nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-020-00451-8

Check for updates

Decoding personal biotic and abiotic airborne exposome

Chao Jiang[1,3]✉, Xinyue Zhang[2,3], Peng Gao[2,3], Qiong Chen[1] and Michael Snyder[2]✉

The complexity and dynamics of human diseases are driven by the interactions between internal molecular activities and external environmental exposures. Although advances in omics technology have dramatically broadened the understanding of internal molecular and cellular mechanisms, understanding of the external environmental exposures, especially at the personal level, is still rudimentary in comparison. This is largely owing to our limited ability to efficiently collect the personal environmental exposome (PEE) and extract the nucleic acids and chemicals from PEE. Here we describe a protocol that integrates hardware and experimental pipelines to collect and decode biotic and abiotic external exposome at the individual level. The described protocol has several advantages over conventional approaches, such as exposome monitoring at the personal level, decontamination steps to increase sensitivity and simultaneous capture and high-throughput profiling of biotic and abiotic exposures. The protocol takes ~18 h of bench time over 2–3 d to prepare samples for high-throughput profiling and up to a couple of weeks of instrumental time to analyze, depending on the number of samples. Hundreds to thousands of species and organic compounds could be detected in the airborne particulate samples using this protocol. The composition and complexity of the biotic and abiotic substances are heavily influenced by the sampling spatiotemporal factors. Basic skillsets in molecular biology and analytical chemistry are required to carry out this protocol. This protocol could be modified to decode biotic and abiotic substances in other types of low or ultra-low input samples.

Introduction

Human diseases are caused by the interplay between intrinsic molecular dynamics and environmental exposures, which can be broadly defined as biological, chemical, physical and even psychological exposures[1–3]. Abnormal environmental exposures, especially during early life, can lead to a variety of acute and chronic diseases, including cancer, allergy/asthma and autoimmune diseases[4–8]. Therefore, it is imperative to develop high-throughput methods to characterize the total environmental exposures with high sensitivity and accuracy.

In 2006, Christopher Wild first introduced the concept of the exposome, which comprises all environmental exposures across one's lifetime[9]. The human exposome is highly dynamic and may be further divided into internal, specific external and general external exposomes to better characterize the nature of exposures[10]. Early exposome studies measured environmental exposures via questionnaires as the self-reported method; environmental monitoring of known toxins and pathogens by targeted methods; monitoring unknown biotic and abiotic stressors by untargeted methods; and Geographic Information System modeling of environmental pollutants using stationary monitoring or wearable technology to quantify personalized physiological reactions to the external exposome[4]. Information of millions of organic chemicals—including pharmaceuticals and personal care products, plasticizers, pesticides, preservatives, flame retardants and microbial metabolites—is already collected in various databases, However, most of the natural organic chemicals remain unannotated[4]. More recently, a larger array of population-based exposome studies have been launched across the globe, such as the European Exposome Project—Human Early Life Exposome (https://www.projecthelix.eu/), exposome projects at Emory University (https://emoryhercules.com/) and the Icahn School of Medicine at Mount Sinai (https://icahn.mssm.edu/research/exposomic) and the Human Health Exposure Analysis Resource program launched by the National Institute of Environmental Health Sciences (https://hhearprogram.org/) to facilitate the analysis of exposure data.

[1]Life Sciences Institute, Zhejiang University, Hangzhou, China. [2]Department of Genetics, Stanford University School of Medicine, Stanford, CA, USA. [3]These authors contributed equally: Chao Jiang, Xinyue Zhang, Peng Gao. ✉e-mail: jiang_chao@zju.edu.cn; mpsnyder@stanford.edu

In the field of medical science, studies of the exposome in the context of human health usually measure the epigenetic changes inside the human body as proxy data of environmental exposures[11,12]. In the fields of environmental science and ecology, traditional methods for profiling outdoor and indoor environmental exposures use high-volume air sampling equipment or occasionally smaller portable devices, but the research focus is either on the microbial or chemical components of the exposome, never measuring both concurrently[13–19]. On a larger scale, previous studies, such as the European EXPOLIS study[20], were focused mostly on the regional scale or even the global scale of air pollution and its effect on population health rather than on precision medicine[21]. In these studies, tens to thousands of biological species or chemicals could be detected in the samples, depending on the adopted analytical method. Specifically, marker gene-based detection methods, such as 16S and 18S ribosomal RNA (rRNA) gene-specific amplification and sequencing, although contributing greatly to the understanding of the identity of airborne microorganisms in various environments[22,23], have also limited the ability to collect pan-domain airborne environmental exposures and to gain knowledge from the vast and dynamic environmental genetic pool. Similarly, targeted chemical analysis of environmental samples does not fully reflect the universe of thousands of chemicals and does not detect emerging contaminants in the exposome[24,25]. Moreover, the biological and chemical exposome are tightly connected, partly owing to the seasonal chemical compounds produced by plants and fungi in the environment, and certain synthetic chemicals can inhibit the growth of specific microbes[1]. However, methods for directly measuring the dynamics and diversity of biotic and abiotic airborne environmental exposures, especially at the individual level, are lacking.

To address various challenges in decoding the personal airborne environmental exposures, we developed a highly sensitive protocol that integrates a wearable device called an exposometer, which is fitted with custom-designed parts, and optimized experimental processes to collect and decode biotic and abiotic environmental exposures (Figs. 1–3; details below). Most previously reported methods are limited to stationary sampling methods and nanograms of extracted genomic DNA materials[23,26,27]. In contrast, our sampling method is highly mobile and increased the sensitivity to the sub-nanogram level by incorporating dedicated decontamination steps[28,29] and isothermal linear amplifications for DNA and RNA before sequencing library preparations (Fig. 4a–d). The exposometer also included a few sensors to measure temperature, humidity, particulate matter (PM) concentration and duration during collection, which can be used for downstream data analysis. Our method for decoding chemicals can detect specific scent molecules released from orange peels in high abundance when compared to the control (Fig. 4e–g), indicating the ability to capture various chemicals. We implemented this protocol to track the personal environmental exposome from 15 individuals for up to 2 years. We identified over 2,500 species and annotated nearly 1,000 compounds, revealing the highly complex and dynamic spatiotemporal personal exposome for the first time[1].

Applications

Implementation of this protocol in various contexts could provide nonpareil exposure data to help better understand diseases that are heavily influenced by environmental exposures, such as allergies/asthma, autoimmune diseases, lung diseases and many types of cancer. Owing to the ubiquity of environmental exposures and their effect on humans, animals, plants and the environment in general, this protocol is broadly applicable to several fields of science, including, but not limited to, medical science, environmental science, public health, environmental DNA/RNA fields and agriculture.

We used the MicroPEM to collect airborne particles. The device is about the size of a smartphone and lightweight enough to be carried around for everyday activities. A typical exposometer uses a power source, either a pump or a fan that draws in air, which goes through a few stages of impactors for size fractionation (MicroPEM comes with PM2.5 or PM10 option). A particle counter collects scattered light to determine the concentration of particles in real time. A filter cassette is placed downstream of the particle counter to collect particles of interest. We added a custom-designed cartridge for chemical collection at the end of airflow and removed the oiled impactors for maximum particle collection. The choice of device is up to the user, as the protocol can be broadly adapted to filters below 47 mm in diameter or larger filters if cut into pieces or pre-treated, as previously described[23]. The experimental part of this protocol can also be adapted in studying low microbial biomass found in human body fluid/swab samples, such as vaginal swabs, lung microbiome and low-biomass skin microbiome. Lastly, the biological and chemical parts of the protocol can be performed separately if researchers deem appropriate.

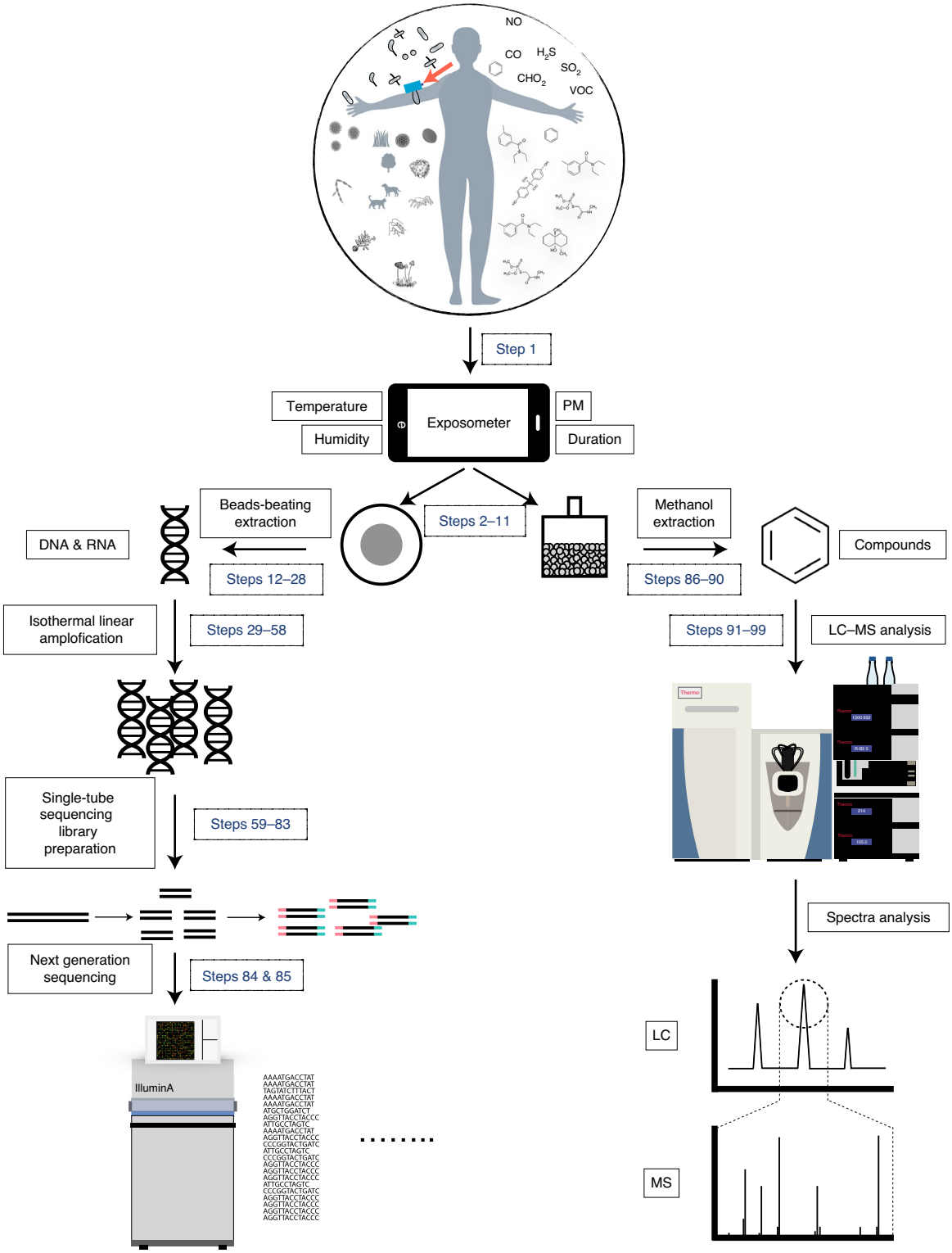


Fig. 1 | An overview of the personal exposome collection, extraction and processing pipeline. Briefly, an exposometer of choice, which preferably also monitors the surrounding temperature, humidity, PM and duration of the collection, is deployed to the individual to collect the personal exposome. A filter (left) and a customized cartridge (right) filled with zeolite beads are used to collect biologicals and chemicals, respectively. DNA/RNA and chemicals are extracted and processed using optimized steps as detailed in the protocol. The nucleic acids and chemical compounds are then quantitated using high-throughput non-targeted methods by NGS and LC–MS platforms, respectively.

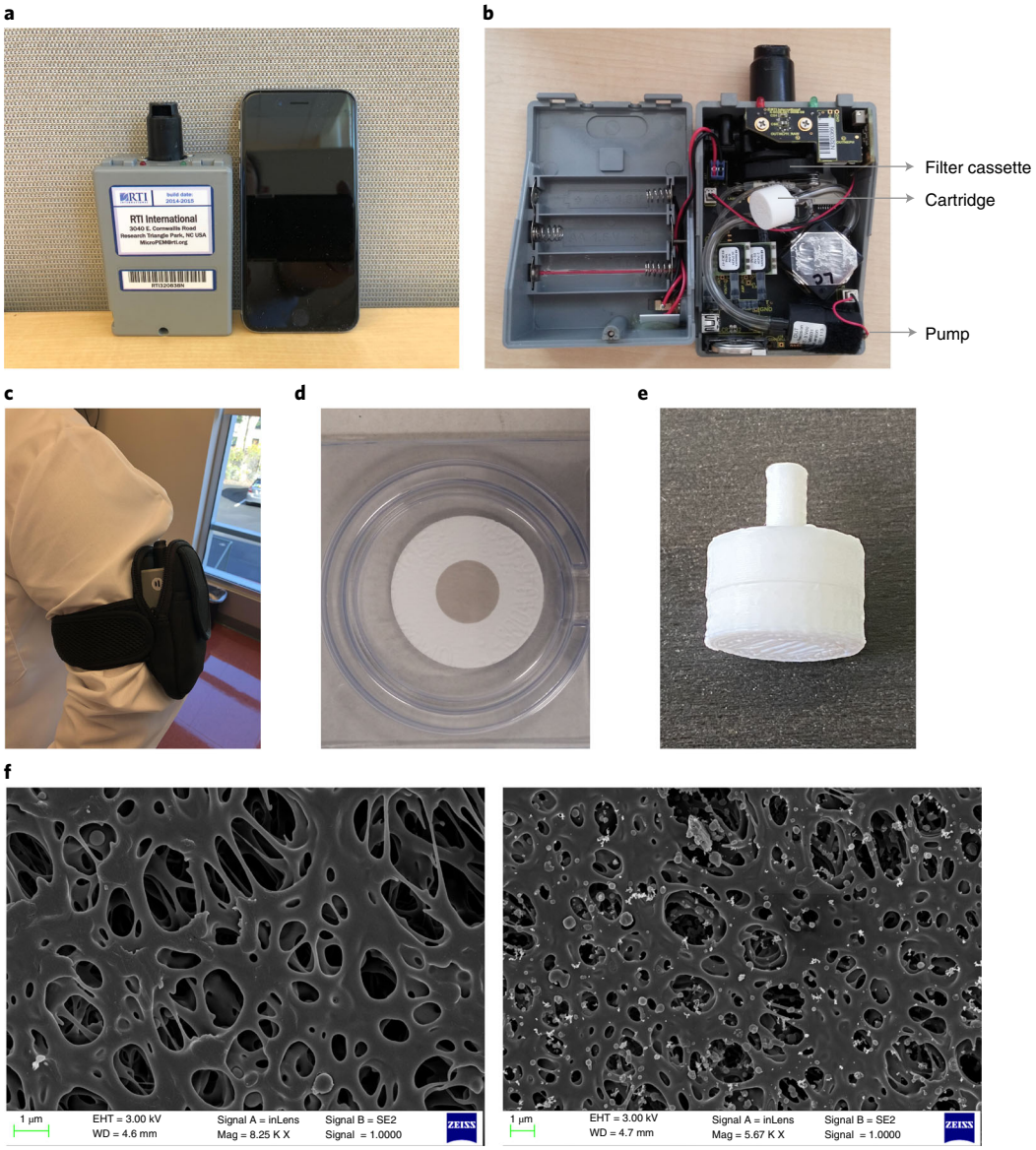


Fig. 2 | Overview of the modified MicroPEM exposometer to include sampling devices for biologicals and chemicals. a, The exposometer MicroPEM is placed next to a smartphone to show its relative size. **b**, Inside view of the modified MicroPEM. A 3D-printed cartridge that houses zeolite beads (molecular sieves for chemical exposures collection) is placed downstream of the filter cassette (for biological exposures collection). The position of the miniaturized pump is also indicated. **c**, The device can be conveniently strapped to the arm for personal exposome collection. Alternatively, the device can be placed next to an individual. **d**, Representative picture of a used filter (the circular gray area in the middle indicates PM collection). **e**, A close-up view of the 3D-printed cartridge for chemical exposure collection. **f**, Unused (left) and used (right) filters under a scanning electron microscope. Scale bars, 1 µm. Figure 2d adapted with permission from ref. [1].

Advantages and limitations

Compared to other exposome monitoring approaches[4,13–15,30], the main advantage of the protocol is the ability to collect and characterize personal environmental exposures in an unprecedented resolution. Our method collects and analyzes both biological and chemical airborne exposomes at the individual level. The main limitation regarding biotics detection is that the current protocol is relatively underpowered in detecting low-abundance organisms owing to the mandatory amplifications steps. Specifically, human RNA viruses were not readily detected in our previous study, although abundant plant and fungal RNA viruses were present[1]. Another limitation is that the extraction protocol could introduce some domain biases because different organisms require different optimizations of physical force and chemical reagents to break apart most efficiently. In terms of the

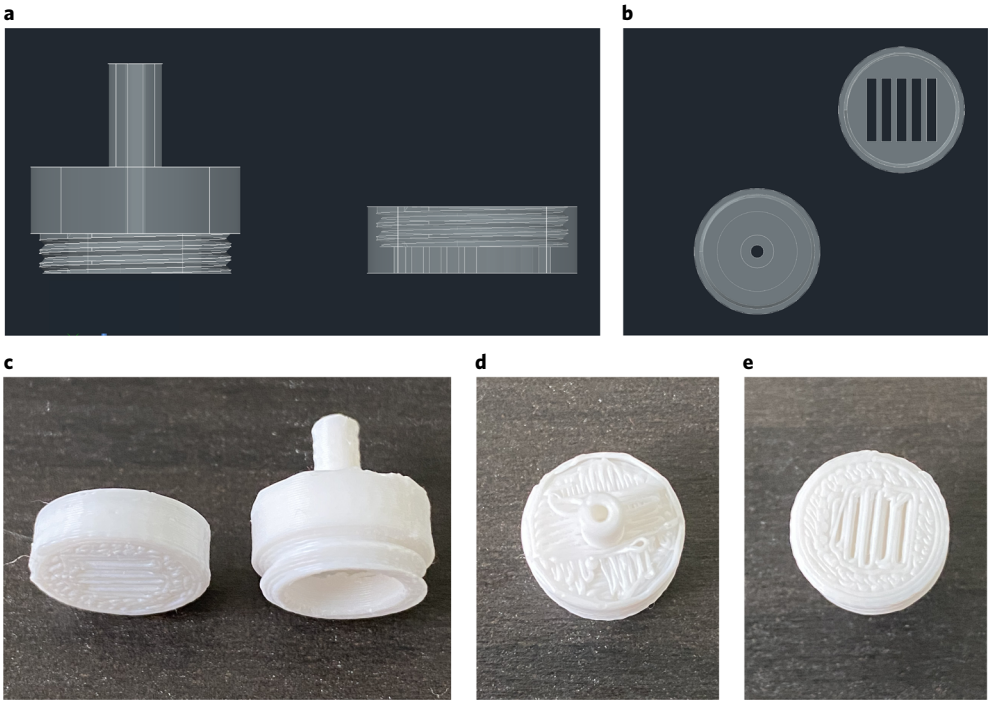


Fig. 3 | The design of the chemical cartridge and a 3D-printed product. a,**b**, The side (**a**) and top (**b**) views of the cartridge design. Note that the cartridge is composed of two screw-on parts (**c**–**e**). The side (**c**), top (**d**) and bottom (**e**) views of a 3D-printed cartridge. The cartridge was designed using AutoCAD software.

abiotic exposome, the main limitation is that, because the chemical cartridge is connected downstream of the biotic collection filter, only air-dissolved compounds from particle-free airflow can be analyzed, and compounds adsorbed to the particulates larger than the pore size will be retained in the biotic collection filter. The chosen adsorbent and extractant might not be able to capture and extract all the chemicals in exposome samples, such as inorganic components[31,32]. Besides, the applied liquid chromatography–mass spectrometry (LC–MS) method might not be able to detect some low-polarity or volatile chemicals as well as trace-level chemicals below the detection limits.

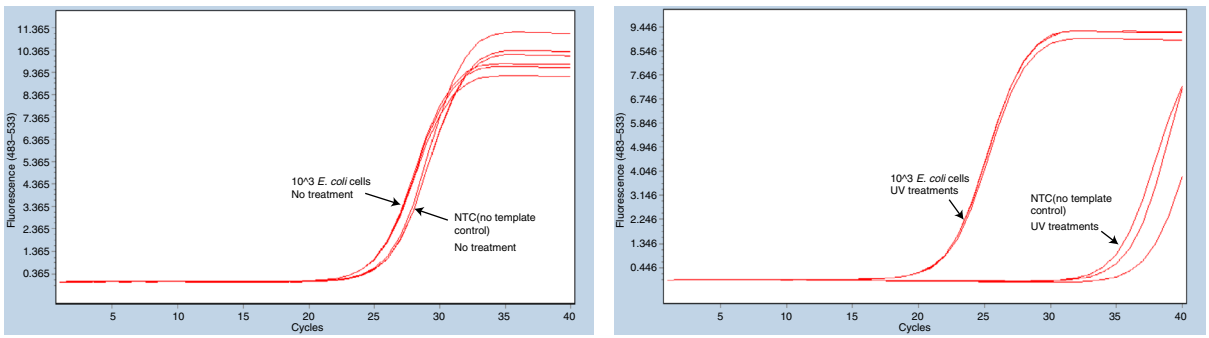
Overview of the protocol

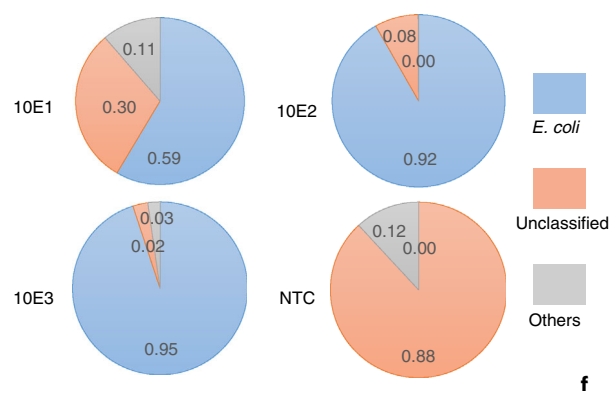
Here we provide detailed steps for the collection of the personal airborne exposome and the extraction of biological and chemical molecules for next-generation sequencing (NGS) and LC–MS analysis (Fig. 1), respectively. The protocol outlines the setup and deployment of the device (Step 1), extraction of DNA/RNA (Steps 2–28) and chemicals from the collected materials (Steps 86–90), isothermal linear amplification of DNA/RNA (Steps 29–58), preparation of the sequencing libraries for NGS (Steps 59–85) and pre-treatment of the chemical collection samples for LC–MS analysis (Steps 91–99). We also discuss the choices of methods for bioinformatic analyses and the databases for data annotations.

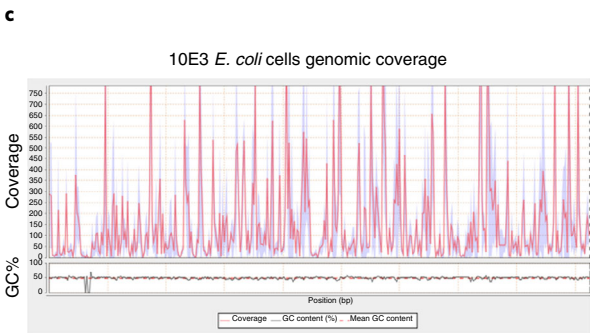
Experimental design

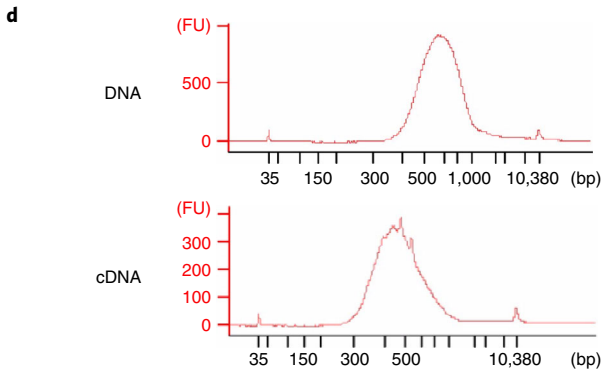
Personal exposome collection

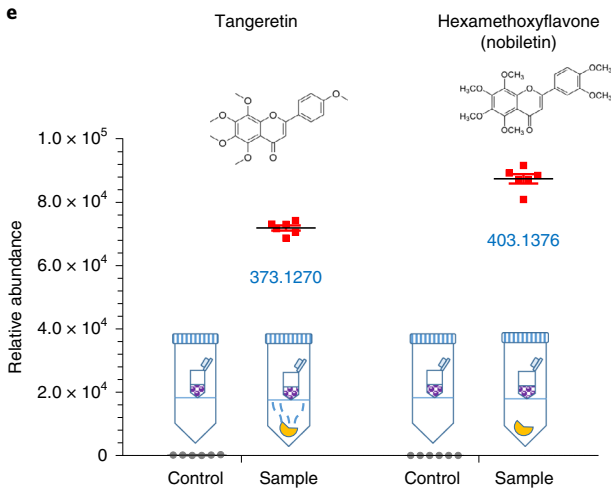
To efficiently collect personal exposome, an ideal device would be wearable, able to collect both biotic and abiotic agents, operate at a flow speed close to the human breathing rate, have low energy consumption and emit minimal noise. There are many research-grade (e.g., TSI SidePak) and consumer-grade (e.g., PurpleAir) PM monitors on the market, but only a few have a pre-designed space for collection filters. The options are even more limited when it comes to the size and weight of the device that is comfortable for daily wearing. In our previous study[1], we used a personal aerosol monitor, MicroPEM, to collect personal airborne exposures (Figs. 1 and 2). The device was modified to allow for the collection of both airborne chemicals and biologicals, and the pre-installed impactors were removed to allow maximum particulates collection. We used zeolite adsorbent to collect

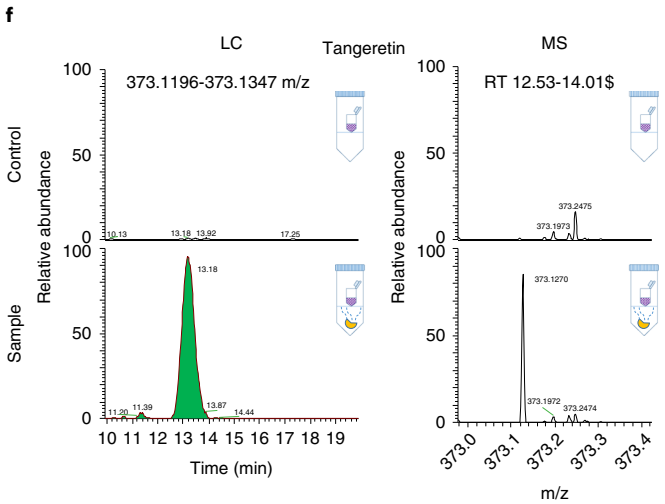

a
Fluorescence (483–533)
Cycles
10^3 E. coli cells
No treatment
NTC(no template control)
No treatment
10^3 E. coli cells
UV treatments
NTC(no template control)
UV treatments

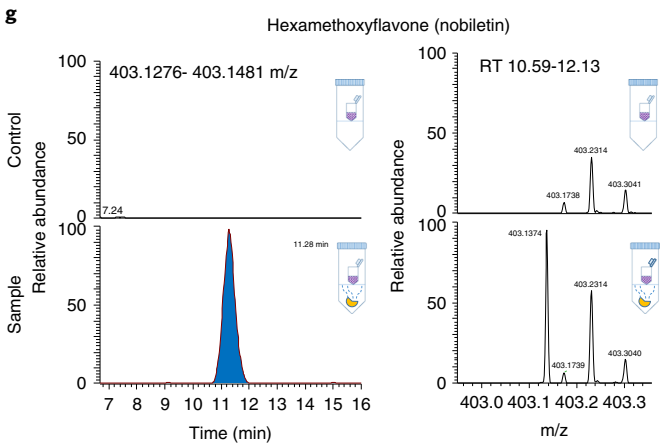

b
10E1
10E2
10E3
NTC
E. coli
Unclassified
Others


c
10E3 E. coli cells genomic coverage
Coverage
GC%
Position (bp)
Coverage
GC content (%)
Mean GC content


d
DNA
cDNA
(FU)
(bp)


e
Tangeretin
Hexamethoxyflavone (nobiletin)
Relative abundance
373.1270
403.1376
Control
Sample


f
LC
Tangeretin
MS
373.1196-373.1347 m/z
RT 12.53-14.01$
Control
Sample
Relative abundance
Time (min)
m/z


g
Hexamethoxyflavone (nobiletin)
403.1276- 403.1481 m/z
RT 10.59-12.13
Control
Sample
Relative abundance
Time (min)
m/z

◀ **Fig. 4 | The personal exposome collection, extraction and processing pipeline are highly sensitive. a**, The UV radiation pre-treatment of the extraction buffers is critical to detect low-abundance biologicals. Without UV pre-treatment, a sample of ~1,000 *Escherichia coli* cells have similar qPCR amplification curves when compared to the no template control (left) but are clearly detectable when extraction buffers were pre-treated with UV radiation (right). Samples were measured in triplicate. **b**, The pipeline can detect as few as ~10 *E. coli* cells in the sample. Different amounts of cells were directly spotted on the filter, dried overnight and processed using the described pipeline. Pie charts show the proportions of assembled bases that were assigned with *E. coli*, Unclassified and Others, respectively. Even with only ~10 *E. coli* cells, there is a clear signal of *E. coli* detection. **c**, The genomic coverage information of the ~1,000 *E. coli* cells sample as described in **b**, visualized by qualimap[43]. **d**, Typical size distributions of a DNA library (top) and a cDNA library (bottom). **e**–**g**, (**e**) To test the detection capacity of the chemical pipeline, a small slice of orange peel was placed inside a 50-ml conical tube in which the zeolite beads were also present (purple rounds). A control tube was set up the same way without the orange peel. Two representative chemicals related to the orange peel—tangeretin and nobiletin—were extracted from the beads in the experimental tube but not from the beads in the control tube ($n = 6$; the line stands for median). Both chemicals were validated independently with purchased standard compounds in **f** and **g**, respectively. Figure 4e adapted with permission from ref. [1].

air-dissolved compounds from particle-free airflow that was concomitantly generated with the collection of biotic samples. Zeolite is widely used as a molecular sieve to collect and remove chemicals. Of our other sample testing, zeolite was able to readily capture flavor compounds emanated by an orange peel placed in close vicinity (Fig. 4e–g). The caveat of the device is that the flow rate is around 5% of the actual human breathing rate; hence, the amount of materials collected is limited. As of now, more powerful aerosol samplers are available for exposome studies, such as the Ultrasonic Personal Aerosol Sampler[33].

Biotic exposome extraction

To allow simultaneous DNA and RNA extractions, we modified the original protocols of the Qiagen DNeasy and RNeasy PowerWater Kits. After removal of non-nucleic acids material, the samples were split into two equal parts for DNA and RNA extraction, respectively. Owing to the low amount of extracted materials, nucleic acids need to be linearly amplified before library preparation. We recommend the Qiagen REPLI-g Single Cell Kit for DNA amplification. For RNA, we use the isothermal amplification kit NuGen Ovation RNA-seq V2. Unlike most kits that selectively enrich messenger RNAs (mRNAs) with poly-A tails, the Ovation kit was designed for the amplification of all non-rRNAs. This is crucial to preserve the complex community structure of the biotic exposome, as bacterial and viral RNAs do not have poly-A tails and would not be amplified with conventional approaches.

Decontamination considerations for biotic extractions

A dedicated biosafety cabinet is highly recommended to reduce contamination, especially for nucleic acid extractions and handling before amplification. Before starting the experimental protocol, carefully wipe the workstation with 70% ethanol. If working with RNA, clean the surface with an RNase decontamination solution. Amplified products should be transferred to a physically separated post-amplification lab space to minimize nonspecific amplification products. Extraction buffers, filters, cartridges and reusable forceps should be placed within 3 cm of the ultraviolet (UV) light source and sterilized with a UV crosslinker each time before use (120 mJ/cm^2). The main purpose of UV radiation is to break all DNA fragments into small pieces that could not be amplified in the later steps. In addition, a negative control (i.e., a blank filter) should always be included for each batch of extractions to assess the level of contamination during experimental procedures.

Library preparation and sequencing

Various library preparation methods and sequencing platforms can be used. We recommend Illumina or BGI sequencing platforms for accuracy and high-throughput data output. Barcoded libraries are prepared using the Roche (originally developed by Kapa) HyperPlus kit and quality checked by Agilent 2100 Bioanalyzer. Multiple libraries can be pooled at equal molarity and sequenced with at least 30–50 million 2 × 150-base pair (bp) pair-ended reads/sample. Duplicated reads are removed from sequenced reads. Reads aligning to the human genome were removed and Blast searched against the pan-domain (or the National Center of Biotechnology Information nucleotide database: https://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/nt.gz) database for taxonomy. The resulting taxonomy classifications can be used for various downstream applications.

Abiotic exposome data quality assurance (QA) and quality control (QC)

To extract most bioactive chemicals out of the chemical cartridge and avoid precipitations during the extraction, we chose methanol as the extractant, which is similar to metabolomics protocols, although, with methanol extraction, certain nonpolar chemicals might be lost.

Also, to ensure LC–MS data quality, we recommend adding internal standards and surrogates depending on the chemicals of interest during the extraction to control for determining extraction efficiency, evaluate LC–MS performance and normalize the LC–MS data. As an example, deuterated polycyclic aromatic hydrocarbons (PAHs) and PAH analogues have been used as internal standards and surrogates, respectively, when determining the concentrations of PAHs in environmental samples[24].

We recommend the following procedures to be taken to ensure QA/QC of the untargeted analysis by LC–MS: (i) sample randomization for chemical extraction and data acquisition; (ii) six injections of a pooled sample, such as equimolar mix of all the samples, to equilibrate the LC–MS system before running the sequence; (iii) injection of a pooled sample every six injections to control for signal deviation with time; (iv) discarding features from solvent blanks; and (v) checking mass accuracy, retention time and peak shape of internal standards in every sample[34,35].

For targeted analysis, we follow the following steps to ensure QA/QC for LC–MS: (i) running the procedural, spike and matrix blanks and sample duplicates routinely to make sure that no interference is detected; (ii) running one laboratory blank and one replicate with every six samples, with samples having ≥ ±15% differences being reanalyzed; (iii) performing a system suitability test before each batch to ensure the acceptable limit of detection for each chemical class; (iv) sample randomization for chemical extraction and data acquisition; and (v) triplicate injection of extracted chemicals from the reference sample at the beginning of the batch[27,36].

Materials

Reagents

Reagents for processing biotics

- β-mercaptoethanol (β-ME) (e.g., Bio-Rad, cat. no. 1610710) **! CAUTION** β-ME is flammable, toxic and might irritate eyes and skin. Store in a fireproof cabinet, wear proper protection and handle with caution in a fume hood.
- Ethyl alcohol (200 proof) (e.g., Gold Shield Distributors) **! CAUTION** Ethyl alcohol is flammable. Store in a fireproof cabinet, handle with caution and dispose of properly.
- Nuclease-free water (e.g., Integrated DNA Technologies, cat. no. 11-05-01-04)
- DNeasy PowerWater Kit (Qiagen, cat. no. 14700-50NF) **▲ CRITICAL** We recommend the Qiagen DNeasy PowerWater Kit owing to its high extraction efficiency.
- RNeasy PowerWater Kit (Qiagen, cat. no. 14900-50NF) **▲ CRITICAL** We recommend the Qiagen RNeasy PowerWater Kit owing to its high extraction efficiency and compatibility for use in conjunction with the DNeasy PowerWater Kit.
- Qubit dsDNA HS (high sensitivity) assay kit (Thermo Fisher Scientific, cat. no. Q32851) **! CAUTION** The Qubit dsDNA HS reagent is a dye that binds to nucleic acids and is provided in a DMSO solution. It should be treated as a potential mutagen. Wear proper hand and eye protection and handle with care.
- Kapa HyperPlus Kit (Roche Sequencing and Life Science, cat. no. KK8514)
- REPLI-g Single Cell Kit (Qiagen, cat.no. 150345) **▲ CRITICAL** We recommend the REPLI-g Single Cell Kit because it uses decontaminated reagents and multiple displacement technology that allow for linear amplification of DNA.
- NuGen Ovation RNA-seq System V2 (Tecan Genomics, cat. no. 7102-32) **▲ CRITICAL** We recommend the NuGen Ovation RNA-seq System V2 because it broadly amplifies all non-rRNAs, including bacterial and viral RNAs. Other kits that use amplifying technologies targeting poly-A tails of mRNAs are not suitable as bacterial, and viral DNAs do not have poly-A tails.
- High Sensitivity DNA Kit (Agilent Technologies, cat. no. G2940CA) **! CAUTION** The dye binds to nucleic acids and should be treated as a potential mutagen. The kit also contains DMSO. Wear proper hand and eye protection and handle with care.
- Agencourt AMPure XP (Beckman Coulter, cat. no. A63881)
- 10 mM Tris-HCl elution buffer, pH 8.5 (e.g., Qiagen, cat. no. 19086)
- RNase decontamination solution (e.g., RNaseZap, Thermo Fisher Scientific, cat. no. AM9780) **! CAUTION** This reagent might irritate the skin, eyes and lungs. Wear proper protection and handle with care.

Reagents for processing abiotics

- LC-grade water (Thermo Fisher Scientific, cat. no. MWX00011)
- LC-grade methanol (Thermo Fisher Scientific, cat. no. AC610090040)
- LC-grade acetonitrile (Thermo Fisher Scientific, cat. no. AC615140025)
- LC-grade ammonium acetate (Thermo Fisher Scientific, cat. no. A11450)
- LC-grade acetic acid (Thermo Fisher Scientific, cat. no. A11350)
- LC-grade ammonium hydroxide (Thermo Fisher Scientific, cat. no. A470250)

Equipment

Common equipment

- 3D printing filament (Acrylonitrile Butadiene Styrene (ABS), Hatchbox, cat. no. 3D ABS-1KG1.75-WHT)
- DNA LoBind tubes (Eppendorf, cat. no. 022431021)
- 0.2-ml polymerase chain reaction (PCR) tubes (FrameStrip 8-well PCR Tube Strip, 4titude, cat. no. 4ti-0785/P)
- Falcon 50-ml conical centrifuge tube (Thermo Fisher Scientific, cat. no. 14-432-22)
- Pipettors (0.1–2, 2–20, 20–200 and 100–1,000 µl (Mettler Toledo Rainin Pipet-Lite LTS, cat. nos. 17014393, 17014392, 17014391 and 17014382)
- 1-L HPLC bottle (Waters)

Equipment for particles collection

- MicroPEM (RTI International)
- MicroPEM Docking Station (RTI International)
- Mass flow meter (TSI Incorporated, cat. no. 4140)
- 3D printer (FlashForge Creator Pro, FlashForge, cat.no. FF-CP3DP)
- Forceps (Thermo Fisher Scientific, cat. no. 09-753-50)
- Polyethersulfone (PES) filter, 25 mm (Sterlitech, cat. no. PES0825100)
- Polytetrafluoroethylene (PTFE) 25-mm Teflon filter (PALL Corporation, cat. no. R2P1025)

Equipment for processing biotics

- 2100 Bioanalyzer (Agilent Technologies, cat. no. G2940CA)
- Centrifuge for 15-ml tubes (able to reach at least 4,000*g*) (Thermo Fisher Scientific, cat. no. 75004537)
- Ultraviolet crosslinker (Spectrolinker UV Crosslinker, Spectronics, cat. no. XL-1500)
- Microcentrifuge (able to reach at least 13,000*g*) (Thermo Fisher Scientific, accuSpin Micro 17, cat. no. 13-100-676)
- Mini centrifuge (Santa Cruz Biotechnology, cat. no. C1008-G)
- Vortex mixer (able to reach at least 3,200 r.p.m) (Thermo Fisher Scientific, cat. no. 02215365)
- Vortex adapter for 5-ml tubes (Qiagen, cat. no. 13000-V1-5)
- DynaMag-96 side magnet (Thermo Fisher Scientific, cat. no. 12331D)
- Qubit 4 fluorometer (Thermo Fisher Scientific, cat. no. Q33226)
- Thermal cycler (Thermo Fisher Scientific, cat. no. A24811)

Equipment for processing abiotics

- Matrix Molecular Sieve 13×, 45–60 mesh (MilliporeSigma, cat. no. 20304)
- Deactivated clear glass insert, 150-µl volume (Waters, SKU: WAT094171DV)
- LC–MS Certified Amber Glass, 12 mm × 32 mm Screw Neck Max Recovery Vial, 2-ml volume (Waters, SKU: 600000754CV)
- Ultra-performance liquid chromatography (Waters)
- Exactive Orbitrap Mass Spectrometer (Thermo Fisher Scientific)
- OPD2 HP-4B LC column (Shodex, Showa Denko)
- OPD2 HPG-4A LC guard column (Shodex, Showa Denko)

Reagent setup

70% (vol/vol) Ethanol

Dilute absolute ethanol with nuclease-free water with a ratio of 70:30 (vol/vol). This solution can be stored at room temperature for up to 1 week.

80% (vol/vol) Ethanol
Dilute absolute ethanol with nuclease-free water with a ratio of 80:20 (vol/vol). Freshly prepare this solution every time before use.

HPLC solution A
Prepare 10 mM ammonium acetate in 50:50 (vol/vol) acetonitrile. Add water in a 1-L HPLC bottle and mix by gently shaking. Solution A can be stored for up to 2 weeks at room temperature (~20–25 °C).

HPLC solution B
Prepare 10 mM ammonium acetate in 90:10 (vol/vol) acetonitrile. Add water in a 1-L HPLC bottle and mix by gently shaking. Solution B can be stored for up to 2 weeks at room temperature (~20–25 °C).

pH modification for positive/negative mode acquisition
Modify HPLC solution A and HPLC solution B with 10 mM acetic acid (pH 4.75) for positive-mode acquisition or 10 mM ammonium hydroxide (pH 9.25) for negative-mode acquisition. Freshly prepare these solutions every time before use.

Equipment setup

Exposometer setup
To minimize environmental contamination, filters and cartridges should be deployed and recovered from MicroPEM in a sterilized biosafety cabinet (Fig. 2a–c). After each run, connect the MicroPEM to a computer (only Windows operating system is supported) and download the data from the Docking Station program. Within the Docking Station program, connect the MicroPEM with a mass flowmeter and adjust the flow rate to ±5% of the target flow rate (0.5 L/min) following instructions from the program. Change to new batteries or fully charged rechargeable batteries before the next deployment.

For biological substances, we used a 3.0-mm pore-size PTFE 25-mm Teflon filter or a 0.8-mm pore-size PES 25-mm filter in the MicroPEM filter cassette to collect aerosol particulates for biotics extraction (Fig. 2d). The original oiled PM2.5 impactors were removed to allow maximum collection.

For organic chemicals, we designed a 3D-printed cartridge (Figs. 2e,f and 3 and Supplementary Data 1) to fill with 200 mg of Matrix Molecular Sieve 13× (zeolite adsorbent) and attached the cartridge to the end of the airflow in the exposometer to collect abiotic chemical air solvent. (Supplementary Data 1 provides an .stl file that can be used for 3D printing using FlashForge Creator Pro. The file may be adjusted if using other 3D printers for the optimal fit of parts.)

LC–MS
Perform the LC–MS analysis in a platform that consists of UPLC coupled to Exactive Orbitrap Mass Spectrometer, using a mix-mode OPD2 HP-4B column (4.6 mm × 50 mm) with a 4.6-mm × 10-mm guard column. Maintain the column temperature and the sample chamber temperature at 45 °C and 4 °C, respectively. Set the flow as follows: flow rate, 0.1 ml/min; gradient, 0–15 min, 99% A, 15–18 minutes, 99% to 1% A; 18–24 min, 1% A; 24–25 min, 1% to 99% A; 25–30 min, 99% A. Set the MS acquisition in profile mode and perform with an electrospray ionization probe, operating with the capillary temperature at 275 °C, sheath gas at 40 units, spray voltage at 3.5 kV for positive mode or 3.1 kV for negative mode, capillary voltage at 30 V, tube lens voltage at 120 V and skimmer voltage at 20 V. Use 100,000 mass resolution for the mass scanning, high dynamic range for AGC target, 500 ms as maximum inject time and 70–1,000 *m/z* as the scan range.

Procedure

Airborne exposure collection ● Timing 1–5 d

1 Either place an exposometer of choice at a location near the individual (e.g., the device may be placed on the table when the individual is working on the computer or clipped to the outside of a backpack when commuting) or ensure that the individual wears the device to collect the personal airborne exposome for 2–5 d. Although the MicroPEM device was used in the previous study[1], the selection of device could be the operator's choice as long as the biological collection filter and the chemical collection cartridge could fit inside the device.

Co-extraction of DNA and RNA from the collection filter ● Timing 4 h

▲ CRITICAL Filters capturing the biotic samples are used for simultaneous DNA and RNA extraction by a modified protocol combining the Qiagen DNeasy PowerWater Kit and the RNeasy PowerWater Kit.

▲ CRITICAL Before starting, use an RNase decontamination solution to clean the processing surface in the biosafety cabinet.

▲ CRITICAL The solutions used in the following steps are provided with the DNeasy and RNeasy kits.

▲ CRITICAL Unless otherwise specified, centrifugation steps are carried out at room temperature.

Before starting

2 Warm PM1 at 55 °C for 10 min. Aliquot 990 µl of PM1 (mix well and make sure that there is no precipitate in the tube) in 1.5-ml tubes and treat the PM1 buffer with UV radiation (254 nm, 120 mJ/cm^2) for 1,800 s (within 3 cm of the UV source).
▲ CRITICAL STEP The high-dosage UV treatment to remove residual DNA/RNA contamination in the buffer(s) is essential to achieve high sensitivity of nucleic acid detection. For different reagent kits, it is strongly recommended that the user checks the DNA/RNA contamination in the buffers with quantitative PCR (qPCR) before the extraction steps.

3 Prepare DNase I stock by adding 550 µl of RNase-free water to the DNase I lyophilized powder. Mix gently and aliquot in 45-µl portions. Store at −20 °C for long-term storage.
▲ CRITICAL STEP Do not freeze or thaw stock enzymes more than three times.

4 Prepare the DNase I solution by thawing the volume of DNase I stock enzyme needed according to the number of samples. Per prep, combine 5 µl of DNase I stock enzyme with 45 µl of DNase Digestion Solution. For example, add 360 µl of DNase Digestion Solution to 40 µl of DNase I stock enzyme for eight samples.

5 Select sample filters and one blank filter as a negative control. Insert the filter into the 5-ml PowerWater Bead Tube provided with the kit.

6 Add 10 µl of β-ME to every 990 µl of PM1. Add 1 ml of Solution PM1 containing β-ME to the PowerWater Bead Tube.
▲ CRITICAL STEP PM1 must be warmed to 55 °C right before use and used while still warm. Precipitates might be present when PM1 is at room temperature.

7 Make sure that the PowerWater Bead Tube cap is securely tightened. Secure the tubes horizontally, with caps pointing toward the center, on the vortex adapter. Vortex at maximum speed (at least 3,200 r.p.m.) for 5 min.

8 Centrifuge the tubes at ≤4,000*g* for 1 min at room temperature (20 °C). Transfer the supernatant to a clean 2-ml collection tube (provided with the kit). Draw up the supernatant using a 200-µl pipette tip by placing it down into the beads.
▲ CRITICAL STEP Expect to recover 600–650 µl of supernatant; some beads might get carried over.

9 Centrifuge at 13,000*g* for 1 min. Avoiding the pellet, transfer the supernatant to a clean 2-ml collection tube.

10 Add 200 µl of Solution IRS to the supernatant in the collection tube and vortex briefly to mix. Incubate at 4 °C for 5 min.

11 Centrifuge the tubes at 13,000*g* for 1 min. Avoiding the pellet, transfer the supernatant to two clean 2-ml collection tubes (for DNA and RNA extraction, respectively). Expect to recover 800–850 µl of supernatant (400–425 µl each for DNA or RNA extraction). Proceed immediately to the next step.

DNA extraction

12 Add 325 µl of Solution PW3 to one of the tubes from Step 11 and inverse-mix by hand briefly.
▲ CRITICAL STEP Do not vortex at this step.

13 Load 500 µl of supernatant onto an MB Spin Column and centrifuge at 13,000*g* for 1 min. Discard the flow-through and repeat until all supernatant has been loaded onto the DNA Spin Filter.

14 Place the MB Spin Column into a clean 2-ml collection tube. Shake to mix Solution PW4 before use. Add 600 µl of Solution PW4 and centrifuge at 13,000*g* for 1 min.

15 Discard the flow-through and add 600 µl of ethanol (provided by the kit) and centrifuge at 13,000*g* for 1 min.

16 Discard the flow-through and centrifuge again at 13,000*g* for 2 min to remove residual wash.

17 Place the MB Spin Column into a clean 2-ml collection tube. Add 52 µl of Solution EB to the center of the white filter membrane. Make sure that the drop of solution is thoroughly absorbed into the filter membrane. Incubate for 2 min at room temperature.
▲ **CRITICAL STEP** The EB buffer can be pre-warmed at 50 °C to increase the yield.

18 Centrifuge at 13,000*g* for 1 min. Discard the MB Spin Column. Aliquot DNA in three DNA/RNA LoBind tubes (15 µl each) to avoid freeze/thaw cycles. Seal the tubes with Parafilm. The DNA is now ready for downstream applications.
■ **PAUSE POINT** Purified DNA can be stored at −80 °C for at least several months.

RNA extraction

19 Add 325 µl of Solution PM3 and 325 µl of Solution PM4 to the other tube from Step 11. Then, inverse-mix by hand briefly. Do not vortex.
▲ **CRITICAL STEP** PM3 and PM4 are binding buffers optimized for RNA.

20 Load 600 µl of supernatant onto an MB RNA Spin Column and centrifuge at 13,000*g* for 1 min. Discard the flow-through and repeat until all the supernatant has been loaded onto the RNA Spin Column.

21 Shake to mix Solution PM5. Add 600 µl of Solution PM5 and centrifuge at 13,000*g* for 1 min. Discard the flow-through.

22 Centrifuge again at 13,000*g* for 1 min to remove residual wash. Place the Spin Filter basket into a clean 2-ml collection tube.

23 Add 50 µl of DNase I Solution (prepared by mixing 45 µl of DNase Digestion Solution and 5 µl of DNase I stock solution) to the center of the MB RNA Spin Column. Incubate at room temperature for 15 min.

24 Add 400 µl of Solution PM7 and centrifuge the column at 13,000*g* for 1 min.

25 Discard the flow-through and add 600 µl of Solution PM5 and centrifuge at 13,000*g* for 1 min.

26 Discard the flow-through and add 600 µl of Solution PM4 and centrifuge at 13,000*g* for 1 min. Discard the flow-through and centrifuge again at 13,000*g* for 2 min to remove residual wash.

27 Place the MB RNA Spin Filter basket into a clean 2-ml collection tube. Add 52 µl of RNase-free water (provided) to the center of the white filter membrane. Incubate for 2 min at room temperature.

28 Centrifuge at 13,000*g* for 1 min. Discard the RNA Spin Filter basket. Aliquot RNA in three DNA/RNA LoBind tubes (15 µl each) to avoid freeze/thaw cycles. Seal the tubes with Parafilm. The RNA is now ready for downstream applications.
■ **PAUSE POINT** Purified RNA can be stored at −80 °C for several months.

DNA amplification and purification ● Timing 10 h (overnight amplification)

▲ **CRITICAL** DNA samples are linearly amplified by the Qiagen REPLI-g Single Cell MDA Amplification Kit with modifications.

▲ **CRITICAL** Perform DNA amplification steps (29–33) in a pre-amplification space (physically separated from the post-amplification space). The purification steps (34–39) should be carried in the post-amplification space to minimize nonspecific amplification products.

▲ **CRITICAL** Solutions used in the following step are provided with the Qiagen REPLI-g kit.

▲ **CRITICAL** Before starting: prepare buffers D1 and N1.

29 D1: Add 7 µl of Reconstituted Buffer DLB to 25 µl of nuclease-free water.

30 N1: Add 9 µl of Stop Solution to 51 µl of nuclease-free water.
▲ **CRITICAL STEP** Volumes given are sufficient for 12 reactions.
■ **PAUSE POINT** Buffers D1 and N1 can be stored at −20 °C for 3 months.

31 Add 5 µl of buffer D1 to 5 µl of DNA sample from Step 18. Mix by gently vortexing and centrifuge briefly. Incubate at room temperature for 3 min.

32 Immediately add 10 µl of stop buffer N1. Mix by gently vortexing and centrifuge briefly.

33 For each reaction of 20 µl of denatured DNA, add 29 µl of REPLI-g sc Reaction Buffer and 2 µl of REPLI-g sc polymerase. Incubate at 30 °C for 8 h.
▲ **CRITICAL STEP** Move the amplified reactions to a physically separated post-amplification lab space before doing the following steps.

34 Clean amplified products using 1.8× volumes of the Agencourt AMPure XP beads to remove reagents from the amplification step. Combine the following:

Component	Volume (µl)
Amplified products	51
Agencourt AMPure XP beads	91.8
Total	**141.8**

35 Mix thoroughly by pipetting up and down ten times. Incubate at room temperature for 10 min to bind DNA to the beads.

36 Place the tubes on a magnet to capture the beads. Incubate for 5 min until the liquid is clear. Carefully remove and discard the supernatant.

37 Keep the tubes on the magnet and add 200 µl of freshly made 80% (vol/vol) ethanol. Incubate the tubes at room temperature for ≥30 s to remove the ethanol. Repeat the rinsing step and try to remove all residual ethanol without disturbing the beads. Dry the beads at room temperature for 5 min.
▲ **CRITICAL STEP** Make a fresh batch of 80% (vol/vol) ethanol every time. The concentration of ethanol is important to ensure the best beads performance.

38 Remove the tubes from the magnet and resuspend the beads in 55 µl of 10 mM Tris-HCl elution buffer. Incubate the tubes at room temperature for 2 min to elute DNA off the beads.

39 Place the tubes on a magnet to capture the beads and incubate until the liquid is clear. Transfer 50 µl of cleared supernatant to new tubes. The final product is ready for library preparation.
■ **PAUSE POINT** Purified and amplified DNA can be stored at −80 °C for at least several months.
? TROUBLESHOOTING

Complementary DNA conversion and amplification ● Timing 6 h

▲ **CRITICAL** RNA samples are linearly amplified by Ovation RNA-seq system V2 (Tecan Genomics) with modifications.

▲ **CRITICAL** Perform steps from first-strand complementary DNA (cDNA) synthesis through Single Primer Isothermal Amplification (SPIA) amplification steps (40–56) in a pre-amplification space (physically separated from the post-amplification space). The remaining steps should be carried in the post-amplification space to minimize nonspecific amplification products.

First-strand cDNA synthesis

40 Before starting: pre-warm the thermal cycler at 65 °C.

41 Add 5 µl of RNA sample from Step 28 to 2 µl of A1 (first-strand primer mix) in a 0.2-ml PCR tube. Place the tubes in a pre-warmed thermal cycler programmed as follows:

Temperature	Time
65 °C	2 min
4 °C	∞
Lid temperature	100 °C

▲ **CRITICAL STEP** 500 pg to 100 ng of RNA sample could be used for the amplification step.

42 Prepare the first-strand master mix by combining 2.5 µl of A2 (first-strand buffer mix) and 0.5 µl of A3 (first-strand enzyme mix). Add 3 µl of the first-strand master mix to each tube containing RNA and A1.

Component	Volume (µl)
RNA and A1 (Step 41)	7
A2	2.5
A3	0.5
Total	**10**

43 Mix by pipetting five times, spin and place on ice. Cool the thermal cycler to 4 °C. Place the tubes in a pre-cooled thermal cycler programmed as follows:

Temperature	Time
4 °C	1 min
25 °C	10 min
42 °C	10 min
70 °C	15 min
4 °C	∞
Lid temperature	100 °C

44 Remove the tubes, spin to collect condensation and place on ice.

Second-strand cDNA synthesis

45 Before starting: pre-cool the thermal cycler to 4 °C.

46 Make a master mix by combining 9.7 µl of B1 (second-strand buffer mix) and 0.3 µl of B2 (second-strand enzyme mix). Add 10 µl of the second-strand master mix to each first-strand reaction tube. Mix by pipetting five times, spin and place on ice.

Component	Volume (µl)
cDNA products (Step 44)	10
B1	9.7
B2	0.3
Total	**20**

47 Place the tubes in a pre-cooled thermal cycler programmed as follows:

Temperature	Time
4 °C	1 min
25 °C	10 min
50 °C	30 min
80 °C	20 min
4 °C	∞
Lid temperature	100 °C

48 Remove the tubes and spin to collect condensation and place on ice.

cDNA purification

49 Purify double-strand cDNA with 1.4× of the Agencourt RNAClean XP beads. Add beads to cDNA products as outlined below:

Component	Volume (µl)
cDNA products (Step 48)	20
Agencourt RNAClean XP beads	28
Total	**48**

50 Follow cleaning Steps 35–37. DO NOT elute.
▲ **CRITICAL STEP** Dried beads are needed for the following steps.

Amplification of purified cDNA with SPIA

51 Before starting: pre-cool the thermal cycler to 4 °C.

52 Make a master mix by sequentially combining 20 µl of C2 (SPIA buffer mix), 10 µl of C1 (SPIA primer mix) and 10 µl of C3 (SPIA enzyme mix).

Component	Volume (µl)
C2	20
C1	10
C3	10
Total	**40**

▲ **CRITICAL STEP** Make sure that C3 is added at the last moment.

53 Add 40 µl of the SPIA master mix to each tube containing the double-strand cDNA bound to the dried beads (Step 50). Use a pipette set to 30 µl and mix thoroughly at least 8–10 times.

54 Place the tubes in a pre-cooled thermal cycler programmed as follows:

Temperature	Time
4 °C	1 min
47 °C	120 min
80 °C	20 min
4 °C	∞
Lid temperature	100 °C

▲ **CRITICAL STEP** Incubation time at 47 °C can vary for 60, 120 or 180 min depending on the amount of target amplified cDNA, the longer the incubation time and the more cDNA amplified.

55 Remove tubes from the thermal cycler, spin to collect condensation and place on ice.
▲ **CRITICAL STEP** DO NOT re-open in the pre-amplification space.

56 Transfer the tubes to a post-amplification space, put on the magnet and let stand for 5 min to clear the solution of beads completely. Carefully transfer 40 µl of the cleared supernatant containing the amplified cDNA to a fresh tube.

Purification of amplified cDNA

57 Purify SPIA amplified cDNA with 0.8 volumes of AMPure XP beads as outlined below:

Component	Volume (µl)
cDNA products	40
Agencourt XP beads	32
Total	**72**

▲ **CRITICAL STEP** The volume of AMPure XP beads might vary depending on the desired cDNA sizes for different downstream applications.

58 Continue beads cleanup following Steps 35–39.
■ **PAUSE POINT** Purified and amplified cDNA can be stored at −80°C for at least several months.
? TROUBLESHOOTING

Library preparation ● Timing DNA: 6 h; cDNA: 6 h

▲ **CRITICAL** Sequencing libraries are prepared with Kapa HyperPlus Kits (Roche Sequencing and Life Science) according to manufacturer instructions with modifications. The protocol for DNA and cDNA library preparation is very similar unless otherwise specified

Enzymatic fragmentation

59 Before starting: pre-cool the thermal cycler to 4 °C.

60 For DNA, follow Option A; for cDNA, follow Option B.

(A) **DNA**

(i) Dilute 500 ng of input DNA (from Step 39) with 10 mM Tris-HCl (pH 8.0–8.5) to a total of 35 µl.

(B) **cDNA**

(i) Dilute 1–1,000 ng of input cDNA (from Step 58) with 10 mM Tris-HCl (pH 8.0–8.5) to a total of 35 µl.

61 Assemble each fragmentation reaction on ice by adding the rest of the components in the order shown below:

Component	Volume (µl)
Diluted DNA	35
Frag buffer (10×)	5
Frag enzyme	10
Total	**50**

62 Vortex gently and spin down briefly. Return the tubes to ice. Proceed immediately to the next step.

63 For DNA, follow Option A; for cDNA follow Option B.

(A) **DNA**

(i) Incubate in the pre-cooled thermal cycler programmed as below:

Temperature	Time
4 °C (pre-cool)	∞
37 °C	8 min
4 °C	∞
Lid temperature	≤50 °C

(B) **cDNA**

(i) Incubate in the pre-cooled thermal cycler programmed as below:

Temperature	Time
4 °C (pre-cool)	∞
25 °C	5 min
4 °C	∞
Lid temperature	≤50 °C

64 Transfer reactions to ice and proceed immediately to the next step.

End repair and A-tailing

65 In the same tubes, assemble each end repair and A-tailing reaction as follows:

Component	Volume (µl)
Fragmented DNA/cDNA	50
End repair and A-tailing buffer	7
End repair and A-tailing enzyme mix	3
Total	**60**

66 Vortex gently and spin down briefly.

67 Incubate the tubes in a thermal cycler programmed as follows:

Temperature	Time
65 °C	30 min
4 °C	∞
Lid temperature	85 °C

Adapter ligation

68 In the same tubes, assemble each adapter ligation reaction as follows:

Component	Volume (µl)
End repair and A-tailing reaction product	60
Adapter stock, 15 µM	5
PCR-grade water	5
Ligation buffer	30
DNA ligase	10
Total	**110**

69 Mix thoroughly, centrifuge briefly and incubate in a thermal cycler programmed as follows:

Temperature	Time
20 °C	30 min
4 °C	∞
Lid temperature	105 °C

Post-ligation cleanup

70 In the same tubes, perform a 0.8× beads cleanup by combining the following:

Component	Volume (µl)
DNA products	110
Agencourt XP beads	88
Total	**198**

71 For DNA, follow Option A; for cDNA, follow Option B.

(A) **DNA**

(i) Follow Steps 35–39 to purify products, and then proceed to Step 72.

(B) **cDNA**

(i) Follow Steps 35–37 to purify products. Then, remove the tubes from the magnet and resuspend the beads in 25 µl of 10 mM Tris-HCl elution buffer. Incubate the tubes at room temperature for 2 min to elute DNA off the beads. Place the tubes on a magnet to capture the beads and incubate until the liquid is clear. Transfer 20 µl of cleared supernatant to new tubes and proceed to Step 76.

DNA: double-size selection (target range, 300–450 bp)

72 Perform the first 0.6× size selection to the samples by adding 30 µl of AMPure XP reagent to the eluted 50-µl DNA library.

73 Place the tubes on a magnet to capture the beads. Incubate for 5 min until the liquid is clear. Transfer 75 µl of supernatant containing library molecules smaller than ~450 bp to new tubes.
▲ **CRITICAL STEP** No beads should be transferred with the supernatant.

74 Perform the second 0.7× size selection to the tubes by adding 5 µl of AMPure XP reagent.

75 Follow Steps 35–37 to purify products. Resuspend the beads in 25 µl of 10 mM Tris-HCl elution buffer. Incubate the tubes at room temperature for 2 min to elute DNA off the beads. Place the tubes on a magnet to capture the beads and incubate until the liquid is clear. Transfer 20 µl of cleared supernatant to new tubes for PCR amplification.

Library amplification

76 Assemble each library amplification reaction as follows:

Component	Volume (µl)
Adapter ligated library (Step 71B(i) or Step 75)	20
2× Kapa HiFi HotStart ReadyMix	25
10× Kapa library amplification primer mix	5
Total	**50**

77 Mix thoroughly and centrifuge briefly.

78 Amplify using the following cycling protocol:

Temperature	Time	Cycles
Initial denaturation 98 °C	45 s	1
Denaturation 98 °C	15 a	DNA: 3–4; cDNA: 5–9
Annealing 60 °C	30 s	
Extension 72 °C	30 s	
Final extension 72 °C	1 min	1
Hold 4 °C	∞	1
Lid temperature	105 °C	

Post-amplification cleanup

79 For DNA, follow Option A; for cDNA, follow Option B.

(A) **DNA**

(i) In each library amplification tube, perform a 0.7× beads cleanup twice by first combining the following:

Component	Volume (µl)
DNA products	50
Agencourt XP beads	35
Total	**85**

(B) **cDNA**

(i) In each library amplification tube, perform beads cleanup steps twice by first combining the following:

Component	1st cleanup volume (µl)	2nd cleanup volume (µl)
cDNA products	50	50
Agencourt XP beads	40	35
Total	**90**	**85**

80 For the first cleanup, follow Steps 35–39 to purify products.

81 For the second cleanup, follow Steps 35–37 to purify products, and then remove the tubes from the magnet and resuspend the beads in 25 µl of 10 mM Tris-HCl elution buffer. Incubate the tubes at room temperature for 2 min to elute DNA off the beads. Place the tubes on a magnet to

capture the beads and incubate until the liquid is clear. Transfer 20 µl of cleared supernatant to new tubes.

82 DNA/cDNA libraries can be quantitated by HS Qubit, and size distribution analysis of DNA/cDNA libraries can be performed by the Agilent 2100 Bioanalyzer. The expected concentrations for libraries should be above 1 ng /µl.
? TROUBLESHOOTING

83 Transfer the reactions to ice and proceed to the next step.
PAUSE POINT DNA and cDNA libraries can be stored at −80°C for at least several months.

Sequencing and analysis ● Timing 5–14 d

84 Choose Illumina or similar second-generation sequencing platforms to generate paired-end reads data, preferably at least 6 GB of data per sample owing to the diversity within the samples.

85 Apply standard metagenomic analytical pipelines, such as Kraken 2 (ref. [37]), Metaphlan2 (ref. [38]) and HUMAaN2 (ref. [39]), to the data if the focus is on the microbiome. Otherwise, customized databases need to be built to include other domains of life before analysis.

Chemical exposome processing pipeline

▲ CRITICAL The chemical exposome processing pipeline can be performed independently from its biological counterpart.

▲ CRITICAL The extractant, LC, LC columns, LC solutions, MS and systems settings are not limited to the models or parameters mentioned in the following steps. Any high-performance spectrometry coupled with high-resolution MS for determining the organic chemicals in exposome samples would work. Also note that no single platform can determine all the presented chemicals in exposome samples.

Organic chemicals collection cartridge extraction ● Timing 1 h

86 Transfer the adsorbent zeolite beads (~200 mg) from the cartridge to a clean LoBind tube.

87 Add 1 ml of methanol (LC-grade) as extractant, which can extract most bioactive chemicals and their bioavailable fractions.

88 Incubate the mixture for 20 min at room temperature.

89 Centrifuge the mixture at 22,000*g* for 20 min at room temperature.

90 Transfer the supernatant to a 150-µl deactivated glass insert housed in 2-ml Amber Glass LC–MS vials for LC–MS analysis, or store it at −20°C for later use.

LC–MS analysis ● Timing 0.5 h for each sample

91 Perform the LC–MS analysis in a platform that consists of a Waters ultra-performance liquid chromatography-coupled Thermo Exactive Orbitrap Mass Spectrometer, using a mix-mode LC column (4.6 mm × 50 mm) with a guard column (4.6 mm ×10 mm).

92 Maintain the column temperature at 45 °C and the sample chamber at 4 °C.

93 Apply HPLC solution A (10 mM ammonium acetate in 50%:50% (vol/vol) acetonitrile:water) and HPLC solution B (10 mM ammonium acetate in 90%:10% (vol/vol) acetonitrile:water) as the binary mobile phase solvents.

94 Modify both solvents with 10 mM acetic acid (pH 4.75) for positive-mode acquisition or 10 mM ammonium hydroxide (pH 9.25) for negative-mode acquisition.

95 Set the flow as follows: flow rate, 0.1 ml/min; gradient, 0–15 min, 99% A, 15–18 min, 99% to 1% A; 18–24 min, 1% A; 24–25 min, 1% to 99% A; 25–30 min, 99% A.

96 Apply an electrospray ionization probe and set the MS acquisition in profile mode.

97 Set the capillary temperature at 275 °C, sheath gas at 40 units, spray voltage at 3.5 kV for positive mode or 3.1 kV for negative mode, capillary voltage at 30 V, tube lens voltage at 120 V and skimmer voltage at 20 V.

98 Set the mass scanning at 100,000 mass resolution, maximum dynamic range, and high automatic gain control target, 500 ms as the maximum injection time and 70–1,000 *m/z* ratio as the scan range.

99 Run the prepared samples in the assigned order.

Troubleshooting

Troubleshooting advice can be found in Table 1.

Table 1 | Troubleshooting table

Step	Problem	Possible reason	Solution
39	Low DNA yield after amplification	Possible inhibitor in the DNA samples	Repeat cleanup the DNA and re-amplify
		Carryover of alcohol from extraction steps	Make sure to perform Step 16 carefully to remove residual buffer
58	Low cDNA yield	RNA was degraded	Store RNA properly and clean workstation with RNase decontamination reagent
		Insufficient amplification	Adjust amplification time in Step 54
82	Presence of adapter dimers	Libraries are not properly cleaned by beads	Perform beads cleanup again (0.8×) to remove adapter dimers
82	Presence of PCR artifacts	Too much input or too many PCR cycles	Reduce input or PCR cycles
82	Presence of beads contamination	Beads were carried over to the final product	Repeat beads cleanup and pipette carefully
82	Wrong size of DNA library	Sample lost during size selection	Use best practice for beads handling during beads cleanup

Timing

Sample collection:
Step 1, days 1–5, airborne exposure collection: 1–5 d
Biological analysis:
Steps 2–28, day 1, co-extraction of DNA and RNA from the collection filter: 4 h
Steps 29–39, day 2, DNA amplification and purification: 10 h
Steps 40–58, day 3, cDNA conversion and amplification: 6 h
Steps 59–83, day 4, DNA and cDNA library preparation: 12 h
Steps 84 and 85, days 5-18, sequencing and analysis: 5–14 d
Chemical analysis:
Steps 86–90, day 1, organic chemicals collection cartridge extraction: 1 h
Steps 91–99, day 1, LC–MS analysis: 0.5 h for each sample

Anticipated results

Biological features detected in the personal exposome

Annotations of the dataset could be performed based on the user's preference, as the extraction methods were intended for capturing as diverse organisms as possible. Typically, the amount of collected nucleic acid using the MicroPEM device is below 1 ng/μl; hence, amplifications are mandatory before library preparations. Typically, the amplifications could achieve as high as 1 μg/μl for DNA and in the range of 1–50 ng/μl for cDNA. The prepared sequencing libraries should be above 1 ng/μl, and the size distributions of the libraries should have peaks above 350 bp for effective sequencing (Fig. 4d). For analysis, depending on the amount of collected materials, a significant amount of human reads could be detected in the sequencing results, potentially due to either environmental contamination during extraction or actual biotic exposures during sample collection. Therefore, we recommend removing reads mapped to the human reference genome before metagenomic taxonomy classifications. The metagenomic reads can then be processed using conventional bioinformatic analytic approaches, such as Kraken 2 (ref. [37]), Metaphlan2 (ref. [38]) and HUMAaN2 (ref. [38]), with modifications of the databases to include different domains of interests. A comprehensive pan-domain database is crucial for decoding the diversity of biotic exposome. We recommend using the National Center of Biotechnology Information nucleotide database or building a customized database including at least plants, fungi, bacteria, viruses and invertebrates. Hundreds to

thousands of species could be detected in the airborne particulate samples, depending on the sampled locations, seasons and duration[1].

Chemical features detected in the personal exposome

More than thousands of chemical features can be detected in a given sample, and some of the features could be annotated when querying against chemical databases. Some examples of the databases for chemical exposome annotation are the Human Metabolome Database[40], the Blood Exposome Database[41] and the Toxin and Toxin Target Database[42]. The density plot of *m/z* ratios for features can be displayed in the positive mode or the negative mode. Features detected in the negative mode have much smaller *m/z* ratios and unimodal distributions, whereas features detected in the positive mode are more uniformly distributed across the spectrum. Max feature abundance should be detected in the positive mode and the negative mode with different HPLC columns, plotted against background abundance in the blank control. Most features should have low abundance in samples and blanks[1]. Feature processing can follow the protocols of metabolomics, including retention time alignment, matched candidates filtration, peak detection and extraction and peak matching. Many chemical databases that are suitable for chemical exposome annotation have been established for chemical annotation at different levels. Purchasing corresponding chemical standards is an optimal solution to identify and quantify each chemical of interest, and at least level 5 annotation is required.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The biotic exposome data generated using the protocol in the relevant study[1] were deposited to the National Center of Biotechnology Information under bioproject ID PRJNA421162.

References

1. Jiang, C. et al. Dynamic human environmental exposome revealed by longitudinal personal monitoring. *Cell* **175**, 277–291 (2018).
2. Huang, R. J. et al. High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **514**, 218–222 (2015).
3. Cohen, S., Janicki-Deverts, D. & Miller, G. E. Psychological stress and disease. *JAMA* **298**,1685–1687 (2007).
4. Vermeulen, R., Schymanski, E. L., Barabási, A. L. & Miller, G. W. The exposome and health: where chemistry meets biology. *Science* **367**, 392–396 (2020).
5. Pfeifer, G. P. Environmental exposures and mutational patterns of cancer genomes. *Genome Med.* **2**, 54 (2010).
6. Heederik, D. & Von Mutius, E. Does diversity of environmental microbial exposure matter for the occurrence of allergy and asthma? *J. Allergy Clin. Immunol.* **130**, 44–50 (2012).
7. Miller, F. W. et al. Epidemiology of environmental exposures and human autoimmune diseases: findings from a National Institute of Environmental Health Sciences Expert Panel Workshop. *J. Autoimmun.* **39**, 259–271 (2012).
8. Fujimura, K. E. et al. House dust exposure mediates gut microbiome Lactobacillus enrichment and airway immune defense against allergens and virus infection. *Proc. Natl Acad. Sci. USA* **111**, 805–810 (2014).
9. Wild, C. P. Complementing the genome with an 'exposome': the outstanding challenge of environmental exposure measurement in molecular epidemiology. *Cancer Epidemiol. Biomarkers Prev.* **14**, 1847–1850 (2005).
10. Wild, C. P. The exposome: from concept to utility. *Int. J. Epidemiol.* **41**, 24–32 (2012).
11. Lindholm, M. E. et al. An integrative analysis reveals coordinated reprogramming of the epigenome and the transcriptome in human skeletal muscle after training. *Epigenetics* **9**, 1557–1569 (2014).
12. Laker, R. C. et al. Transcriptomic and epigenetic responses to short-term nutrient-exercise stress in humans. *Sci. Rep.* **7**, 15134 (2017).
13. Cao, C. et al. Inhalable microorganisms in Beijing's PM2.5 and PM10 pollutants during a severe smog event. *Environ. Sci. Technol.* **48**, 1499–1507 (2014).
14. Lax, S. et al. Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science* **345**, 1048–1052 (2014).
15. Dannemiller, K. C. et al. Next-generation DNA sequencing reveals that low fungal diversity in house dust is associated with childhood asthma development. *Indoor Air* **24**, 236–247 (2014).
16. Barberán, A. et al. Continental-scale distributions of dust-associated bacteria and fungi. *Proc. Natl Acad. Sci. USA* **112**, 5756–5761 (2015).

17. Abbatt, J. P. D. & Wang, C. The atmospheric chemistry of indoor environments. *Environ. Sci. Process. Impacts* **22**, 25–48 (2020).
18. Oliveira, M., Slezakova, K., Delerue-Matos, C., Pereira, M. C. & Morais, S. Children environmental exposure to particulate matter and polycyclic aromatic hydrocarbons and biomonitoring in school environments: a review on indoor and outdoor exposure levels, major sources and health impacts. *Environ Int*. **124**, 180–204 (2019).
19. Cissé, O. H., Ma, L., Jiang, C., Snyder, M. & Kovacs, J. A. Humans are selectively exposed to *Pneumocystis jirovecii*. *mBio* **11**, e03138-19 (2020).
20. Hänninen, O. O. et al. The EXPOLIS study: implications for exposure research and environmental policy in Europe. *J. Expo. Anal. Environ. Epidemiol.* **14**, 440–456 (2004).
21. Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. & Pozzer, A. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **525**, 367–371 (2015).
22. Behzad, H., Gojobori, T. & Mineta, K. Challenges and opportunities of airborne metagenomics. *Genome Biol. Evol.* **7**, 1216–1226 (2015).
23. Jiang, W. et al. Optimized DNA extraction and metagenomic sequencing of airborne microbial communities. *Nat. Protoc*. **10**, 768–779 (2015).
24. Gao, P., da Silva, E. B., Townsend, T., Liu, X. & Ma, L. Q. Emerging PAHs in urban soils: concentrations, bioaccessibility, and spatial distribution. *Sci. Total Environ*. **670**, 800–805 (2019).
25. Xiang, P. et al. Cellular responses of normal (HL-7702) and cancerous (HepG2) hepatic cells to dust extract exposure. *Chemosphere* **193**, 1189–1197 (2018).
26. Leung, M. H. Y., Wilkins, D., Li, E. K. T., Kong, F. K. F. & Lee, P. K. H. Indoor-air microbiome in an urban subway network: diversity and dynamics. *Appl. Environ. Microbiol.* **80**, 6760–6770 (2014).
27. Hospodsky, D. et al. Characterizing airborne fungal and bacterial concentrations and emission rates in six occupied children's classrooms. *Indoor Air* **25**, 641–652 (2015).
28. Salter, S. J. et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **12**, 87 (2014).
29. Weiss, S. et al. Tracking down the sources of experimental contamination in microbiome studies. *Genome Biol.* **15**, 564 (2014).
30. Cui, Y. et al. The exposome: embracing the complexity for discovery in environmental health. *Environ. Health Perspect.* **124**, A137–A140 (2016).
31. Gao, P. et al. Spatial and temporal changes of P and Ca distribution and fractionation in soil and sediment in a karst farmland-wetland system. *Chemosphere* **220**, 844–650 (2019).
32. da Silva, E. B. et al. Background concentrations of trace metals As, Ba, Cd, Co, Cu, Ni, Pb, Se, and Zn in 214 Florida urban soils: different cities and land uses. *Environ. Pollut*. **264**, 114737 (2020).
33. Volckens, J. et al. Development and evaluation of an ultrasonic personal aerosol sampler. *Indoor Air* **27**, 409–416 (2017).
34. Contrepois, K. et al. Cross-platform comparison of untargeted and targeted lipidomics approaches on aging mouse plasma. *Sci. Rep*. **8**, 17747 (2018).
35. Gao, P. et al. Human exposure to polycyclic aromatic hydrocarbons: metabolomics perspective. *Environ. Int.* **119**, 466–477 (2018).
36. Gao, P. et al. Emerging and legacy PAHs in urban soils of four small cities: concentrations, distribution, and sources. *Sci. Total Environ*. **685**, 463–470 (2019).
37. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. **20**, 257 (2019).
38. Truong, D. T. et al. MetaPhlAn2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).
39. Franzosa, E. A. et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat. Methods* **15**, 962–968 (2018).
40. Wishart, D. S. et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res*. **46**, D608–D617 (2018).
41. Barupal, D. K. & Fiehn, O. Generating the blood exposome database using a comprehensive text mining and database fusion approach. *Environ. Health Perspect*. **127**, 97008 (2019).
42. Wishart, D. et al. T3DB: the toxic exposome database. *Nucleic Acids Res*. **43**, D928–D934 (2015).
43. García-Alcalde, F. et al. Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics* **28**, 2678–2679 (2012).

Acknowledgements

We thank X. Wang and X. Li for their contributions to the development of the methods. This work was supported by Zhejiang University Life Sciences Institute start-up funds and the Leona M. and Harry B. Helmsley Charitable Trust (grant G-2004-03820).

Author contributions

C.J. conceived the study. M.S. supervised the study. C.J., X.Z. and P.G. drafted and revised the protocol with input from Q.C. and M.S.

Competing interests

Two provisional patents were filed (pending application numbers 62/488119 and 62/488256).

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41596-020-00451-8.

Correspondence and requests for materials should be addressed to C.J. or M.S.

Peer review information *Nature Protocols* thanks Lea Maitre and Simone Morais for their contribution to the peer review of this work.

Received: 15 May 2020; Accepted: 20 October 2020;
Published online: 13 January 2021

Related links

Key references using this protocol

Jiang C. et al. *Cell* **175**, 277–291 (2018): https://doi.org/10.1016/j.cell.2018.08.060

Cissé O. et al. *mBio* **11**, e03138-19 (2020): https://doi.org/10.1128/mBio.03138-19

nature research

Corresponding author(s): Jaime Grutzendler

Last updated by author(s): Aug 18, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	LAS X for Leica confocal microscope Prairie View for Bruker two-photon microscope
Data analysis	ImageJ/Fiji for image visualization and analysis MATLAB for data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The mouse lines described in this protocol can be obtained from JAX and the rest upon request from the developer (see Materials - Animals). All the images presented in the figures are derived from raw data with minor contrast modifications. For the calcium time lapse imaging the raw movies can be obtained from supplementary data in (Hill et al. 2015). All data is available upon request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	N/A
Data exclusions	N/A
Replication	N/A
Randomization	N/A
Blinding	N/A

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging