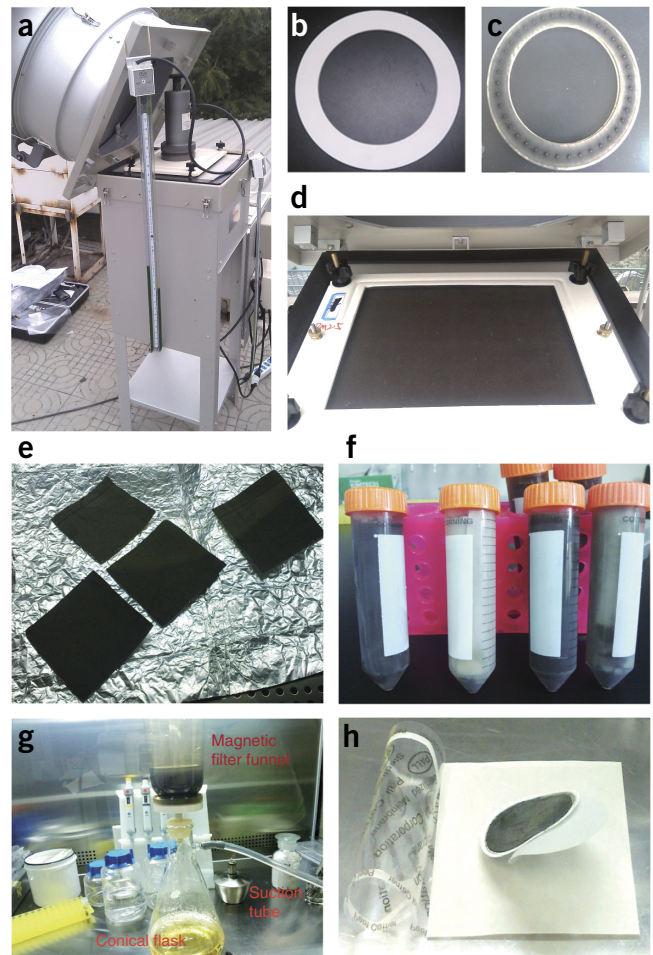
**Sample pretreatment (Steps 6–12).** Owing to the low mass of collectable particulate matter samples and the DNA-binding properties of silica fibers in the Tissuquartz filter, we designed a series of sample pretreatment steps to optimize the yield of genomic DNA (Fig. 1). Considering the different mass of input materials from the collected PM<sub>2.5</sub> and PM<sub>10</sub> samples, we typically use one-fourth of a PM<sub>10</sub> filter ( $\sim 103.04$  cm<sup>2</sup>) and 1.5 or 2 pieces of PM<sub>2.5</sub> filters (collected by two samplers on the same day) for sample pretreatment and subsequent DNA extraction. The filters are cut into four equal-sized pieces (8.96 cm  $\times$  11.5 cm each; Fig. 2e), which are rolled and inserted into 50-ml Falcon tubes filled with sterilized 1 $\times$  PBS buffer (Fig. 2f). The particulate matter samples are then pelleted at 4 °C by low-speed centrifugation at 200g for 3 h. After brief, gentle vortexing, the resuspension is filtered through a 0.2- $\mu\text{m}$  Supor 200 PES Membrane Disc Filter (Fig. 2g,h). The above wash and filter recollection (pretreatment) steps are necessary to help separate most of the biological contents from the quartz filters before DNA extraction for improved extraction efficiency. In addition, the pretreatment of particulate matter sample reduces the sample volume needed for DNA extraction, and thus it allows the use of AMPure XP beads for genomic DNA extraction at a reasonable cost. Quantification of the recovered genomic DNA and quantitative PCR (qPCR) results suggested that the extraction efficiency of microbial DNA from particulate matter was optimized by applying the pretreatment step and by using AMPure XP beads compared with other methods (Table 1 and Fig. 3a). After the wash and filter recollection, we cut the PES Membrane Disc Filter into small pieces for DNA extraction using commercial kits. All the above steps should be carried out in a decontaminated biosafety cabinet.

**DNA extraction (Steps 13–24).** A variety of DNA extraction kits for handling environmental samples have been developed. We recommend performing DNA extraction of pretreated particulate matter samples using the MO-BIO PowerSoil DNA Isolation Kit, with a few key modifications to the manufacturer’s instructions. First, the samples are heated to 65 °C (instead of at room temperature (25 °C)) in PowerBead tubes (included in the PowerSoil kit) for 15 min; this is followed by vortexing for 15 min. The remaining steps of the extraction are performed according to the standard PowerSoil DNA isolation protocol, except that the column purification is replaced by AMPure XP bead purification for improved DNA yield, as shown by comparisons between different extraction methods (Table 1 and Fig. 3a).

**Quantification and quality assessment of extracted genomic DNA (Step 25).** The extracted genomic DNA should be checked for concentration and quality before library preparation. Because the concentrations of the extracted genomic DNA are typically

## PROTOCOL

**Figure 2** | Experimental setup for particulate matter collection and sample pretreatment. (a–c) Setup of a VFC high-volume air sampler (Step 4) (a), with a clean PM<sub>2.5</sub> inlet gasket (b) and a gasket after several days of particulate matter collection (c). (d) A Tissuquartz filter after 23 h of particulate matter collection. (e,f) A Tissuquartz filter is cut into four equal-sized pieces (Step 8) (e), before being rolled and inserted into 50-ml Falcon tubes filled with sterilized 1× PBS buffer, and then pelleted at 4 °C by low-speed centrifugation at 200g for 3 h (Step 9) (f). (g) The setup for re-filtration of the resuspension using a magnetic filter funnel (Step 10). (h) The re-collected sample on a 0.2-μm Supor 200 PES membrane disc filter (Step 11).



below the reliable detection limit of common DNA quantification instruments such as the NanoDrop Spectrophotometer (the detection limit of which is typically in the range of a few ng/μl), we recommend using a fluorescent double-stranded DNA (dsDNA)-binding dye assay on the Qubit Fluorometer, which has a detection limit of dsDNA at 10–100 pg/μl. The low concentrations of DNA also do not permit visualization on a regular agarose gel, and thus gel electrophoresis is not required. If needed, we recommend using the Lonza FlashGel system, which can be used to visualize DNA bands at quantities as low as 0.1 ng. qPCR of the genomic DNA samples using 16S rRNA gene universal primers can be used to assess the quality of the extracted DNA (a step that is recommended but not required). Amplification curves with lower threshold cycle (C<sub>t</sub>) values suggest the presence of more bacterial DNA in the extracted genomic DNA and thus better sample quality. All the extracted DNA samples can be stored at –80 °C until further use.

**Sequencing library preparation (Steps 26–34).** Various high-throughput sequencing platforms can be used for metagenomic sequencing, and the library preparation methods vary accordingly<sup>24,25</sup>. We recommend using the Illumina sequencing platforms for the high data throughput and low per-base cost. The sequencing library construction and template preparation can be performed using the NEBNext ultra DNA library prep kit (by following the steps for input DNA <50 ng) or by using the modified protocol provided in this paper. In our previous study<sup>8</sup>, we constructed barcoded, paired-end libraries with an insert size of ~500 bp for each sample. To ensure sample consistency, the entire volume of extracted DNA sample, or an aliquot of 5 ng of DNA from each sample (for DNA samples >5 ng), is used for library preparation. The lowest quantity of input genomic DNA

from which we have successfully prepared Illumina sequencing library was 3.84 ng (ref. 8). The quality of the libraries should be validated by the Agilent 2100 platform before Illumina sequencing; the existence of a single, wide peak at 500–800 bp in electropherogram suggests good library quality (Fig. 3b–d).

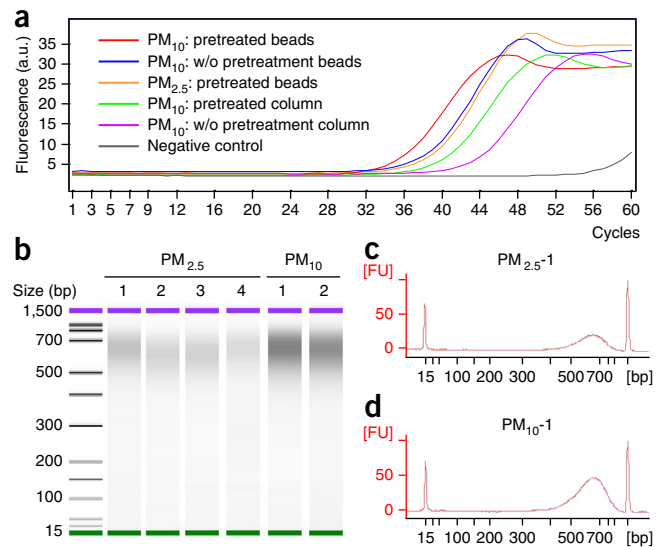
**DNA sequencing and data analysis (Steps 35 and 36).** We recommend using the Illumina MiSeq platform for validating the quality of the above-prepared libraries by generating small data sets of 500 Mb–1 Gb for each sample (optional) and the HiSeq 2000 or 2500 sequencing platform for high-throughput sequencing.

**TABLE 1** | DNA extracted by different pretreatment, extraction and purification methods.

Sample date	Sample type	Filter size	Net weight (mg)	Volume (μl)	Conc. (ng/μl)	Quant. (ng)	Extraction methods used
2012.12.01	PM <sub>2.5</sub>	1.5	97	50	0.195	9.75	PowerSoil, AMPure XP beads
2012.12.03	PM <sub>10</sub>	1/4	462	50	0.632	31.6	PowerSoil, AMPure XP beads
2012.12.03	PM <sub>10</sub>	1/4	462	50	0.382	19.1	PowerMax, AMPure XP beads
2012.12.03	PM <sub>10</sub>	1/4	462	50	0.198	9.9	PowerSoil, column
2012.12.03	PM <sub>10</sub>	1/4	462	50	0.073	3.65	PowerMax, column

The PowerSoil kit was used to extract DNA that had been pretreated according to this protocol; the PowerMax kit was used to extract DNA from samples without pretreatment. AMPure XP beads or columns were the two different methods of DNA purification tested.

**Figure 3** | Quality assessment of extracted genomic DNA and validation of the quality of prepared sequencing libraries. **(a)** qPCR results from genomic DNA samples extracted as described in **Table 1** using 16S rRNA gene universal primers (Step 25). PM<sub>10</sub>-pretreated beads: DNA extracted from pretreated PM<sub>10</sub> sample (using the MO-BIO PowerSoil kit) and purified using AMPure XP beads; PM<sub>10</sub> without (w/o) pretreatment beads: DNA extracted from PM<sub>10</sub> sample by shredding the quartz filter without pretreatment (using the MO-BIO PowerMax kit) and purified using AMPure XP beads; PM<sub>10</sub>-pretreated column: DNA extracted from pretreated PM<sub>10</sub> sample (using the MO-BIO PowerSoil kit) and purified using traditional column purification; PM<sub>10</sub> w/o pretreatment column: DNA extracted from PM<sub>10</sub> sample by shredding the quartz filter without pretreatment (using the MO-BIO PowerMax kit) and purified using traditional column purification; PM<sub>2.5</sub> pretreated beads: DNA extracted from pretreated PM<sub>2.5</sub> sample (using the MO-BIO PowerSoil kit) and purified using AMPure XP beads. qPCR of the above five samples resulted in successful 16S rRNA gene amplification, whereas the negative control did not. Amplification curves with lower C<sub>t</sub> values suggest the presence of more bacterial DNA in the extracted genomic DNA and thus better sample quality. The qPCR data (color-coded and ranked top-down from the lowest C<sub>t</sub> to the highest), together with the DNA concentration data presented in **Table 1**, suggest that the method of using pretreatment with bead purification provides higher DNA yield than all the other methods tested above. The addition of the pretreatment step improves the yield of DNA extraction, and using beads for DNA purification is more effective than using column. **(b–d)** Agilent 2100 Bioanalyzer analysis of six examples of prepared good-quality sequencing libraries (Step 34), including those of four PM<sub>2.5</sub> samples and two PM<sub>10</sub> samples (all DNA samples were extracted using the pretreated beads method). **(b)** Electrophoresis of the six libraries: a single, smeary band at 500–800 bp suggests good library quality (the green and purple bands correspond to Bioanalyzer-called lower and upper markers, respectively). **(c,d)** Electropherograms of two of the libraries (PM<sub>2.5</sub>-1 and PM<sub>10</sub>-1, as shown in **b**) with 15-bp and 1,500-bp markers; a single, wide peak at 500–800 bp suggests good library quality. a.u., arbitrary units; FU, fluorescence units.



For HiSeq sequencing, each sample can be barcoded (added during library preparation), and equal quantities of barcoded libraries can be multiplexed during sequencing. The use of other alternative high-throughput sequencing platforms is possible, and the manufacturer's instructions on the library preparation and sequencing should be followed. Attention should be paid to the throughput and requirement of input genomic DNA. To appreciate the diversity and to sample a sufficient amount of microbial genes, we recommend generating at least 5–7 Gb of data per sample for the study of complex airborne microbial communities. Methods for data analysis vary greatly, depending on the specific needs of the study<sup>26–28</sup>, but they typically involve removal of the adapter sequences, low-quality reads and duplicate reads, followed by alignment or assembly of the short reads. For the purpose of identifying microbial species<sup>8</sup>, HiSeq reads are aligned to a cohort of nonredundant National Center for Biotechnology Information (NCBI) complete genomes using the Short Oligonucleotide Analysis Package (SOAP) alignment tool<sup>29</sup>, which is typically faster to run than the Basic Local Alignment Search Tool (BLAST) or the BLAST-like Alignment Tool (BLAT). Genome coverage is calculated using the SOAP.coverage package.

**Contamination prevention.** Owing to the low mass and low biological content, particulate matter samples are sensitive to laboratory and other environmental contaminations, and thus extensive efforts to avoid contamination should be taken in the sample collection, pretreatment and DNA extraction steps. The Tissuquartz filters used for sample collection should be sterilized by baking in a muffle furnace at 500 °C for 5 h before sampling. The sterilized filters can be individually packaged in sterilized aluminum foil and stored in self-sealing plastic bags until use. The filter holder and all the tools used for changing new filters should also be cleaned with 75% (vol/vol) ethanol, or they should be autoclaved (if possible) every day to avoid cross-contamination. In the sample pretreatment and DNA extraction steps, the reusable scissors, forceps and filter funnels should be sterilized before use. The use of sterile surgical gloves and face masks is also recommended. To assess the degree of potential contamination, negative-control experiments should be conducted by placing a sterilized filter in the sampler without collecting a sample, followed by an identical sample pretreatment and DNA extraction process to that used when samples are collected.

**MATERIALS**

**REAGENTS**

- Absolute ethanol (e.g., reagent alcohol; Sigma-Aldrich, cat. no. 270741) **! CAUTION** Absolute ethanol is highly flammable. Store it in a fireproof cabinet, and use it with caution and away from open flame.
- Double-distilled water
- Ultrapure PBS (e.g., PBS, 10×, pH 7.4; Life Technologies, cat. no. AM9625)
- UltraPure DNase/RNase-free distilled water (Life Technologies, cat. no. 10977) **! CAUTION** The DNA isolation kit contains organic solvents

and harmful reagents. Wear gloves when you are using this kit; in case of contact, wash thoroughly with water. **▲ CRITICAL** Although other sources of distilled water may be used, in many steps of this protocol it is important to use DNase-free water owing to the small quantity of extracted DNA.

- PowerSoil DNA Isolation Kit (MO-BIO, cat. no. 12888) **▲ CRITICAL** We recommend using the MO-BIO PowerSoil DNA Isolation Kit for its high DNA extraction efficiency when treating particulate matter samples. The use of other alternative DNA isolation kits has not been tested with this protocol.



## PROTOCOL

- Agencourt AMPure XP (Beckman Coulter, cat. no. A63881)
- Qubit dsDNA high-sensitivity (HS) assay kit (Life Technologies, cat. no. Q32851) **! CAUTION** The kit contains dyes that bind nucleic acid and should be treated as a potential mutagen. Some of the reagents contain DMSO. Use the kit with appropriate care.
- LightCycler 480 SYBR Green I master mix (Roche, cat. no. 04707516001) **! CAUTION** SYBR Green binds nucleic acid, and it should be treated as a potential mutagen. Use it with appropriate care.
- Agilent DNA 1000 kit (Agilent Technologies, cat. no. 5067-1504) **! CAUTION** The kit contains dyes that bind nucleic acid. Some of the reagents contain DMSO. Use the kit with appropriate care.
- NEBNext End Repair Module (NEB, cat. no. B6052S)
- NEBNext da-Tailing Module (NEB, cat. no. E6053S)
- Quick Ligation Kit (NEB, cat. no. M2200S)
- Q5 high-fidelity 2× master mix (NEB, cat. no. M0492S)
- TruSeq PE cluster kit v3-cBot-HS (Illumina, cat. no. PE-401-3001)
- TruSeq SBS kit v3-HS (Illumina, cat. no. FC-401-3001)
- Annealing buffer for DNA oligos (5×; Beyotime, cat. no. D0251)
- 515F primer, 10 μM (5'-GTGCCAGCMGCCGCGTAA-3', Invitrogen)
- 806R primer, 10 μM (5'-GGACTACHVGGGTWTCTAAT-3', Invitrogen)
- Illumina adapter forward primer, 50 μM (5'-GATCGGAAGAGCACACGCTCTGAACCTCCAGTCAC-3', Invitrogen)
- Illumina adapter reverse primer, 50 μM (5'-TACACTCTTCCCTACACGACGCTCTTCCGATCT-3', Invitrogen)
- Illumina universal primer, 10 μM (5'-AATGATACGGCAGCACCGGAGTCTACACTCTTCCCTACACGACGCTCTTCCGATCT-3', Invitrogen)
- Illumina index primer, 10 μM (e.g., 5'-CAAGCAGAAGACGGCATACGAGATGATGACCGTGGACTGGAGTTCAGACGTGTGCTCTCCGATCT-3'; the underlined 6-bp sequence is the index sequence of Illumina index 11; Invitrogen)

### EQUIPMENT

#### Particulate matter collection

- VFC high-volume air samplers with PM<sub>2.5</sub> or PM<sub>10</sub> fractionating inlets (Thermo Scientific, cat. no. RFP5-1287-063)
- Biosafety cabinet (e.g., 1300 Series Class II, Type A2 biological safety cabinet; Fisher Scientific, cat. no. 13-261-339)
- Muffle furnace (e.g., Thermolyne benchtop muffle furnaces, Thermo Scientific, cat. no. 10-554-3)
- Autoclave (e.g., Kinematica HMC vertical autoclave; Fisher Scientific, cat. no. 08-451-302)
- Ultra-low-temperature freezer (e.g., Isotemp basic -86 °C ultra-low-temperature freezer; Fisher Scientific, cat. no. IE1386V)
- Precision balance (e.g., Mettler Toledo, cat. no. PL303-IC)
- Tissuquartz filters, 2500 QAT-UP, 8 × 10 inches (PALL Life Sciences, cat. no. 7204)
- Aluminum foil (e.g., Aluminum Foil; Fisher Scientific, cat. no. 01-213-103)
- Self-sealing plastic bags (e.g., zipper seal sample bags, 33 × 23 cm; Fisher Scientific, cat. no. 01-816E)
- Long forceps (e.g., general-purpose extra-long forceps, 20 cm; Fisher Scientific, cat. no. 10-316A)
- Surgical gloves (e.g., Cardinal Health Protexis latex classic surgical gloves; Fisher Scientific, cat. no. 19-163-104)
- Face mask (e.g., High Five face masks; Fisher Scientific, cat. no. 19-167-899)

#### Sample pretreatment

- Irradiated Falcon tubes (e.g., Falcon 50-ml conical centrifuge tubes; Fisher Scientific, cat. no. 14-432-22)
- Sterilized culture dishes (e.g., Falcon standard tissue culture dishes; Fisher Scientific, cat. no. 08-772-E)
- 0.2-μm Supor 200 PES membrane disc filter (PALL Life Sciences, cat. no. 66234)
- Filter funnels (e.g., Nalgene polysulfone reusable bottle-top filter; Thermo Scientific, cat. no. DS0320-5045, or PALL Life Sciences 47-mm magnetic filter funnels, cat. no. 4238)
- Glass bottle with screw thread (e.g., reusable glass media bottles with cap; Thermo Scientific, cat. no. FB-800-250) for polysulfone filter, or filter flask (e.g., Filter flask Buchner conical shape 1l; Fisher Scientific, cat. no. 12363886) for magnetic filter funnel
- Vacuum pump (e.g., MaximaDry oil-free vacuum pumps; Fisher Scientific, cat. no. 13-880-15)

- Refrigerated centrifuge with rotor and adapters for 50-ml tubes (e.g., Centrifuge 5810 R; Eppendorf, cat. no. 022628171)
- Vortexer (e.g., fixed-speed vortex mixer; Fisher Scientific, cat. no. 02-215-361)
- Operating scissors (e.g., standard dissecting scissors; Fisher Scientific, cat. no. 08-951-20)

### DNA extraction

- Vortexer and adapter (MO-BIO, cat. nos. 13111-V and 13000-V1)
- Microcentrifuge (e.g., Centrifuge 5424; Eppendorf, cat. no. 022620428)
- Water bath (e.g., Isotemp economy analog-control water bath; Fisher Scientific, cat. no. 15-460-2Q)
- Refrigerator, 4 °C (e.g., Isotemp general-purpose series lab refrigerators; Fisher Scientific, cat. no. 11670236)
- DynaMag-2 magnet (Life Technologies, cat. no. 12321D)
- Qubit 2.0 fluorometer (Life Technologies, cat. no. Q32866)
- LightCycler 480 Instrument II and multiwell plates 96 (Roche, cat. nos. 05015278001 and 05102413001)
- 1.5-ml clean microcentrifuge tubes (e.g., 1.5-ml Maxymum Recovery Snaplock microcentrifuge tube; Corning, cat. no. MCT-150-L-C)
- Pipettors (e.g., Research Plus pipettes, 0.1–2.5, 0.5–10, 2–20, 20–200 and 100–1,000 μl; Eppendorf, cat. nos. 3120000011, 3120000020, 3120000097, 3120000054 and 3120000062)
- PCR clean pipette tips (e.g., epT.I.P.S. Racks pipette tips, 0.1–10, 2–200 and 50–1,000 μl; Eppendorf, cat. nos. 022491806, 022491831 and 022491857)

### Sequencing library preparation

- S220 focused ultrasonicator and Snap-Cap microTUBE (Covaris, cat. nos. S220 and 520045)
- Thermal cycler (e.g., Veriti 96-well thermal cycler; Applied Biosystems, cat. no. 4375786)
- 2100 Bioanalyzer (Agilent Technologies, cat. no. G2940CA)
- 0.2-ml PCR tubes (e.g., 0.2-ml Thin-wall PCR Tubes with Flat Cap; Axygen, cat. no. PCR-02-C)

### DNA Sequencing

- Illumina cBot (Illumina)
- Illumina HiSeq 2000 or HiSeq 2500 (Illumina)

### REAGENT SETUP

**Ethanol solution for equipment cleaning, 75% (vol/vol)** Dilute absolute ethanol to a 3:4 ratio in double-distilled water, which can be stored in closed containers at room temperature for up to 1 week before use.

**Ethanol solution for DNA wash, 70% (vol/vol)** Dilute absolute ethanol to a 7:10 ratio in ultrapure water. Freshly prepare the solution before every use.

**PBS buffer, 1×** Dilute 10× ultrapure PBS solution to a 1:10 ratio in ultrapure water. Freshly prepare the solution before every use.

**Illumina adapter, 10 μM** Anneal the Illumina adapter Forward and Reverse primers using the following reagents and program:

Reagent	Volume (μl)
Annealing buffer for DNA oligos (5×)	20
Illumina adapter forward primer (50 μM)	20
Illumina adapter reverse primer (50 μM)	20
Ultrapure water	40
Total	100

	Temperature	Time
Denature	95 °C	2 min
Anneal	95 °C to 25 °C, 0.0125 °C/s	~90 min
Hold	4 °C	∞

PROCEDURE

Particulate matter collection ● TIMING 1 d

- 1| Wrap up to 45 Tissuquartz filters (sufficient for collection for 15 d: one for collecting PM<sub>10</sub>, two for PM<sub>2.5</sub> each day) in aluminum foil individually, and then sterilize them in a muffle furnace at 500 °C for 5 h. Store the sterilized filters in self-sealing plastic bags.
- 2| Before particulate matter collection, clean the gasket of the PM<sub>2.5</sub> fractionating inlet and remove the deposited large particles with double-distilled water and 75% (vol/vol) ethanol. Clean the PM<sub>10</sub> fractionating inlet if needed. Calibrate the PM<sub>2.5</sub> and PM<sub>10</sub> samplers to ensure that the flow rates are at 1.13 m<sup>3</sup>/min (±10%).
- 3| Take the Tissuquartz filters in aluminum foil out of the self-sealing plastic bags and weigh them on a precision balance at milligram accuracy.
- 4| Load the Tissuquartz filters into the filter cartridge, secure the filter holder in the high-volume air sampler and start particulate matter collection for 23 h (Fig. 2a,d).
  - ▲ **CRITICAL STEP** Frequently clean the PM<sub>2.5</sub> and PM<sub>10</sub> sampler inlet gaskets when you are sampling particulate matter at locations with high mass concentration; e.g., clean them once a day or every 2 d.
- 5| After the sample collection is complete, unsecure the filter holder from the high-volume air sampler and transfer it into a biosafety cabinet immediately. Gently open the filter holder, pick up the filter with sterilized forceps and lay it flat on the aluminum foil. Fold the filter with the sample collection side facing inward, and then weigh it again.
  - ▲ **CRITICAL STEP** Use sterile surgical gloves and wear face masks when you are handling the high-volume air sampler and the collected samples.
  - ? **TROUBLESHOOTING**
  - **PAUSE POINT** The samples can be stored in self-sealing plastic bags at –80 °C for several months.

Sample pretreatment ● TIMING 4 h

- ▲ **CRITICAL** Steps 6–12 should be carried out in a decontaminated biosafety cabinet.
- 6| Clean the outside surface of the self-sealing plastic bags containing the particulate matter samples with 75% ethanol (vol/vol) before putting it into the biosafety cabinet.
  - 7| Open the self-sealing plastic bag in the biosafety cabinet, take out the aluminum foil from the bag and open it, and then unfold the filter with sterilized forceps.
  - 8| Cut out one-quarter of the filter with sterilized scissors, roll it with sterilized forceps and insert the rolled-up filter into a 50-ml Falcon tube (Fig. 2e,f).
    - ▲ **CRITICAL STEP** The Tissuquartz filters are fragile, especially after freezing, and thus it is important to avoid extensive tearing of the filter membrane while cutting.
  - 9| Add 50 ml of ultrapure 1× PBS buffer into the tube, screw the lid on tightly and centrifuge it for 3 h at 200g at 4 °C to pellet the particulate matter.
    - ▲ **CRITICAL STEP** 200g is an optimized centrifugation speed. Centrifugation at too high a speed will result in disruption of the Tissuquartz filter, and thus it will lower the extraction efficiency.
  - 10| Remove a piece of 0.2-µm Supor 200 PES membrane disc filter from its pack with sterilized forceps and put it onto the gauze of a sterilized filter funnel, screw the funnel on a 1-l glass bottle and connect the funnel to a vacuum pump. When you are using the magnetic filter funnel, connect the filter funnel to a 1-l filter flask connected to a vacuum pump (Fig. 2g).
  - 11| After centrifugation, gently vortex the tube from Step 9 and pour the resuspension into the funnel from Step 10 without disturbing the collecting filter, and start suction filtration until all of the liquid filters through the PES filter (Fig. 2h).
    - ▲ **CRITICAL STEP** Control the suction speed so that a volume of ~50 ml of buffer suspension goes through the PES filter in 1–2 min. We recommend air-drying the membrane briefly before cutting it into small pieces in the next step, as wet filter membranes tend to stick to the scissors.

? TROUBLESHOOTING

## PROTOCOL

**12|** Switch off the pump, transfer the PES filter into a sterilized culture dish with sterilized forceps, cut the filter into small pieces (~0.20 cm<sup>2</sup> each) with sterilized scissors and put them into a clean PowerBead tube from the MO-BIO PowerSoil DNA isolation kit.

### DNA extraction ● TIMING 2 h

**13|** Extract DNA using the MO-BIO PowerSoil DNA isolation kit and Agencourt AMPure XP beads, modifying the manufacturer's instructions as detailed in the following steps. Add an aliquot of 60 µl of Solution C1 into each PowerBead tube containing the PES pieces, and incubate the tube in a water bath at 65 °C for 15 min.

▲ **CRITICAL STEP** Screw the cap of the PowerBead tube tightly to make sure that the tube does not dry up when heated in a water bath at 65 °C.

**14|** Secure the PowerBead tubes on a vortexer using the MO-BIO vortex adapter, and then vortex the tubes at maximum speed for 15 min. Alternatively, the tubes can be fastened onto an ordinary vortexer with tape.

**15|** Centrifuge PowerBead tubes at 10,000g for 30 s at room temperature. Transfer the supernatant from each tube to a clean 2-ml collection tube from the kit.

▲ **CRITICAL STEP** The centrifuge speed should not exceed 10,000g, or it may damage the PowerBead tubes and cause leakage.

**16|** Add an aliquot of 250 µl of Solution C2 to each collection tube and vortex it briefly for 5 s. Incubate the mixture at 4 °C for 5 min. Centrifuge the tubes for 60 s at 10,000g at room temperature. The precipitate should appear black in color. Transfer the supernatant from each tube to a clean 2-ml collection tube, making sure to avoid the pellet.

**17|** Add an aliquot of 200 µl of Solution C3 to each collection tube and vortex it briefly. Incubate at 4 °C for 5 min. Centrifuge the tubes for 60 s at 10,000g at room temperature. The precipitate should appear milk-white or white in color. Taking care to avoid the pellet, transfer the supernatant from all tubes into a clean 2-ml collection tube. Distribute the 700 µl supernatant equally into two 1.5-ml clean microcentrifuge tubes.

**18|** Vortex to mix the AMPure XP beads thoroughly at room temperature before use. Add 500 µl of AMPure XP beads to each tube, and mix thoroughly by pipetting ten times. Incubate the tubes at room temperature for 5 min.

▲ **CRITICAL STEP** Use AMPure XP beads instead of a column to improve the DNA recovery efficiency.

**19|** Place the microcentrifuge tubes on a magnetic rack for 3 min until AMPure XP beads are separated from the solution, forming a plaque on the side of the microcentrifuge tube until the solution clears.

**20|** Aspirate the clear solution slowly from the microcentrifuge tubes and discard it, without disturbing the AMPure XP beads.

**21|** Add an aliquot of 500 µl of 70% (vol/vol) ethanol. Rinse the beads by turning the tubes 90°, and allow the beads to re-collect on the side of the tubes. Turn the tubes eight more times, and then discard the ethanol by aspiration. Repeat this wash step once.

**22|** Remove the microcentrifuge tubes from the magnetic rack. Quickly spin down the beads, and then place the microcentrifuge tubes back on the magnetic rack and remove any remaining liquid. Keep the microcentrifuge tubes on the magnetic rack with the cap open. Air-dry the beads at room temperature for 5 min.

▲ **CRITICAL STEP** Do not over dry the beads, as this can substantially affect the DNA recovery efficiency.

**23|** Remove the microcentrifuge tubes from the magnetic rack. Add an aliquot of 25 µl of DNase-free water to each of the two tubes and pipette to mix ten times. Incubate the tubes at room temperature for 2 min.

▲ **CRITICAL STEP** The use of 25 µl instead of a higher volume of ultrapure water for elution results in sufficient DNA concentration for downstream analysis and library preparation.

**24|** Place the microcentrifuge tubes onto the magnetic rack for 1 min to separate the beads from the solution. Transfer the combined eluate from both tubes to a new clean microcentrifuge tube.

25| Pipette 1  $\mu$ l of the eluate from each tube, and quantify the DNA concentration using a Qubit 2.0 fluorometer with the dsDNA HS assay kit. One may also perform qPCR (optional) using 16S rRNA gene universal primers, such as the 515F and 806R primer set<sup>30</sup>, using the following reagents and thermal cycling conditions:

Reagent	Volume ( $\mu$ l)
LightCycler 480 SYBR Green I master mix	10
515F primer (10 $\mu$ M)	0.3
806R primer (10 $\mu$ M)	0.3
Template	1
ddH <sub>2</sub> O	8.4
Total	20

	Cycles	Temperature ( $^{\circ}$ C)	Time
Preincubation	1	95	5 min
Amplification	60	95	10 s
		50	20 s
		72	30 s
Melting curve	1	95	5 s
		65	1 min
		97	(acquisition)
Cooling	1	40	10 s

### ? TROUBLESHOOTING

■ **PAUSE POINT** The extracted genomic DNA can be stored at  $-80^{\circ}$  C for several months, although repeated freeze-thawing of the sample is not recommended.

### Sequencing library preparation ● TIMING 6 h

▲ **CRITICAL** The following protocol is for library preparation for sequencing on the Illumina MiSeq or HiSeq platforms. Library preparation for other alternative sequencing platforms should follow the protocol provided by the manufacturer. Pay attention to the amount of input DNA when choosing alternative library preparation and sequencing platforms.

26| Transfer the extracted genomic DNA solution of  $\sim 50$   $\mu$ l from Step 24 to a snap-cap microTUBE tube, and sonicate the sample using the S220 Focused-ultrasonicator using the parameters below:

<b>Peak incident power (W)</b>	<b>175</b>
Duty factor	5%
Cycles per burst	200
Treatment time (s)	35
Temperature ( $^{\circ}$ C)	7
Water level-S220	12
Water level-E220	6
Sample volume ( $\mu$ l)	50
E220-intensifier (pn500141)	Yes



## PROTOCOL

**27|** Transfer the DNA sample from the microTUBE into a 0.2-ml nuclease-free PCR tube, and complete the end-repair, dA-tailing, adapter ligation and PCR reactions, as described below. Alternatively, one may use the NEBNext ultra DNA library prep kit for Illumina (by following the steps for <50 ng of input DNA) for the library preparation.

**28|** For end repair, first mix the following reagents in a 0.2-ml nuclease-free PCR tube and incubate the mixture in a thermal cycler for 30 min at 20 °C. Use a 1.6× volume (80 μl) of AMPure XP beads to purify the end-repaired product before adding 43 μl of ultrapure water for dilution.

Reagent	Volume (μl)
Fragmented DNA	42.5
NEBNext End Repair Reaction Buffer (10×)	5
NEBNext End Repair Enzyme Mix	2.5
Total	50

**▲ CRITICAL STEP** Use AMPure XP beads to improve the DNA recovery efficiency.

**29|** Take an aliquot of 42 μl of end-repaired product from Step 28 for adding the dA tail. Mix the following reagents in a 0.2-ml nuclease-free PCR tube and incubate the mixture in a thermal cycler for 30 min at 37 °C. Use a 1.6× volume (80 μl) of AMPure XP beads to purify the product before adding 22 μl of ultrapure water for dilution.

Reagent	Volume (μl)
End-repaired DNA	42
NEBNext dA-Tailing Reaction Buffer (10×)	5
Klenow fragment (3'→5' exo-)	3
Total	50

**30|** Take an aliquot of 21 μl of the product from Step 29 for adapter ligation. Mix the following reagents in a 0.2-ml nuclease-free PCR tube and incubate the mixture in a thermal cycler for 30 min at 20 °C. Use a 1.0× volume (50 μl) of AMPure XP beads to purify the product before adding 24 μl of ultrapure water for dilution.

Reagent	Volume (μl)
DNA with added dA tail	21
2× Quick Ligation Buffer	25
Quick T4 DNA ligase	3
Illumina adapter (10 μM)	1
Total	50

**31|** Take an aliquot of 23 μl of the product from Step 30 for limited-cycle PCR. Mix the following reagents in a 0.2-ml nuclease-free PCR tube.

Reagent	Volume (μl)
DNA with Illumina adapters	23
Q5 high-fidelity 2× master mix	25
Illumina universal primer (10 μM)	1
Illumina index primer (10 μM)	1
Total	50

Perform limited-cycle PCR using the following program:

Cycles	Temperature (°C)	Time
1	98	30 s
12	98	10 s
	65	30 s
	72	50 s
1	72	5 min
1	4	∞

**▲ CRITICAL STEP** Perform limited-cycle PCR (no more than 12 cycles) to avoid amplification bias and loss of sequence diversity.

**32|** After PCR, add a 0.5× volume (25 µl) of AMPure XP beads to the reaction tube, mix well and incubate the tube at room temperature for 10 min. After incubation, transfer the mixture to a new 1.5-ml microcentrifuge tube, and place the tube on the magnetic rack to separate the beads from the supernatant. Wait for the solution to clear, before transferring the supernatant containing the library DNA to a new tube and discard the beads with the unwanted large DNA fragments.

**▲ CRITICAL STEP** Although other size-selection methods may be used, we recommend using the AMPure XP beads for optimized recovery efficiency. Pay particular attention to Steps 32 and 33, in which either the beads or the supernatant is discarded, respectively.

**33|** Add 5 µl of AMPure XP beads to the tube containing the supernatant from Step 32, mix them well and incubate the tube at room temperature for 10 min. Place the tube on a magnetic rack to separate the beads from the supernatant. Wait for the solution to clear, before discarding the supernatant with the unwanted small DNA fragments. Add 70% (vol/vol) ethanol to wash the beads twice. Air-dry the beads before adding 25 µl of ultrapure water to dilute the product.

**▲ CRITICAL STEP** Do not overdry the beads, as this can markedly affect the DNA recovery efficiency.

**34|** Take an aliquot of 1 µl of the product from Step 33, and analyze it on the Agilent 2100 Bioanalyzer to validate the library quality (**Fig. 3b–d**).

**? TROUBLESHOOTING**

**■ PAUSE POINT** The prepared libraries can be stored at –20 °C for several months.

**DNA SEQUENCING ● TIMING 8 d**

**35|** Use the TruSeq PE cluster kit v3-cBot-HS for cluster generation and the TruSeq SBS kit v3-HS for sequencing with a 2 × 101 strategy.

**▲ CRITICAL STEP** Because of the apparent sequence diversity of the airborne microbial community, we recommend generating at least 5–7 Gb of data per sample. When you are choosing other sequencing platforms, note that the quantity of available input airborne microbial genomic DNA is typically limited to 5–50 ng.

**36|** Remove the adapter sequences, low-quality reads and duplicate reads. For the analysis of microbial species, align the Illumina reads to a cohort of nonredundant NCBI complete genomes using the SOAP alignment tool (release 2.21t). Calculate the genome coverage using the SOAP.coverage package (version 2.7.7), in which only uniquely aligned reads are calculated. One can identify bacterial, fungal and viral species with coverage of ≥5%, ≥0.5% (average alignment of all chromosomes) and ≥1%, respectively.

**? TROUBLESHOOTING**

Troubleshooting advice can be found in **Table 2**.



**TABLE 2** | Troubleshooting table.

Step	Problem	Possible reason	Solution
5	Collected PM <sub>2.5</sub> mass is equal to or higher than collected PM <sub>10</sub> mass	The PM <sub>2.5</sub> inlet gasket is not cleaned in time, which leads to an increase of the inlet cutoff size and large particle bounce	We recommend cleaning the PM <sub>2.5</sub> inlet gasket every day or every 2 d, especially for sampling particulate matter at locations with high mass concentration
	There is an abnormal mass of the collected PM <sub>2.5</sub> and PM <sub>10</sub> samples	There may have been a precision balance malfunction, or sampler flow rates are different from the nominal flow rate (1.13 m <sup>3</sup> /min)	Recalibrate and check whether the precision balance is functioning properly. Calibrate the sampler flow rate at least once per month
11	Failure to filter the buffer suspension through the PES filter	There are too many quartz fibers in the buffer suspension, which become trapped on the PES filter during suction filtration	We recommend 200g as the centrifugation speed in Step 9 in order to ensure minimum disruption of the quartz filter
25, 34	Reduced DNA recovery efficiency	Overdrying the AMPure XP beads	We recommend air-drying the beads at room temperature for 5 min
25	No or too little extracted genomic DNA	Insufficient input sample, ineffective sample pretreatment or ineffective DNA extraction	We recommend checking the net weight of the input particulate matter sample, making sure that the sample pretreatment steps are carried out effectively, and ensuring that the steps of the DNA extraction are carried out carefully
34	Low-quality sequencing library as a result of adapter contamination	Ineffective removal of adapter sequences	Use AMPure XP beads to purify the library again

## ● TIMING

Steps 1–5, day 1, particulate matter collection: 1 d

Steps 6–12, day 2, sample pretreatment: 4 h

Steps 13–25, day 2, DNA extraction: 2 h

Steps 26–34, day 2, sequencing library preparation: 6 h

Steps 35 and 36, days 3–10, DNA sequencing: 8 d

## ANTICIPATED RESULTS

The protocol described here enables the extraction of airborne microbial genomic DNA from collected particulate matter for metagenomic sequencing. In our experience, most of the particulate matter samples collected daily in the city of Beijing weigh in the range of tens to hundreds of milligrams, and each sample typically yields 5–50 ng of genomic DNA. By using the improved library preparation protocol described here, this quantity is sufficient for direct library preparation without requiring 16S rRNA gene amplification or whole-genome amplification. So far, we have applied this protocol to the extraction of airborne microbial genomic DNA from a total of 30 particulate matter samples (half of which were PM<sub>2.5</sub> and the other half were PM<sub>10</sub>). ~90% of these samples yielded genomic DNA of ≥5 ng, were qPCR positive and produced good-quality sequencing libraries (**Fig. 3**). We show that although it is possible to extract DNA directly from the collected particulate matter samples by shredding the quartz filter without pretreatment, the extraction efficiency can be affected (**Table 1, Fig. 3a**), probably owing to the presence of DNA-binding quartz fibers. Moreover, extracting DNA using the MO-BIO PowerMax kit without pretreatment would require the use of >10-fold more AMPure XP beads per sample owing to the larger sample volume, and thus we recommend adding the wash and filter recollection steps for the higher extraction efficiency and lower cost. We also show that the use of AMPure XP beads to replace a column for DNA recovery in the DNA extraction and library preparation steps improves the yield (**Table 1, Fig. 3a**); otherwise, higher amounts of input materials would be required for library preparation and metagenomic sequencing.

**ACKNOWLEDGMENTS** We thank M. Kayani, C. Liu and C. Cao for helpful discussions and for help taking the experimental setup photos. This work was supported in part by funding from the Ministry of Science and Technology of China (973 grant no. 2015CB553402), the National Natural Science Foundation of China (grant nos. 31470532, 21190054, 21221004), the Tsinghua University-Peking University Center for Life Sciences (CLS), a Tsinghua-Janssen Scholarship and research grant and the Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases.

**AUTHOR CONTRIBUTIONS** W.J., P.L., B.W., J.F. and J.L., with inputs from T.F.Z., J.J. and G.T., performed the experiments and analyzed the data. T.F.Z. conceived and organized the study. W.J. and T.F.Z., with inputs from the other coauthors, wrote the paper.

**COMPETING FINANCIAL INTERESTS** The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Eisen, J.A. Environmental shotgun sequencing: its potential and challenges for studying the hidden world of microbes. *PLoS Biol.* **5**, e82 (2007).
- Yooseph, S. *et al.* Genomic and functional adaptation in surface ocean planktonic prokaryotes. *Nature* **468**, 60–66 (2010).
- Christner, B.C. *et al.* A microbial ecosystem beneath the West Antarctic ice sheet. *Nature* **512**, 310–313 (2014).
- Qin, J. *et al.* A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**, 59–65 (2010).
- Fierer, N. *et al.* Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proc. Natl. Acad. Sci. USA* **109**, 21390–21395 (2012).
- Fierer, N. *et al.* Comparative metagenomic, phylogenetic and physiological analyses of soil microbial communities across nitrogen gradients. *ISME J.* **6**, 1007–1017 (2012).
- Foster, J.S. *et al.* Molecular and morphological characterization of cyanobacterial diversity in the stromatolites of Highborne Cay, Bahamas. *ISME J.* **3**, 573–587 (2009).
- Cao, C. *et al.* Inhalable microorganisms in Beijing's PM<sub>2.5</sub> and PM<sub>10</sub> pollutants during a severe smog event. *Environ. Sci. Technol.* **48**, 1499–1507 (2014).
- Bertolini, V. *et al.* Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy. *Appl. Microbiol. Biotechnol.* **97**, 6561–6570 (2013).
- Bowers, R.M., McLetchie, S., Knight, R. & Fierer, N. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *ISME J.* **5**, 601–612 (2011).
- Brodie, E.L. *et al.* Urban aerosols harbor diverse and dynamic bacterial populations. *Proc. Natl. Acad. Sci. USA* **104**, 299–304 (2007).
- Deleon-Rodriguez, N. *et al.* Microbiome of the upper troposphere: species composition and prevalence, effects of tropical storms, and atmospheric implications. *Proc. Natl. Acad. Sci. USA* **110**, 2575–2580 (2013).
- Robertson, C.E. *et al.* Culture-independent analysis of aerosol microbiology in a metropolitan subway system. *Appl. Environ. Microbiol.* **79**, 3485–3493 (2013).
- Bowers, R.M. *et al.* Seasonal variability in bacterial and fungal diversity of the near-surface atmosphere. *Environ. Sci. Technol.* **47**, 12097–12106 (2013).
- Steffen, M.M. *et al.* Comparative metagenomics of toxic freshwater cyanobacteria bloom communities on two continents. *PLoS ONE* **7**, e44002 (2012).
- Rinke, C. *et al.* Obtaining genomes from uncultivated environmental microorganisms using FACS-based single-cell genomics. *Nat. Protoc.* **9**, 1038–1048 (2014).
- Leung, M. & Chan, A.H. Control and management of hospital indoor air quality. *Med. Sci. Monit.* **12**, SR17–23 (2006).
- Peiris, J.S. *et al.* Coronavirus as a possible cause of severe acute respiratory syndrome. *Lancet* **361**, 1319–1325 (2003).
- Uyeki, T.M. & Cox, N.J. Global concerns regarding novel influenza A (H7N9) virus infections. *N. Engl. J. Med.* **11**, 11 (2013).
- He, K. *et al.* The characteristics of PM<sub>2.5</sub> in Beijing, China. *Atmos. Environ.* **35**, 4959–4970 (2001).
- Wang, S. *et al.* Quantifying the air pollutants emission reduction during the 2008 Olympic games in Beijing. *Environ. Sci. Technol.* **44**, 2490–2496 (2010).
- Yang, F. *et al.* Characteristics of PM<sub>2.5</sub> speciation in representative megacities and across China. *Atmos. Chem. Phys.* **11**, 5207–5219 (2011).
- Tolocka, M.P., Peters, T.M., Vanderpool, R.W., Chen, F.L. & Wiener, R.W. On the modification of the low-flow-rate PM<sub>10</sub> dichotomous sampler inlet. *Aerosol. Sci. Technol.* **34**, 407–415 (2001).
- Metzker, M.L. Sequencing technologies - the next generation. *Nat. Rev. Genet.* **11**, 31–46 (2010).
- Metzker, M.L. Emerging technologies in DNA sequencing. *Genome Res.* **15**, 1767–1776 (2005).
- Segata, N. *et al.* Metagenomic microbial community profiling using unique clade-specific marker genes. *Nat. Methods* **9**, 811–814 (2012).
- Segata, N. *et al.* Computational meta'omics for microbial community studies. *Mol. Syst. Biol.* **9**, 666 (2013).
- Meyer, F. *et al.* The metagenomics RAST server - a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* **9**, 1–8 (2008).
- Li, R. *et al.* SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* **25**, 1966–1967 (2009).
- Caporaso, J.G. *et al.* Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **108**, 4516–4522 (2011).