UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Medicine and Life Sciences
School for Life Sciences
Master of Biomedical Sciences

Master's thesis

The association between indoor black carbon levels and bacterial communities measured with Hedera helix

Jorinde Bielen

Thesis presented in fulfillment of the requirements for the degree of Master of Biomedical Sciences, specialization Environmental Health Sciences

SUPERVISOR :
Prof. dr. Tim NAWROT

CO-SUPERVISOR :
dr. Hannelore BOVE

MENTOR :
Mevrouw Katrien WITTERS
dr. Hannelore BOVE

2018
2019

Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Medicine and Life Sciences
School for Life Sciences
Master of Biomedical Sciences

Master's thesis

The association between indoor black carbon levels and bacterial communities measured with Hedera helix

Jorinde Bielen

Thesis presented in fulfillment of the requirements for the degree of Master of Biomedical Sciences, specialization Environmental Health Sciences

SUPERVISOR :
Prof. dr. Tim NAWROT

MENTOR :
Mevrouw Katrien WITTERS
dr. Hannelore BOVE

CO-SUPERVISOR :
dr. Hannelore BOVE

Content

List of abbreviations

Afr	Africa
Amr	Americas
AQG	Air quality guidelines
ASV	Amplicon sequence variant
BC	Black carbon
CI	Confidence interval
CNS	Central nervous system
COPD	Chronic obstructive pulmonary disease
CVD	Cardiovascular disease
DADA2	Divisive Amplicon Denoising Algorithm
Eur	Europe
FEV1	Forced expiratory volume in 1 second
FVC	Forced vital capacity
GIS	Geographical information system
HIC	High-income countries
IQR	Interquartile range
LMIC	Low- and middle-income countries
MgSO4	Magnesium sulfate
NO_2	Nitrogen dioxide
PAH	Polyaromatic hydrocarbon
PCR	Polymerase Chain Reaction
PERMANOVA	Permutational analysis of variance
PM	Particulate matter
Sear	South-East Asia
TPAF	Two-photon excited autofluorescence
UFP	Ultrafine particles
WHO	World Health Organization
Wpr	Western Pacific
ZOL	Ziekenhuis Oost-Limburg

Acknowledgments

This thesis is the completion of my master biomedical sciences – environmental health sciences at Hasselt University. A thesis is never the work of one on his own, so I would like to thank everyone involved for their support and effort.

First, I would like to thank my promotor Prof. Dr. Tim Nawrot for the opportunity to perform my senior internship at the Centre for Environmental Sciences. Despite his busy schedule, he found the time to discuss the results and to help improve my thesis. Also, a special thanks to my supervisors Katrien Witters and Dr. Hannelore Bové for the feedback on this thesis and to help me with all the questions and problems I faced. Katrien, thank you for the motivational text message despite being on maternity leave.

Further, I would like to thank Dr. Esmée Bijnens and Vincent Stevens for their help with the statistics and to teach me how to work with the different statistical software programs. Also, thank you, Esmée, for giving feedback on my thesis and to help me to get the best out of it. I would also like to thank my second examiner Prof. Dr. Ann Cuypers for the advice you gave during my senior internship.

I would like to thank my fellow senior students at the Centre for Environmental Sciences, Stijn Vos, Kris Kunnen, Laurien Geebelen, An-Julie Trippas, and Ingeborg Pequeur, for the nice breaks together and our attempts to win some money in Qwistet.

Last, I would like to thank my parents for their support and understanding during my senior internship and for listening to me explaining my project even though you did not always understand. Thank you for supporting me not only during the senior internship but also during my entire bachelor and master training.

Abstract

Background: Air pollution is a serious problem in society and millions of people die prematurely from illnesses attributed to air pollution. Further, since most of our time is spent in the indoor environment, the focus will be on indoor air quality. An important air pollutant is particulate matter (PM), of which black carbon (BC) or so-called soot is the component of interest here. BC particles are fine particles with diameters smaller than 2.5 μm. They are produced during incomplete combustion reactions from diesel engines, stoves burning biomass, etc. These particles might also be associated with changes in microbial diversity since literature has shown an association between air pollution and microbial diversity. The hypothesis is that differences in the indoor BC levels result in a change in the diversity of the *Hedera helix*'s bacterial microbiome.

Materials and methods: In this study, indoor BC exposure is assessed with a biomonitoring system. The biomonitor used was the *Hedera helix* or common ivy since its leaves have several properties that allow BC accumulation. To measure the amount of BC present on the leaves, a two-photon femtosecond laser was used to visualize the particles and a MATLAB algorithm to calculate the number of particles and the area of the leaf covered by the particles. The validation of this biomonitoring system is the first objective. The second objective is to investigate the association between indoor BC and changes in microbial diversity on the leaves of the *Hedera helix*.

Results and discussion: First, no significant correlation ($r=0.35$, $p=0.15$, $n=19$) was found between the particle area on the leaves and an indoor BC aethalometer measurement. However, when using a mixed model taking into account the percentage of residential area present significance was reached ($p=0.03$). The percentage of residential area was added to help explain the particle area since the aethalometer only measured for 30 minutes while the plant accumulated BC for 42 days. Second, the association between the particle area and bacterial alpha-diversity was also not significant (Chao1 $p=0.38$, ASV richness $p=0.38$, Shannon-Wiener index $p=0.76$, Simpson index $p=0.88$). Obtaining significance for this association is hard since only data of 7 individuals were available indicating that non-significance could be due to low power. The PERMANOVA analysis also showed no association between the beta-diversity (Bray-Curtis dissimilarity matrix) and the particle area ($p=0.59$) but for only 7 individuals nearly half the variation could be explained by the particle area ($R2=0.48$).

Conclusion: From these results, we can conclude that the biomonitoring model can measure indoor BC but that more research is necessary to confirm the use of the model. Further, there was no association found between the alpha-diversity and beta-diversity, and the particle area but more power is required to obtain significant results. However, the PERMANOVA analysis of the beta-diversity showed that it is interesting to further investigate the association between BC and microbial diversity.

1 Introduction

Nowadays, air pollution is a hot topic in environmental research, but this was not always the case. The Lethal London Fog of 1952 is seen as the catalyst for research on air pollution epidemiology. The Lethal London Fog took place from December 5th to December 9th in 1952 and led to 3000 more deaths, a three times higher death rate than normal, for that time period. Moreover, it took until the end of February for the death rates to return to normal levels. Although data on air pollution from the period of the Fog is limited there was a significant correlation between the measured pollutant levels (sulfur dioxide and total suspended matter), and mortality and morbidity of cardiovascular and respiratory diseases (1). It is thus important to take the lessons learned from the London Fog and to monitor our air quality.

1.1 Air Pollutants

The air pollutant on which will be focused most here is particulate matter (PM). PM consists of a large variety of particles such as sulfates, nitrates, ammonia, sodium chloride, black carbon, mineral dust (such as asbestos), metal residues, and biological components (bacteria, fungi, viruses, pollen, and plant fibers) (2-5). This is subdivided in different fractions based on the particle diameter going from large particles (>10 µm), coarse particles or PM_{10} (<10 µm), fine particles or $PM_{2.5}$ (<2.5 µm) to ultrafine particles (<0.1 µm, UFP) (Figure 1.1) (2, 3, 6).

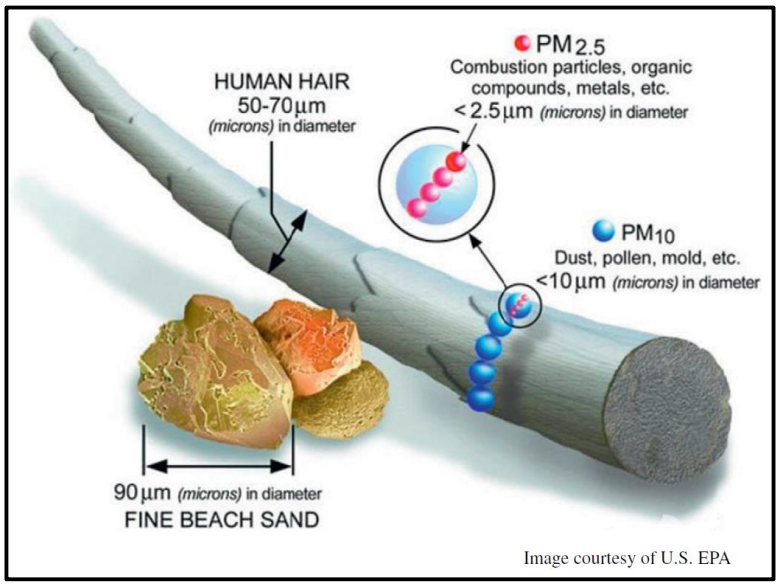


***Figure 1.1**: Comparison of the particle diameter of PM_{10} and $PM_{2.5}$ with human hair and sand particles (5).*

The World Health Organization (WHO) tracks PM air pollution globally, focusing on the PM10 and PM2.5 fractions which have a significant impact on health, and issued recommendation levels for them based on the impact on health. They recommend an annual mean of 10 µg/m3 and a 24-hour mean of 25 µg/m3 for PM2.5, and an annual mean of 20 µg/m3 and a 24-hour mean of 50 µg/m3 for PM10 (4). However, when looking at the compliance of these recommendations, they found that only 16% of the assessed population (43% of the global urban population) complied to the annual mean level recommendation, but this can be due to the fact that the recommendations are not always economically feasible (4). When looking at the interim targets for annual mean levels, which were made to allow an easier progression towards the issued guidelines, still a lot of people live in very polluted air. For interim target 3 (30 µg/m3 for PM10 and 15 µg/m3 for PM2.5) 27%, for interim target 2 (50 µg/m3 for PM10 and 25 µg/m3 for PM2.5) 46%, and for interim target 1 (70 µg/m3 for PM10 and 35 µg/m3 for PM2.5) 56% of the urban population assessed lives in compliance with the target levels (4). This indicates that it is known that approximately 18% of the global urban population lives in highly polluted air. Furthermore, the high-income countries of America, Western Pacific and Europe complied the best, and the high-income countries of the Eastern Mediterranean, low- and middle-income countries in Western Pacific and the South-East of Asia complied the worst (4). The change in air pollution levels over 5 years (2008-2013) showed a global increase of 8% in the assessed cities indicating that air pollution is still an actual worldwide problem that needs attention (4).

1.1.1 Black carbon

In this study, the focus is on one specific component of PM named black carbon (BC) or so-called soot. The particles are smaller than 2.5 µm, making it a component of $PM_{2.5}$, and consist of an elemental carbon core coated with various substances such as polyaromatic hydrocarbons (PAHs) (7-10). This indicates that BC can act as a carrier for other harmful substances. Further, BC is produced during incomplete combustion reactions such as those from diesel-powered vehicles, stoves burning biomass, forest fires, etc. (11). The major source of BC in Europe is transportation and living close to high-traffic roads increases the BC exposure (11).

1.1.2 Impact on health

PM and BC are important air pollutants since they are associated with various health effects. Exposure to PM can lead to serious health problems. The size of the particle determines its health effects because smaller particles, like UFPs, can penetrate deeper into the alveoli (2). The particles from the large fraction are usually removed at the upper airways, the PM_{10} particles can enter the lungs beyond the trachea and large bronchi, and the $PM_{2.5}$ particles can be deposited in the smaller airways and

alveoli (Figure 1.2) (3). From the alveoli, the particles can enter into the bloodstream and affect other organ systems (3). PM exposure, especially $PM_{2.5}$, has been associated with pulmonary, and cardiovascular diseases (3, 6). The WHO states that millions of people die prematurely from illnesses attributed to air pollution such as stroke, ischemic heart disease, Chronic Obstructive Pulmonary Disease (COPD) and lung cancer in adults, and pneumonia in children under 5 years of age (4, 12).

Further, the effects of BC are similar to those from $PM_{2.5}$, since BC is a component of $PM_{2.5}$, and BC has been associated with cancer, respiratory diseases, cardiovascular diseases (CVD) and central nervous system (CNS) effects (7, 11, 13, 14). BC can cause cancer, such as respiratory and urinary tract cancers, by influencing gene expression profiles and PAH related carcinogenicity (13). BC's association with respiratory diseases, such as asthma and COPD, is mediated through triggering inflammation (13, 14). The CVDs associated with BC exposure, such as myocardial infarction and coronary artery disease, are caused by BC influencing the endothelium and triggering inflammation (7, 11, 13, 14). The effects on the CNS effect are mostly seen in cognitive development (11). This all indicates that due to the impact on health by PM and BC, intensive research and monitoring is required.

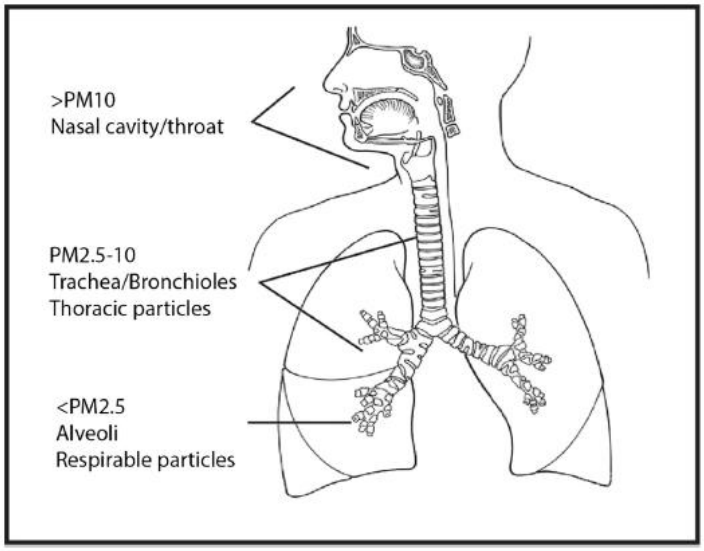


Figure 1.2: Deposition of PM in the respiratory tract (15).

1.2 Indoor air pollution

The previous part described the impact air pollution can have, but this impact is determined by the time spent in contact with the air pollution and the place where people spend most of their time is the indoor environment (16). Of our time spent indoors, 65% is spent at home making it important to maintain good air quality at home (16). However, many air pollutants can be found in the home environment such as carbon dioxide, carbon monoxide, nitrogen dioxide (NO_2), volatile organic

compounds, PM, PAHs, and environmental tobacco smoke (2, 3, 17, 18). Indoor sources of PM and BC include cigarette smoke, heating or cooking, cleaning, and PM and BC produced outdoors (2, 3). The outdoor pollution can enter indoors through ventilation and vice versa (16). This makes indoor air quality season depending since people tend to ventilate more in summer when the temperatures are higher (16).

1.3 Indoor Air Pollution Monitoring

In order to get a good view of the impact of PM and BC, long-term monitoring with high spatial resolution is ideal. Nowadays, we have several different types of samplers which can be divided into active and passive samplers (15). Active samplers actively suck in air that is assessed for air pollutants. Passive samplers collect air pollutants by means of gravitational forces, inertia, electrostatic attraction or convective diffusion (15). However, these devices are quite expensive with an average price of 3000 euros per device, which is a problem in epidemiological studies were hundreds of participants need to be monitored (19-22). Another disadvantage is that some samplers, namely active samplers, require a stable power source and frequent maintenance which does not allow them to be operated unattended and/or in remote areas (15). All these disadvantages indicate that an alternative is wanted, and this was found in the use of plants as a biomonitoring system.

1.3.1 Monitoring particulate matter and black carbon using plants

To use plants as a biomonitor for PM including BC, the ability of plants to accumulate PM on their leaves has to be assessed and several studies already confirmed this ability (22-25). It was shown that plants were able to intercept 17.7 to 74.1 % of the PM in the air indicating that they can be used as a biomonitor (25). Furthermore, the properties of the leaf found to be important for PM accumulation included leaf density, stomatal roughness, shape and the waxy layer of the leaf (23-27). The rougher the surface and the more secrete or waxy layer it contains the more particles can stick and accumulate on the leaf (23, 26, 27). Also, a higher number of stomata and the presence of hairs, trichomes, increases the accumulation of particles (23). The accumulation of BC occurs mostly in the waxy layer since this layer is especially effective in accumulating $PM_{2.5}$ (26, 27). However, factors such as rain and wind can influence the accumulation of PM and BC but, since this is only true for outdoor monitoring, plants can ideal monitors indoors (24, 25).

1.3.2 Monitoring using *Hedera helix*

In this study, there is one plant species of interest, namely, *Hedera helix* or common ivy. Common ivy has various advantages that make it very useful for this study (26, 27). One advantage is that common ivy demands little space and is more ideal for locations with high pollution levels (27). This is explained by the fact that the PM accumulation is not very high making common ivy less vulnerable to the harmful effects of PM (27). Another advantage is its evergreen nature which allows year-round sampling, sampling over longer periods of time and higher PM accumulation (28). Moreover, ivy can collect high levels of $PM_{2.5}$ and acts like a particle sink with the highest accumulation on the outer leaves. This can be explained by the large surface area and moist surfaces of the ivy leaves (29). The leaves also have other properties such as a rough wax surface, trichomes, and stomata present on the lower (abaxial) surfaces indicating an effective accumulation of PM (30, 31). In the end, it can be concluded that *Hedera helix* is suitable for indoor PM and BC monitoring due to these described advantages.

1.3.3 Detecting black carbon particles on the leaves

For measuring BC particles on the leaves, a technique developed at Hasselt University was used (32). In this technique, a two-photon femtosecond laser is used for the illumination of the BC particles on the ivy leaves causing them to generate a bright, white-light signal (Figure 1.3). This emission shows an intense and steep peak instead of emitting over a period of time. This is due to the particle returning quickly to its ground state instead of over a period (32).

Next to the BC particles, the leaf will also emit light which is called two-photon excited autofluorescence (TPAF). This emission is different since it will give a green-light instead of a strong white-light emission. The difference makes it possible to differentiate the signal from the particle with that of the TPAF. To differentiate between the TPAF and the particle two different channels (red and green) can be used since the particles' signal (white light) will be visible in both channels and TPAF will only be visible in one channel. Furthermore, a peak-find algorithm in MATLAB (MathWorks, Natick, Massachusetts, U.S.A.) can identify the peak intensities from the particles and by comparing whether they are found in both channels it can determine whether it is a particle of the background. In the end, the algorithm gives the number of particles found in the image and the surface the particles occupy (Figure 1.3) (32).

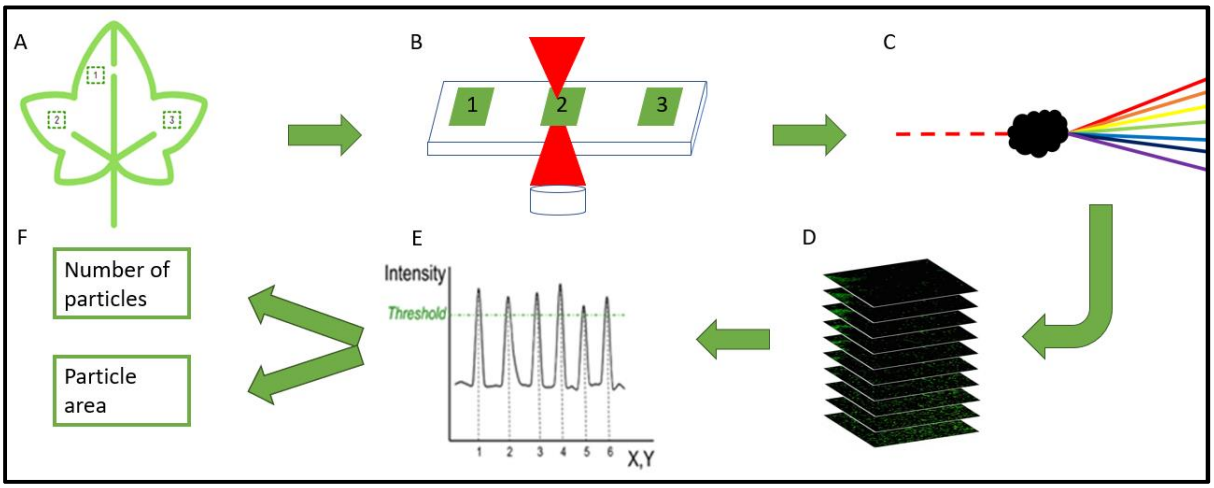


Figure 1.3: The workflow of the detection technique. *(A) Three spots per leaf were cut out and tapped on a coverslip. (B) Under the confocal microscope, a two-photon femtosecond laser is used to visualize the sample. (C) The femtosecond laser causes the black carbon particles to emit a bright white-light signal while the leaf produces two-photon excited autofluorescence (green light). (D) The light signals from the sample are captured into a picture and for each depth in the sample, a picture is made and stacked (Z-stack). (E) A MATLAB algorithm counts the peaks from the BC particles found in the images. (F) The output of the algorithm consists of the number of particles counted and the area covered by particles in the image.*

1.4 Black carbon and the microbiome

When monitoring BC exposure, we can link the BC levels to changes in the microbiome of *Hedera helix*, more specifically the epiphytes, bacteria present on the leaves. Researching the effects of BC on bacteria can help to elucidate the possible health effects of BC. However, the effect of BC on bacteria is not well studied, but the effect of PM is more studied. Furthermore, several studies have confirmed an association between PM and bacteria from the lung microbiome (33-35) and only one study found an association between BC present in the lungs and the lung microbiome (36). This indicates that air pollution and BC can influence bacterial communities. However, this involves another microbiome than that of interest in this study.

A study of Smets *et al.* investigated the difference in the bacterial community of ivy leaves between different urban and non-urban sites (37). They found that both urban and non-urban had an equal mean richness and diversity, but that certain taxa were more dominant in the urban sites and vice versa (37). However, this is performed in the outdoor environment while our focus is on the indoor environment and the plants in our study will also accumulate bacteria from the indoor environment next to their own microbiome. These bacteria are influenced in a different way as was investigated by Weikl *et al.* They found that indoor bacterial community variation was significantly influenced by ventilation of the living room through windows in winter, the type of living room floor, more than three occupants living in the home, and the home situated higher than the first floor (38). The type of living room floor with the highest variation were small-scale structured floors (smooth with rugs for

example) since different structures can harbor different bacteria (38). The most determining factor for community variation and diversity was the time of the year, which consisted of a shift in bacterial community variation (38). These studies indicate that it is interesting to investigate what changes occur when the bacterial community of ivy leaves is exposed to indoor BC.

1.5 This study

Since air pollution is a serious health threat, long-term and large-scale monitoring are advised. However, this is not economically feasible due to high costs. Therefore in this study, a plant-based model is used to monitor BC levels and to associate this with the bacterial community of the plant.

In this study, we investigated the association between indoor BC levels and the diversity of the microbiome of *Hedera helix*. It is hypothesized that differences in the indoor BC levels lead to a change in the diversity of the *Hedera helix's* bacterial microbiome. Hence, two objectives were set forward: first, to determine if the BC levels measured with the plant model correlates with several traffic indicators, land use variables and ambient air pollution variables associated with BC, and with aethalometer data on indoor BC. Second, to research if there is an association between the BC levels measured with the plant and the microbial diversity of the plant.

2 Materials and Methods

To validate the use of common ivy as a biomonitor and to study the association between indoor BC and bacterial diversity present on the leaves of the ivy, plants were distributed among the homes of participants. After a period, the leaves of the plants were collected and used to determine the levels of BC exposure and bacterial diversity. During the collection, BC was also measured with an aethalometer.

2.1 Study population

The study population used in this study is a subset (n=154) of the ENVIR*ON*AGE birth cohort (n=1700). The birth cohort recruits mother-child pairs at hospital East-Limburg (ZOL) since February 2010 to investigate the effects of the environment on the development of the children. The women are recruited during labor in ZOL hospital. The subset used in this study were mother-child pairs who completed the follow-up, step 2 of the cohort, when the child was 4 years of age. The parents of these children were called and asked if they wanted to participate and at a first house visit, an informed consent was signed by the parents. Participants were excluded if they moved away after the follow-up or planned to move away during the study period. Another exclusion criterium was if they were doing or planning to do home improvement works during the study period. In this study also 10 participants living ≤ 30 m and 10 living ≥ 1500 m from the nearest major road were selected, all other participants analyzed in this study were selected based on the availability of an indoor BC aethalometer measurement.

2.2 Origin, placement, and collection of the plants

Hedera helix plants were bought from Boomkwekerij Frijns (Margraten, the Netherlands) and were kept in the greenhouse of Hasselt University (60% humidity, day-night cycle from 7 am to 10 pm, day temperature 23°C, night temperature 18°C, and a light intensity of 300 W m^{-2}) until the day of placement. The plants were placed in the homes of the participants in the spring (March-June) of 2017 and 2018.

On the day of placement, the plants were washed with sterile magnesium sulfate ($MgSO_4$) to remove most of the dirt, dust, and bacteria present on the leaves. Under the laminar air flow, a control leaf was taken using a sterile scalpel and sterile forceps, and the plant was wrapped in a plastic bag to prevent contamination with BC from the environment. The control leaf was taken to know the BC present on the leaf before the exposure in order to calculate the total amount of BC accumulated at

the exposure site. At the home of the participant, two plants were unwrapped and placed in the living room, on a spot with indirect sunlight, ideally on a windowsill. The participants were told to water the plants occasionally, preferably once a week, and not to touch, dust or move the plant.

After approximately 42 days, a second house visit was conducted. During this visit, four leaves were taken from one plant using a sterile scalpel and sterile forceps, and each leaf was put in a separate Petri dish with the front facing the Petri dish. The Petri dishes were stored away from humidity, heat, cold, or direct sunlight. In addition, also leaves for microbiome analysis were collected. Two to five leaves were transferred into one to five pre-weighed 50 ml Falcon tubes with 30 ml 1x sampling buffer per ID (Supplemental protocol 1) using sterile forceps. The tubes were weighed, for the quantification of total leaf material collected, and stored at 4°C until further analysis.

2.3 Aethalometer measurement

Next to a BC measurement with the plant, for some participants also a BC measurement with an aethalometer (microAeth/AE51, aethlabs, San Francisco, California, U.S.A.) was performed during the second house visit. The aethalometer measured the BC each minute during 30 minutes of measurement with a flow rate of 99 ml/min. The data received from the device can give negative values which were smoothed using the Optimized Noise-reduction Averaging algorithm with an incremental light attenuation of 0.01.

2.4 Epiphyte isolation, DNA isolation, and sequencing

The bacterial analysis used the 50 ml Falcon tubes with leaves from the sampling. Bacteria were isolated from the leaf samples using a standardized protocol, after which DNA from the bacteria was isolated and sequenced.

2.4.1 Isolation of epiphytes

After the collection of the leaf samples into the sampling tubes, the tubes were weighed to know the total amount of leaf material collected. The samples were stored in a cooling chamber (4°C) until the isolation was performed. The bacteria were isolated by shaking the tubes for three hours followed by four rounds of centrifugation (Supplemental protocol 1).

As an additional step, the bacteria from the leaves were cultivated on 1/10 869 rich medium plates and purified for each bacterial strain found (Supplemental protocol 2 and Supplemental table 2). For each of the strains seen it was recorded what their appearance was and how many colonies were seen

for each ID. The purification consisted of cultivating the strain's colony separately (in four-fold if possible) on 869 rich medium plates (Supplemental protocol 3 and Supplemental table 2). When the colonies were pure they were put on a 15%$_w$ glycerol stock in the freezer (-45°C) for further use (Supplemental table 3).

2.4.2 DNA isolation

The bacterial DNA was isolated using the Macherey-Nagel NucleoSpin® Soil kit (Macherey-Nagel GmbH & Co. KG; Düren, Germany) (Supplemental protocol 4). Sample lysis was performed with a Retsch Mixer Mill MM 400 (Retsch; Aartselaar, Belgium) at 12 Hz for 10 minutes, followed by centrifugation at 16 000 g and incubation. The lysate was filtered, and DNA was bound to a silica membrane. Then the silica membrane was washed four times using 650 µl SW2 and dried. At last, the DNA was eluted using 50 µl Buffer SE, and quality and quantity were checked spectrophotometrically before 70 µl was stored at -45°C and the rest (approximately 30 µl) at -20°C.

2.4.3 Sequencing of the bacterial 16S rRNA gene

The sequencing was prepared with the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, California, USA) and Nextera XT Index Kit v2 Set A (Illumina, San Diego, California, USA) (Supplemental protocol 5). First, an amplicon PCR was performed using primers for only the V4 region of the 16S rRNA gene (Supplemental table 4, Supplemental table 5, and Supplemental table 8) and was followed by a PCR clean-up (Supplemental protocol 5). This clean-up was performed using AMPure beads (20 µl for 25 µl PCR product) which were dried for 2.5 minutes and 27.5 µl 10 mM Tris pH 8.5 was used to retrieve the clean PCR product. Next, an index PCR was performed to tag each sample library with a unique barcode, allowing them to be pooled during the sequencing, followed by a second clean-up with AMPure beads (Supplemental protocol 5, Supplemental table 6, Supplemental table 7, Supplemental table 8, and Supplemental table 9). After the index PCR, gel electrophoresis was done to determine whether the PCR had gone up and if the product was of the correct length. At last, the libraries were quantified, normalized and pooled (Supplemental protocol 5).

The sequencing was done with the PhiX Control Kit v3 and the MiSeq Reagent Kit v3 for the Illumina MiSeq system (Illumina, San Diego, California, USA) (Supplemental protocol 6 and Supplemental table 8). The pooled final DNA library and the PhiX control were denatured by adding 0.2 N NaOH and centrifuging at 280 g. After this pre-chilled HT1 was added and the DNA was diluted to its final concentration. Finally, the DNA library and PhiX control were combined and heat denatured before being loaded into the MiSeq reagent cartridge. The sequencing method consisted of next-generation-

sequencing with index tagged libraries (Supplemental figure 1). In this sequencing method, during the clustering step, the DNA single strands are attached to grafted oligo's (P7 and P5) on the flow cell and after attachment, the strands on the P5 oligo's were removed leaving only the P7 strands in the flow cell. Now, the read one primer is annealed and the strand is sequenced followed by the index one primer annealing and sequencing. Next, the strand anneals to the grafted P5 oligo and index two is sequenced followed by amplification of the strand. The strands attached to the P7 oligo's are removed leaving only the strands attached to the P5 oligo's. Last, the read two primer anneals and the new strand is sequenced.

2.5 Confocal microscopy

BC was measured on the leaves using confocal microscopy. First, the leaves, both control and exposed, were photographed to calculate the total leaf surface with MATLAB (MathWorks; Natick, Massachusetts, U.S.A.). Leaves were then flattened out and taped into separate Petri dishes and stored away from humidity, heat, cold, or direct sunlight until further analysis. When the leaves were dry, three spots were cut out (Figure 1.3) and taped on a microscope coverslip with the backside of the leaves facing the glass.

These cut-outs were then analyzed with an LSM510 META NLO (Carl Zeiss, Jena, Germany) mounted on an Axiovert 200 M (Carl Zeiss, Jena, Germany) with a two-photon femtosecond pulsed laser (MaiTai DeepSee, Spectra-Physics, Santa Clara, USA) exciting at 810 nm using a 10x objective (Plan-Neofluar 10x/0·3, Carl Zeiss). The speed of the laser was set at 2.51 µs/pixel, and six different spots per cut out were chosen for analysis and for each spot a Z-stack of 6.62 µm thick slices was made (Figure 1.3). The two channels used for the detection of the emission light were the red (400-410 nm) and green (450-650 nm) channels (Supplemental protocol 7). These images were then analyzed using a MATLAB algorithm.

2.6 Data processing

The previous experiments generated large sets of data that were processed using specific programs. This processing is explained in this section including the acquisition of GIS data to correlate indoor BC levels with certain outdoor elements influencing the ambient BC levels.

2.6.1 Sequencing data

The data generated by the sequencing was processed with the DADA2, Divisive Amplicon Denoising Algorithm, package in R. This algorithm identifies the species present based on amplicon sequence variants (ASVs). ASVs are determined by a de novo process discriminating biological sequences from errors. The reference database used to assign the taxonomy was the Silva reference database version 128 train set. After the taxonomy was assigned to each ASV, total number of reads, alpha-diversity measures (Shannon-Wiener index, Simpson's index, Chao1 index, and ASV richness) and beta-diversity (Bray-Curtis dissimilarity matrix) were calculated for each ID. These calculations were made on the data stripped from sequences identified as eukaryotic, archaic, mitochondrial or chloroplast, or that could not be identified at the kingdom level.

2.6.2 Confocal microscopy data

The confocal images were processed viewing in the red (400-410 nm) and green (450-650 nm) channels and particles were identified by the algorithm if the threshold was exceeded in both channels (threshold for channel 1 was 0 and for channel 2 was 410) (Figure 1.3). The output of the algorithm was the total surface of the leaf that was taken in by BC particles (μm^2), the total number of particles present on the leaf (aggregates not taken into account), and the total surface (in pixels) of the image. Only the area covered by BC particles was used in the further analysis since the number of particles is an underestimation because the algorithm cannot determine the difference between 1 particle and an aggregate of particles.

2.6.3 GIS data

Using ArcMap ArcGIS software (Esri BeLux S.A., Wemmel, Belgium) several variables such as traffic indicators, land use variables and air pollution data were determined for all participants. For the traffic indicators, the distance to the nearest roads was calculated for highways, national roads and major roads (closest highway or national road). For the land use variables, the surface area (square meters) of residential areas, natural areas, and agricultural areas were calculated within a buffer of the address using CORINE Landcover 2012. For each participant, it was determined whether they lived in an urban, suburban or rural area, and what the annual mean levels of BC, NO_2, $PM_{10,}$ and $PM_{2.5}$ of 2015 were.

2.7 Statistics

For data management and statistical analyses, we used SAS 9.4 Statistical software (SAS Institute Inc., Cary, NC, USA). First, for the area of leaf covered by BC particles, outliers were identified, and normal distribution was tested with the Shapiro-Wilk test. IDs with more than 5 outliers in particle area were removed from the analysis and the outliers in the other IDs were left in. Next, the particle area was correlated with BC aethalometer levels, traffic indicators, land use variables, annual mean air pollution concentrations, the total number of sequence reads and alpha-diversity measures. The Pearson Correlation Coefficient was used when the data was normally distributed, and the Spearman Rank Correlation Coefficient was used when the data was not normally distributed. Mixed models were performed to take the multiple measurements of particle area for each ID into account. The mixed models were made for aethalometer levels, traffic indicators, land use variables, and air pollution and for each model the estimate and the 95% confidence intervals (CI) were determined. Mixed modeling was performed adjusted for covariates selected *a* priori. The models with the aethalometer data were corrected for the leaf and spot of BC particle area measurement. All other models were corrected for the leaf and spot of BC particle area measurement and the number of leaves measured. Last, to test whether the particle area was associated with the beta-diversity a PERMANOVA analysis was performed in R 3.5.3 statistical Software (R Foundation for Statistical Computing, Vienna, Austria). For the PERMANOVA the IDs were grouped based onto the mean particle area into a low (<100 000), medium (100 000 -200 000), high (200 000 – 300 000), or very high (>300 000) group. Statistical significance was reached when the p-values were lower than 0.05.

3 Results

Data on BC present on the leaves was collected for 34 participants, and for all participants, traffic indicators, land use variables and annual mean air pollution concentrations were calculated, and for 19 participants there was aethalometer data available. A table containing the average area taken by BC particles on the leaves for each ID is available in the supplement (Supplemental table 10). One participant was removed since too many outliers were identified in the area covered by BC particles, giving 33 IDs available for analysis.

General data on the study population is presented in Table 3.1 and in Figure 3.1.the locations of the participants' homes are visualized. Participants initial recruited from ENVIR*ON*AGE was at the East-Limburg Hospital in Genk (Belgium). The catchment area of the hospital includes the province of Limburg, (Flanders, Belgium) and combines urban, suburban and rural areas. Most participants are living in a rural area with a high percentage of agricultural areas but a low percentage of natural areas in 1000 m.

Table 3.1: Characteristics of the study population

Characteristic	
General	n=34
Degree of urbanicity	
Urban area (%)	8 (23.53)
Suburban area (%)	4 (11.76)
Rural area (%)	22 (64.71)
Participants moved	11 (32.35%)
Traffic indicators	
Distance to major roads (m)	733.5 ±907.1
Land use variables	
Residential areas in 1000 m (%)	42.50 ±18.56
Agricultural areas in 1000 m (%)	39.42 ±24.14
Natural areas in 1000 m (%)	12.62 ±13.09
Outdoor air pollution	
Annual mean BC (µg/m^3)	1.054 ±0.15
Annual mean NO_2 (µg/m^3)	15.11 ±3.22
Annual mean PM_{10} (µg/m^3)	16.83 ±1.50
Annual mean $PM_{2.5}$ (µg/m^3)	11.24 ±0.68
Indoor BC	
Average BC area on leaves (µm^2)	2.39 x10^5 ±126516
Average BC, aethalometer measurement (ng/m^3)	886.4 ±457.1

Data presented are means ± standard deviation or number (percentage). BC: black carbon; NO_2: nitrogen dioxide; PM_{10}: particulate matter with diameter <10 µm; $PM_{2.5}$: particulate matter with diameter <2.5 µm.

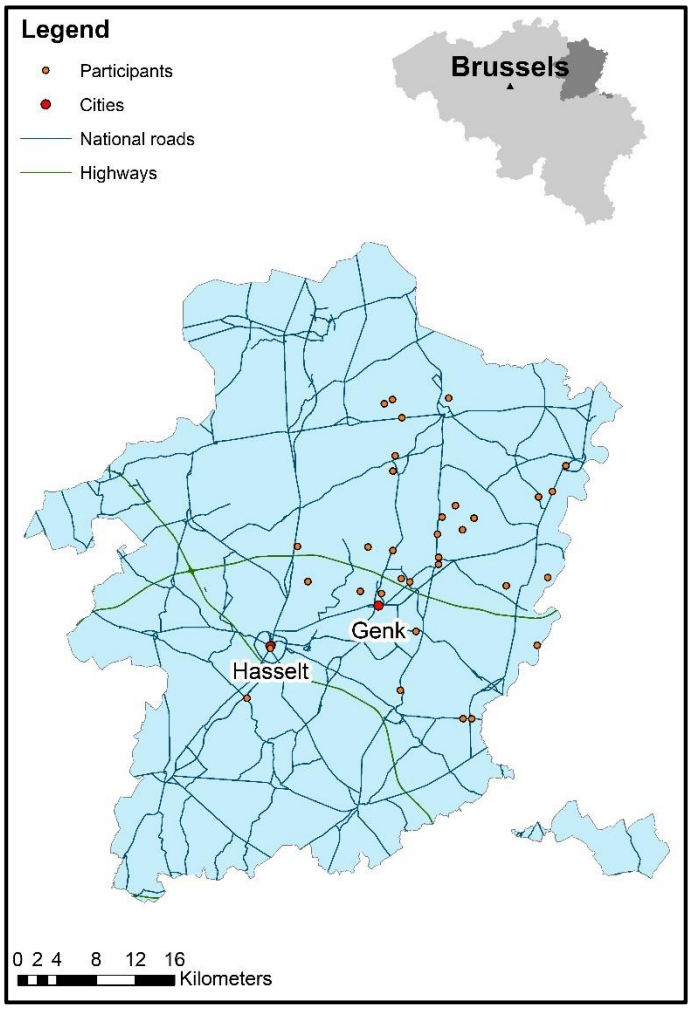


Figure 3.1: A map of Limburg showing the location of the participants' homes, major cities, national roads, and highways.

3.1 Validation of the plant model

To validate if the plant model is useable, aethalometer measurements of 30 minutes were performed for 19 participants. The unadjusted model shows a positive association (r=0.35, p=0.15, n=19) between the BC concentrations measured by the aethalometer and the mean BC particle area measured on the leaves(Figure 3.2). After adjusting for the multiple measurements as well as the leaf and spot of the particle area measurement, a positive association was still observed between the BC concentrations measured by the aethalometer and the BC particle area measured on the leaves (Table 3.2). This was seen in an IQR increase (566.17 ng/m^3) in BC measured by the aethalometer associated with an increase of 6.38×10^4 µm² in particle area (95% CI: -1.40×10^4 to 1.42×10^5 µm², p= 0.11, n=19). After additional adjustments for the percentage of residential areas found in 1000 m around the participants address, an IQR increase (566.17 ng/m^3) in BC measured by the aethalometer is associated with a significant increase of 7.78×10^4 µm² (95% CI: 6.57×10^3 to 1.49×10^5, p= 0.03, n=19) (Table 3.2).

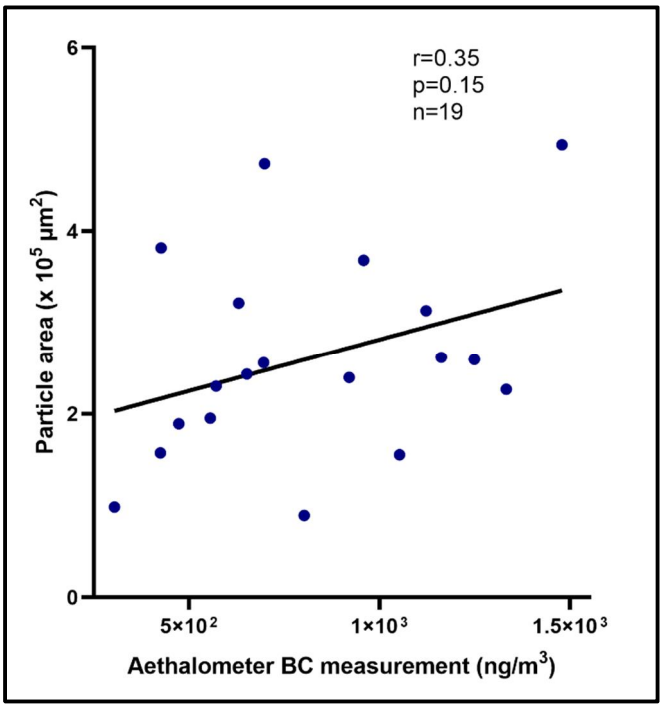


Figure 3.2: The Spearman Rank correlation between the aethalometer measurement (ng/m³) and the mean particle area of BC (µm²) measured on the H. helix leaves.

Table 3.2: The mixed models for the aethalometer measurement explaining the particle area of BC measured on the H. helix leaves.

Model	Estimated change (µm²)	95% CI	P-value
Aethalometer*	$6.38x10^4$	$-1.40x10^4$ to $1.42x10^5$	0.11
Aethalometer adjusted for percentage of residential areas within 1000 m buffer*	$7.78x10^4$	$6.57x10^3$ to $1.49x10^5$	0.03

** Adjusted for the leaf and spot of the particle area measurement.*

3.2 Indoor black carbon and traffic indicators

The association between the shortest distance to a major road and the measured mean BC particle area on the leaves was not significant in an unadjusted model (r=-0.08, p= 0.65, n= 33) (Figure 3.3). After adjusting for multiple measurements, the leaf and spot of the particle area measurement, and the number of leaves measured, a 100 m increase in distance to a major road was associated with a decrease of $2.97x10^3$ µm² in particle area (95% CI: $-6.49x10^3$ to 558 µm², p= 0.10, n=33).

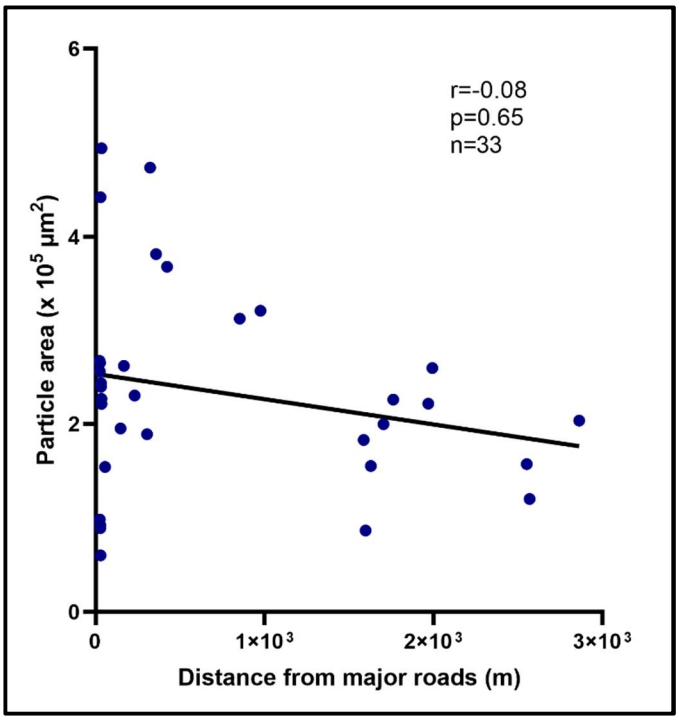


Figure 3.3: The Spearman Rank correlation between the distance from a major road (m) and the mean particle area of BC (µm²) measured on the H. helix leaves.

3.3 Indoor black carbon and land use variables

For the association between the BC particle area and land use variables, the percentage of residential, agricultural and natural areas in a 1000 m buffer around the home address of the participants were used. All models were corrected for multiple measurements, the leaf and spot of the particle area measurement and the number of leaves measured (Table 3.3). An IQR increase (27.65%) in residential areas was associated with a $4.83x10^4$ µm² increase in particle area (95% CI: -515 to $9.70x10^4$, p= 0.05, n=33). For the agricultural area model, an IQR increase (38.38%) gave an increase of $-3.77x10^4$ µm² in indoor BC (95% CI: $-9.10x10^4$ to $1.56x10^4$, p= 0.17), while the natural area model gave for an IQR increase (17.20%) an increase in indoor BC of 4016 µm² (95% CI: $-4.06x10^4$ to $4.86x10^4$, p= 0.86, n=33).

3.4 Indoor black carbon and outdoor air pollution

The particle area of the leaves was also modeled with the annual mean levels of outdoor BC, NO_2, PM_{10} and $PM_{2.5}$ (µg/m^3) from 2015 for all 33 participants (Table 3.3). An IQR increase (0.19 µg/m^3) in outdoor BC was associated with an increase of $1.79x10^3$ µm² BC particle area on the leaves (95% CI=$-4.06x10^4$ to $4.42x10^4$ µm², p=0.93, n=33). For NO_2 an IQR increase (4.12 µg/m^3) was associated with a change of -19224 µm² in BC particle area (95% CI=$-6.27x10^4$ to $2.43x10^4$ µm², p=0.39, n=33). The models of PM_{10} and $PM_{2.5}$ gave, respectively, for an IQR change of 1.28 and 0.95 µg/m^3 a change in BC particle area of

-1.02x10^4, and 7.12x10^3 µm² (95% CI=-3.87x10^4 to 1.83x10^4 µm² and -4.01x10^4 to 5.44x10^4 µm², p= 0.48 and 0.77, respectively for PM_{10} and $PM_{2.5}$, n=33).

Table 3.3: The mixed models of traffic indicators, land use variables, and air pollution levels explaining the particle area of BC (µm²) measured on the H. helix leaves.

Model	Estimated change (µm²)	95% CI	P-value
Traffic indicators (increase in 100 m)			
Distance to major roads (m)*	-2.97x10^3	-6.49 x10^3 to 558	0.10
Land use indicators (increase in IQR)			
Percentage of residential area within a 1000 m buffer*	4.83x10^4	-515 to 9.70x10^4	0.05
Percentage of agricultural area within a 1000 m buffer*	-3.77x10^4	-9.10x10^4 to 1.56x10^4	0.17
Percentage of natural area within a 1000 m buffer*	4.02x103	-4.06x10^4 to 4.86x10^4	0.86
Air pollution (increase in IQR)			
Annual mean BC 2015 (µg/m^3)*	1.79x103	-4.06x10^4 to 4.42x10^4	0.93
Annual mean NO_2 2015 (µg/m^3)*	-1.92x10^4	-6.27x10^4 to 2.43x10^4	0.39
Annual mean PM_{10} 2015 (µg/m^3)*	-1.02x10^4	-3.87x10^4 to 1.83x10^4	0.48
Annual mean $PM_{2.5}$ 2015 (µg/m^3)*	7.12x103	-4.01x10^4 to 5.44x10^4	0.77

**Adjusted for the leaf and spot of the particle area measurement and the number of leaves measured.*

3.5 The microbiome of *Hedera helix* leaves

From all successfully sequenced IDs (n=51) only 7 participants had data on BC particle area available. The other participants with BC particle area data were not successfully sequenced. To assess whether sequencing was deep enough a rarefaction plot was made (Supplemental figure 2) and shows that sequencing was deep enough since the graph goes to a plateau level. Also, for all sequenced IDs (n=51) the top 20 of most abundant taxa was constructed and the relative abundances of the phyla in all samples together were calculated (Figure 3.4 and Figure 3.5). The most abundant genera were the *Burkholderia-Paraburkholderia, Enterobacter, Gilliamella, Pantoea,* and *Pseudomonas* and the most abundant phyla were the *Proteobacteria* (97.63%)*, Firmicutes* (1.15%)*,* and *Actinobacteria* (0.71%).

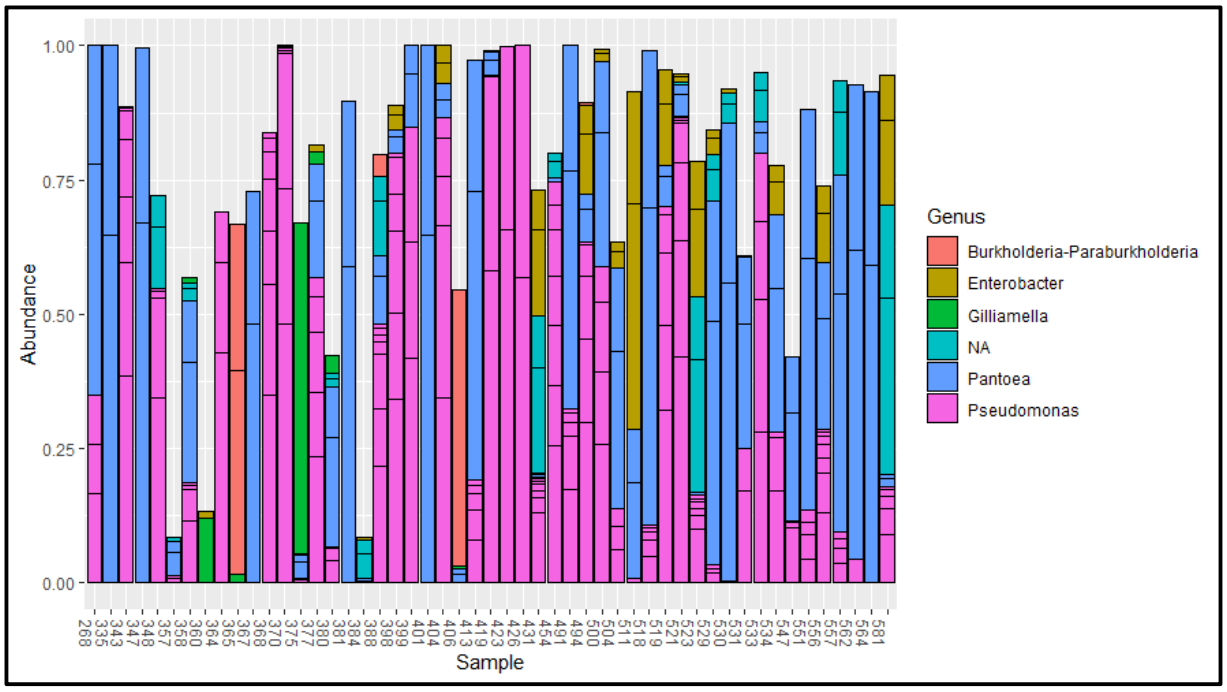


Figure 3.4: The abundances of genera from the top 20 ASVs found in all sequenced samples. *NA: Unclassified genus.*

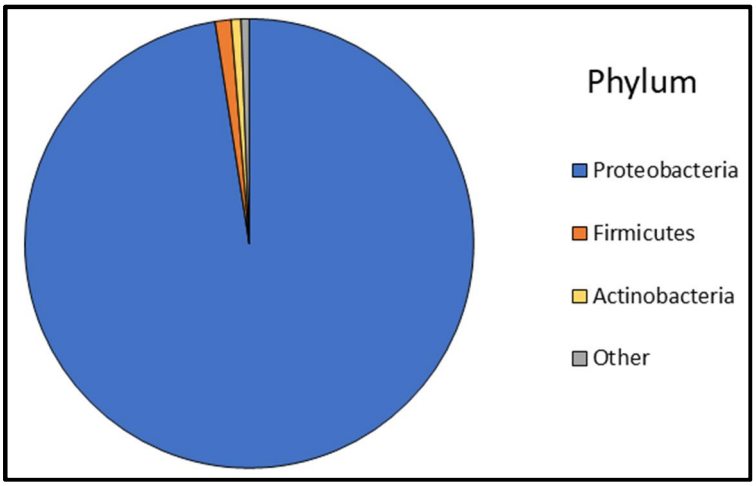


Figure 3.5: The relative abundances of the phyla from all samples together.

3.6 Black carbon levels associated with the bacteria from leaves

From all sequenced IDs (n=7) with BC particle area data available alpha diversity measures (Chao1, ASV richness, Shannon-Wiener index, and Simpson's index) were calculated, and associations were made between the alpha-diversity and number of sequence reads, and the BC particle area on the leaves. No significant association was observed between the number of sequence reads and the mean BC particle area (r= -0.21, p= 0.64, n=7) (Figure 3.6). Although not significant, a negative association was observed between mean particle area and alpha-diversity, based on 4 different measures (Figure 3.7);

Chao1 (r= -0.39, p= 0.38, n=7), number of ASVs (r= -0.39, p= 0.38, n=7), Shannon-Wiener index (r= -0.14, p= 0.76, n=7) Simpson index (1-D) (r= -0.07, p= 0.88, n=7).

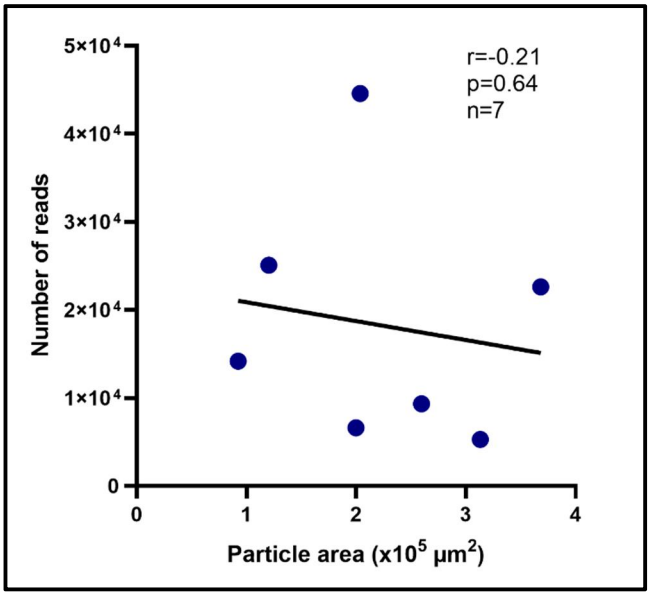


Figure 3.6: The Spearman Rank correlation between the mean particle area and the number of sequence reads.

For the PERMANOVA analysis, the subjects were split into different groups based on the mean particle area. The low group included mean particle areas lower than 100 000 (n=1), the medium group between 100 000 and 200 000 (n=2), the high group between 200 000 and 300 000 (n=2), and the very high group higher than 300 000 (n=2). First, the homogeneity of the variance in distance from the centroid, geometric mean, for each cluster group of BC level was analyzed. This showed that the variance was unequal between the groups but having equality between the groups with only 7 participants in total is nearly impossible. Next, it was tested whether the variance in composition was significantly different. The variance in composition was not significantly different (p= 0.59) but 48% of the variance could be explained by the BC groups (r=0.48).

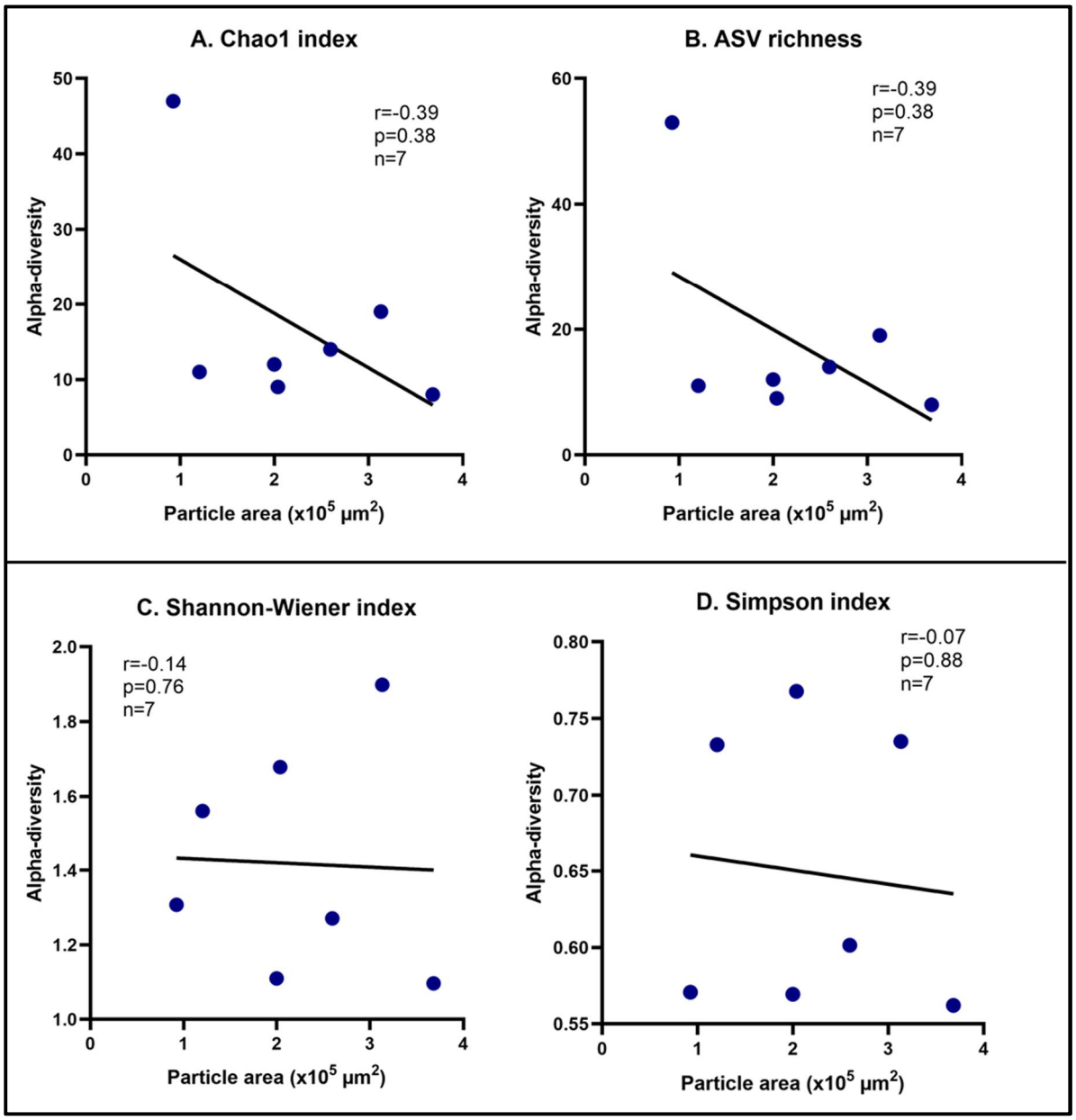


Figure 3.7: The Spearman Rank correlation between the alpha-diversity measures and the mean particle area. *(A) The correlation with the Chao1 index, (B) the correlation with the ASV richness, (C) the correlation with the Shannon-Wiener index, and (D) the correlation with the Simpson (1-D) index.*

4 Discussion

This study was set up to determine whether the *Hedera helix* is a good biomonitor for indoor BC and if indoor BC is associated with changes in microbial diversity of the *Hedera helix* leaf microbiome. The association between indoor BC measured with an aethalometer and indoor BC particle area covered on the leaves was positive and significant when taking into account the percentage of residential areas (estimate=$7.78x10^4$ μm^2, p=0.03). Further, no significant associations were found between microbial alpha- and beta-diversity, and the BC particle area on the leaves. However, in the alpha-diversity measures a negative trend (Chao1: r=-0.39, p=0.38, ASV richness: r=-0.39, p=0.38, Shannon-Wiener: r=-0.14, p=0.76; Simpson: r=-0.07, p=0.88) was seen as expected and the beta-diversity could be explained for 48% by the BC particle area (R^2=0.48, p=0.59).

First, it had to be validated if the use of *Hedera helix* as a biomonitor tool for indoor BC was feasible. Therefore, the association between aethalometer measurements of indoor BC and area covered by BC particles on the leaves was investigated. A positive association was observed between indoor BC aethalometer measurements and BC particle area on the leaves, but this was not significant in an unadjusted model (r=0.35, p=0.15). A mixed model taking the multiple measurements into account adjusted for the leaf, spot of the particle area and the percentage of residential area measurement showed a significant association (estimate=$7.78x10^4$ μm^2, p=0.03). The adjusted model without the residential areas gave a non-significant association (estimate=$6.38x10^4$, p=0.11). We expected that the aethalometer measurements and particle area on the leaves would correlate significantly without the percentage of residential areas since both are used to measure the indoor BC. However, the difference between the two methods, being the accumulation time for BC, was too large.

A study by Baldacchini *et al.* indicated that using leaves as an outdoor PM monitor can be useful although it was not as uniform as an aethalometer (39). They found that the particles' surface density of fine and coarse particles, 0.3-0.2 µm and 2.5-10 µm diameters respectively, and the composition of the on the leaf deposited particles was different between street and park sites. However, the unadjusted model gave a p-value of 0.18 and the adjusted model gave a p-value of 0.34 for the sampling site, indicating that the differences were not significant like for this study's unadjusted and adjusted only for leaf and spot of particle area measurement models (39).

A plausible explanation for the non-significant results in the unadjusted and adjusted only for leaf and spot of particle area measurement models is that the aethalometer only measured the indoor BC for 30 minutes while the plant accumulated BC particles for 42 days on average. Also, there was low

statistical power with only 19 participants which makes it harder to see a significant difference. Consequently, to adjust for BC exposure not measured by the aethalometer we added the percentage of residential areas in a 1000 m buffer around the residential location to the model since this indicates urbanicity of the location and urban sites have higher BC levels (40). However, it is promising that there is a significant association found since this indicates that it is interesting to further investigate the use of *Hedera helix* as a biomonitor which will reduce the costs of long-term monitoring in large epidemiological studies.

Further, the associations between the indoor BC measured on the leaves and traffic indicators, land use variables, and outdoor air pollution were assessed. The negative association with the distance from a major road was not significant (estimate=-2.97x10^3 µm^2, p=0.10), but it is near statistical significance. A association would be expected since the closer to a major road the higher the BC concentrations (41). A study on the influence of road proximity on indoor air pollutant concentrations found significant associations for $PM_{2.5}$, BC, Fe, Mo, and Sr concentrations (41). The study of Baxter *et al.* found significant associations between indoor BC, and ambient concentrations and distance to nearest designated truck route (42).

However, these studies do not explain the non-significant result found in this study. This can be the result of a different ventilation behavior between participants since ventilation can significantly influence the indoor concentrations (2, 43). A study on the influence of ventilation on the indoor air quality found that ventilating for one hour before rush hour significantly reduced indoor PM concentrations (44). Thus, it is plausible that some participants chose to ventilate less or avoided ventilation during rush hour. Further, another study has found a significant association between local traffic and the indoor concentration, and confirmed that ventilation has a significant impact on this relation (42). This indicates that when assessing the association between traffic indicators and indoor BC concentrations extensive information on the ventilation habits of the participants is required. The non-significant result for the association can also be explained by the low statistical power of this study while other studies had a higher statistical power.

The association with the land use variables was also not significant but showed promising relationships. The model with the residential areas showed an expected positive association between the measured BC particle area and the percentage of residential areas in 1000 m around the residential address that was almost significant indicating that the more urban the location the higher the indoor BC (estimate=4.83x10^4 µm^2, p=0.05). A study performed in nursing schools also found that at an urban location, the indoor PM concentrations where higher (45). Furthermore, a study of Bressi *et al.* showed

that $PM_{2.5}$ concentrations were higher in urban and suburban sites, which could be explained by local anthropogenic sources (40). In contrast, another study in Poland on differences in air pollution between urban and rural sites found that PM concentrations were higher at rural sites and not at urban sites for both indoor and outdoor air (46). The differences between the studies can be explained by the fact that both traffic and anthropogenic sources influence the concentrations as seen in the study of Bressi *et al.* (40). The contrasting study did their measurements in the winter season and during the winter season in rural Poland low-quality coal and biomass is used for heating (46). The non-significant result in our study is most likely due to the low statistical power (only data from 33 participants available) since the association with traffic was also not significant.

The model with the agricultural areas showed a negative association (estimate=-3.77×10^4 μm^2, p=0.17) and the model with the natural areas showed a positive association (estimate=4.02×10^3 μm^2, p=0.86) between the measured BC particle area and the percentage of agricultural or natural areas around the residential address with both being not significant. The negative trend from the agricultural model is expected since rural areas, which have more agricultural areas near, have lower BC concentrations since there are fewer anthropogenic and traffic sources present (40, 45). Furthermore, a study performed by Pavilonis *et al.* showed that PM_{10} did not increase in agricultural areas during the harvest season indicating that difference in PM concentrations with residential areas are mostly due to traffic density and anthropogenic sources (47). Also, the higher vegetation levels in rural areas can influence the PM and BC concentrations since they can act as physical barriers between home and road, and can capture significant amounts of PM and BC (23, 29). The agricultural areas indicate how rural the participants are living and act as the opposite of the residential areas indicating the degree of urbanicity. This can explain the nearly significant negative association since the residential areas showed a nearly significant positive association. The positive association of the natural areas might be influenced by the low presence of natural area (average= 12.22%) near the participants' locations.

The outdoor air pollution models were not significant and the NO_2 and PM_{10} models gave a negative association while a positive association is expected (BC: estimate=1.79×10^3 μm^2, p= 0.93; NO_2: estimate=-1.92×10^4 μm^2, p=0.39; PM_{10}: estimate=-1.02×10^4 μm^2, p=0.48; $PM_{2.5}$: estimate=7.12×10^3 μm^2, p=0.77) However, other studies have also found a non-significant result for the association between indoor and outdoor air pollution (42, 48, 49). The study of Mohammadyan *et al.* did not find a significant association between indoor and outdoor concentrations for each PM fraction (<10 µm, <2.5 µm, <1.0 µm), of which also some negative associations were seen (49). However, another study on indoor and outdoor air pollution found that the outdoor concentration could explain the variability of indoor BC, which was mostly due to traffic (42). Another explanation for the results is that the indoor

concentrations are not entirely determined by the outdoor concentrations, but also have sources indoors, such as stoves, smoking, cooking and cleaning, that need to be taken into account (2, 50). This means that the participants that have a stove, or who smoke indoors accumulate higher concentrations of indoor BC even though the outdoor concentrations might be lower than for the other participants. Furthermore, the outdoor air pollution concentrations used in this study were not measured at the residential address of the participants but determined with an air pollution grid based on annual mean air pollution levels from 2015. Furthermore, the outdoor air pollution levels were from 2015 while the study was performed in 2017 and 2018 which could have influenced the results, but the low statistical power is a more plausible cause for the non-significant results.

Second, the microbiome on the leaves of the *Hedera helix* was determined to investigate the association between microbial diversity and indoor BC particle area. The most abundant genera found were the *Burkholderia-Paraburkholderia, Enterobacter, Gilliamella, Pantoea,* and *Pseudomonas* genera and the most abundant phyla were the *Proteobacteria, Firmicutes,* and *Actinobacteria*. The *Burkholderia* genus is commonly found in peat soils and is known for including species with nitrogen fixating abilities (51, 52). The *Enterobacter* genus includes known human pathogens and the genus is known to be found in the phyllosphere of plants (53, 54). The *Pantoea* and *Pseudomonas* genera are common epiphytes, and especially *Pseudomonas* is found in high abundances in the phyllosphere (54, 55). A study by Smets *et al.,* who investigated the effect of an urban environment on the leaf microbiome of *Hedera helix* plants, found different most abundant taxa (37). It has to be taken into account that they used wild *H. helix* species found outdoors, while in this study cultivated ivy was used in an indoor environment. A study investigating the microbial differences between cultivated and wild blueberry plants showed that the species richness and community composition of the phyllosphere from the plants differed between wild and cultivated species (56). Moreover, Smets *et al.* sampled the leaves in February, winter, while this study sampled the leaves during spring, which can also explain the difference in most abundant taxa. The microbiome changes between the different seasons of the year (38, 57).

The association between the BC particle area on the leaves and the number of reads counted during sequencing was negative but not significant (r=-0.21, p=0.64). The alpha-diversity measures, Chao1, ASV richness, Simpson's index and Shannon-Wiener index, also had a negative but non-significant association (Chao1: r=-0.39, p=0.38, ASV richness: r=-0.39, p=0.38, Shannon-Wiener: r=-0.14, p=0.76; Simpson: r=-0.07, p=0.88). The non-significant results are in line with the study of Smets *et al.* where they found no significant difference in taxa richness and the Shannon index, p-values of 0.54 and 0.67 respectively (37). Another study on the effect of urban air pollution on the phyllosphere microbiome

also found no significant difference in alpha-diversity between urban and rural sites (58). Furthermore, it is in line with a study on BC exposure and the lung microbiome in which no significant difference in alpha-diversity was seen between high and low particulate groups (36). The non-significance of our results can be explained by the low statistical power obtained (7 participants), although the other studies indicate that there are no differences in alpha-diversity to be detected.

The PERMANOVA analysis of the beta-diversity (Bray-Curtis dissimilarity matrix) showed that the variance in bacterial composition was not significantly different (p=0.59) between the groups of indoor BC levels but a large portion of the variance could be explained by the BC groups (R^2=0.48). Furthermore, the study of Smets *et al.* has shown a significant difference (p= <0.001) in beta-diversity between urban and non-urban locations, but they have a weaker association (R^2=0.38) (37). Another study like the previous one also found a significant difference between the location of plant species and the beta-diversity with p=0.001 (58). The study performed by Rylance *et al.* did not find a significant difference in beta-diversity between the high and low BC particle groups at both genus and phylum level (p=0.209 and p=0.397, respectively). They did find significant differences in relative abundances between several selected genera and some genera increase while others decrease with higher BC particle levels. This indicates that, although it was not significant in diversity measures, there is an effect of BC on the microbiome. The reason that the results found in this study were not significant is most likely the very low statistical power (7 participants) while there might be a strong association.

5 Conclusion and outlook

With the low number of measured samples, it cannot be concluded that the indoor biomonitoring system using the Hedera helix can be used as an alternative for the aethalometer, nor that there is an association between the bacterial diversity and the indoor BC. However, this study has shown that the BC particle area found on the leaves of the *Hedera helix* can give an indication about the indoor BC levels when taking the urbanicity into account making it a possible alternative for aethalometer. And even though there was no significant association with the microbial diversity, the results show that it is promising to further investigate this. The major limitation in this study was the lack of statistical power, but not all samples collected for this study were processed which indicates that when the other samples are processed significant results can be found. This study is also limited in the information available on ventilation and indoor sources available, and the lack of a continuous aethalometer measurement during the accumulation period of the plant.

In future research, the power will increase with the data from all the samples that could not be processed and analyzed during this project and will allow to detect smaller differences and to more clearly elucidate the associations investigated. A future experiment should also include a continuous aethalometer measurement during the period the plant is accumulating BC and a questionnaire to get detailed information on the timing, amount and type of ventilation, and on the presence and use of indoor sources for BC. Also, more focus should be on the association between the indoor BC and changes in microbial diversity since this is a topic that is not yet studied intensively.

References

1. Bell ML, Davis DL. Reassessment of the lethal London fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. Environmental health perspectives. 2001;109 Suppl 3:389-94.
2. Nazaroff WW. Indoor particle dynamics. Indoor air. 2004;14 Suppl 7:175-83.
3. Viegi G, Simoni M, Scognamiglio A, Baldacci S, Pistelli F, Carrozzi L, et al. Indoor air pollution and airway disease. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2004;8(12):1401-15.
4. WHO. Ambient air pollution: a global assessment of exposure and burden of disease. Geneva, Switzerland: World Health Organization; 2016.
5. Li Z, Wen Q, Zhang R. Sources, health effects and control strategies of indoor fine particulate matter (PM2.5): A review. The Science of the total environment. 2017;586:610-22.
6. Mitchell CS, Zhang JJ, Sigsgaard T, Jantunen M, Lioy PJ, Samson R, et al. Current state of the science: health effects and indoor environmental quality. Environmental health perspectives. 2007;115(6):958-64.
7. Luben TJ, Nichols JL, Dutton SJ, Kirrane E, Owens EO, Datko-Williams L, et al. A systematic review of cardiovascular emergency department visits, hospital admissions and mortality associated with ambient black carbon. Environment international. 2017;107:154-62.
8. Ni M, Huang J, Lu S, Li X, Yan J, Cen K. A review on black carbon emissions, worldwide and in China. Chemosphere. 2014;107:83-93.
9. Janssen NA, Hoek G, Simic-Lawson M, Fischer P, van Bree L, ten Brink H, et al. Black carbon as an additional indicator of the adverse health effects of airborne particles compared with PM10 and PM2.5. Environmental health perspectives. 2011;119(12):1691-9.
10. Long CM, Nascarella MA, Valberg PA. Carbon black vs. black carbon and other airborne materials containing elemental carbon: physical and chemical distinctions. Environmental pollution (Barking, Essex : 1987). 2013;181:271-86.
11. Grahame TJ, Klemm R, Schlesinger RB. Public health and components of particulate matter: the changing assessment of black carbon. Journal of the Air & Waste Management Association (1995). 2014;64(6):620-60.
12. WHO. Household air pollution and health Geneva, Switzerland: World Health Organization; 2018 [updated 8 May 2018. Available from: https://www.who.int/en/news-room/fact-sheets/detail/household-air-pollution-and-health.
13. Niranjan R, Thakur AK. The Toxicological Mechanisms of Environmental Soot (Black Carbon) and Carbon Black: Focus on Oxidative Stress and Inflammatory Pathways. Frontiers in immunology. 2017;8:763.
14. Nichols JL, Owens EO, Dutton SJ, Luben TJ. Systematic review of the effects of black carbon on cardiovascular disease among individuals with pre-existing disease. International journal of public health. 2013;58(5):707-24.
15. Elmes M, Gasparon M. Sampling and single particle analysis for the chemical characterisation of fine atmospheric particulates: A review. Journal of environmental management. 2017;202(Pt 1):137-50.
16. Morawska L, Ayoko GA, Bae GN, Buonanno G, Chao CYH, Clifford S, et al. Airborne particles in indoor environment of homes, schools, offices and aged care facilities: The main routes of exposure. Environment international. 2017;108:75-83.
17. Dales R, Liu L, Wheeler AJ, Gilbert NL. Quality of indoor residential air and health. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2008;179(2):147-52.
18. Zhang J, Smith KR. Indoor air pollution: a global health concern. British medical bulletin. 2003;68:209-25.
19. Airthinx. 2019 [Webshop for air quality monitors]. Available from: https://airthinx.io/.

20. Ecomesure. 2019 [Webshop for air quality monitors]. Available from: https://eshop.ecomesure.com/index.php.
21. Aeroqual. 2019 [Webshop for air quality monitors]. Available from: https://www.aeroqual.com/.
22. Vazquez S, Martin A, Garcia M, Espanol C, Navarro E. Metal uptake of Nerium oleander from aerial and underground organs and its use as a biomonitoring tool for airborne metallic pollution in cities. Environmental science and pollution research international. 2016;23(8):7582-94.
23. Zhang WK, Wang B, Niu X. Study on the Adsorption Capacities for Airborne Particulates of Landscape Plants in Different Polluted Regions in Beijing (China). International journal of environmental research and public health. 2015;12(8):9623-38.
24. Xie C, Kan L, Guo J, Jin S, Li Z, Chen D, et al. A dynamic processes study of PM retention by trees under different wind conditions. Environmental pollution (Barking, Essex : 1987). 2018;233:315-22.
25. Saleh IH, Abdel-Halim AA. Cosmogenic beryllium-7 in soil, rainwater and selected plant species to evaluate the vegetal interception of atmospheric fine particulate matter. Isotopes in environmental and health studies. 2018;54(4):392-402.
26. Dzierzanowski K, Popek R, Gawronska H, Saebo A, Gawronski SW. Deposition of particulate matter of different size fractions on leaf surfaces and in waxes of urban forest species. International journal of phytoremediation. 2011;13(10):1037-46.
27. Przybysz A, Saebo A, Hanslin HM, Gawronski SW. Accumulation of particulate matter and trace elements on vegetation as affected by pollution level, rainfall and the passage of time. The Science of the total environment. 2014;481:360-9.
28. Castanheiro A, Samson R, De Wael K. Magnetic- and particle-based techniques to investigate metal deposition on urban green. The Science of the total environment. 2016;571:594-602.
29. Sternberg T, Viles H, Cathersides A, Edwards M. Dust particulate absorption by ivy (Hedera helix L) on historic walls in urban environments. The Science of the total environment. 2010;409(1):162-8.
30. Baldini E, Facini O, Nerozzi F, Rossi F, Rotondi A. Leaf characteristics and optical properties of different woody species. Trees. 1997;12(2):73-81.
31. METCALFE DJ. Hedera helix L. Journal of Ecology. 2005;93(3):632-48.
32. Bové H, Steuwe C, Fron E, Slenders E, D'Haen J, Fujita Y, et al. Biocompatible Label-Free Detection of Carbon Black Particles by Femtosecond Pulsed Laser Microscopy. Nano Letters. 2016;16(5):3173-8.
33. Yang B, Xiao C. PM2.5 exposure significantly improves the exacerbation of A549 tumor-bearing CB17-SCID mice. Environmental toxicology and pharmacology. 2018;60:169-75.
34. Li N, He F, Liao B, Zhou Y, Li B, Ran P. Exposure to ambient particulate matter alters the microbial composition and induces immune changes in rat lung. Respiratory research. 2017;18(1):143.
35. Wang L, Cheng H, Wang D, Zhao B, Zhang J, Cheng L, et al. Airway microbiome is associated with respiratory functions and responses to ambient particulate matter exposure. Ecotoxicology and environmental safety. 2019;167:269-77.
36. Rylance J, Kankwatira A, Nelson DE, Toh E, Day RB, Lin H, et al. Household air pollution and the lung microbiome of healthy adults in Malawi: a cross-sectional study. BMC microbiology. 2016;16(1):182.
37. Smets W, Wuyts K, Oerlemans E, Wuyts S, Denys S, Samson R, et al. Impact of urban land use on the bacterial phyllosphere of ivy (Hedera sp.). Atmospheric environment. 2016;147:376-83.
38. Weikl F, Tischer C, Probst AJ, Heinrich J, Markevych I, Jochner S, et al. Fungal and Bacterial Communities in Indoor Dust Follow Different Environmental Determinants. PloS one. 2016;11(4):e0154131.
39. Baldacchini C, Castanheiro A, Maghakyan N, Sgrigna G, Verhelst J, Alonso R, et al. How Does the Amount and Composition of PM Deposited on Platanus acerifolia Leaves Change Across Different Cities in Europe? Environmental science & technology. 2017;51(3):1147-56.

40. Bressi M, Sciare J, Ghersi V, Bonnaire N, Nicolas J, Petit J-E, et al. A one-year comprehensive chemical characterisation of fine aerosol (PM 2.5) at urban, suburban and rural background sites in the region of Paris (France). Atmospheric Chemistry and Physics. 2013;13(15):7825-44.
41. Huang S, Lawrence J, Kang CM, Li J, Martins M, Vokonas P, et al. Road proximity influences indoor exposures to ambient fine particle mass and components. Environmental pollution (Barking, Essex : 1987). 2018;243(Pt B):978-87.
42. Baxter LK, Clougherty JE, Paciorek CJ, Wright RJ, Levy JI. Predicting residential indoor concentrations of nitrogen dioxide, fine particulate matter, and elemental carbon using questionnaire and geographic information system based data. Atmospheric environment (Oxford, England : 1994). 2007;41(31):6561-71.
43. . !!! INVALID CITATION !!! {MacNeill, 2016 #83;Nazaroff, 2004 #48}.
44. MacNeill M, Dobbin N, St-Jean M, Wallace L, Marro L, Shin T, et al. Can changing the timing of outdoor air intake reduce indoor concentrations of traffic-related pollutants in schools? Indoor air. 2016;26(5):687-701.
45. Nunes RA, Branco PT, Alvim-Ferraz MC, Martins FG, Sousa SI. Particulate matter in rural and urban nursery schools in Portugal. Environmental pollution (Barking, Essex : 1987). 2015;202:7-16.
46. Mainka A, Zajusz-Zubek E. Indoor Air Quality in Urban and Rural Preschools in Upper Silesia, Poland: Particulate Matter and Carbon Dioxide. International journal of environmental research and public health. 2015;12(7):7697-711.
47. Pavilonis BT, Anthony TR, T O'Shaughnessy P, Humann MJ, Merchant JA, Moore G, et al. Indoor and outdoor particulate matter and endotoxin concentrations in an intensely agricultural county. Journal of Exposure Science and Environmental Epidemiology. 2013;23(3):299.
48. Challoner A, Pilla F, Gill L. Prediction of Indoor Air Exposure from Outdoor Air Quality Using an Artificial Neural Network Model for Inner City Commercial Buildings. International journal of environmental research and public health. 2015;12(12):15233-53.
49. Mohammadyan M, Ghoochani M, Kloog I, Abdul-Wahab SA, Yetilmezsoy K, Heibati B, et al. Assessment of indoor and outdoor particulate air pollution at an urban background site in Iran. Environmental monitoring and assessment. 2017;189(5):235.
50. Clougherty JE, Houseman EA, Levy JI. Source apportionment of indoor residential fine particulate matter using land use regression and constrained factor analysis. Indoor air. 2011;21(1):53-66.
51. Nie Y, Li L, Wang M, Tahvanainen T, Hashidoko Y. Nitrous oxide emission potentials of Burkholderia species isolated from the leaves of a boreal peat moss Sphagnum fuscum. Bioscience, Biotechnology, and Biochemistry. 2015;79(12):2086-95.
52. Sun H, Terhonen E, Koskinen K, Paulin L, Kasanen R, Asiegbu FO. Bacterial diversity and community structure along different peat soils in boreal forest. Applied soil ecology. 2014;74:37-45.
53. Aydogan EL, Moser G, Muller C, Kampfer P, Glaeser SP. Long-Term Warming Shifts the Composition of Bacterial Communities in the Phyllosphere of Galium album in a Permanent Grassland Field-Experiment. Front Microbiol. 2018;9:144.
54. Thapa S, Prasanna R. Prospecting the characteristics and significance of the phyllosphere microbiome. Annals of Microbiology. 2018;68(5):229-45.
55. Vorholt JA. Microbial life in the phyllosphere. Nature reviews Microbiology. 2012;10(12):828-40.
56. Stanwood JM, Dighton J. Seasonality and management, not proximity to highway, affect species richness and community composition of epiphytic phylloplane fungi found on (wild and cultivated) Vaccinium spp. Fungal ecology. 2011;4(4):277-83.
57. Toju H, Kurokawa H, Kenta T. Factors Influencing Leaf- and Root-Associated Communities of Bacteria and Fungi Across 33 Plant Orders in a Grassland. Front Microbiol. 2019;10:14.
58. Espenshade J, Thijs S, Gawronski S, Boye H, Weyens N, Vangronsveld J. Influence of Urbanization on Epiphytic Bacterial Communities of the Platanus x hispanica Tree Leaves in a Biennial Study. Front Microbiol. 2019;10.

Supplement

In this section, supplementary information is given. This can either be supplementary materials and methods, which contains standardized protocols, or supplementary data, which contains figures and tables.

Attachment 1: Supplemental protocols

In this section, the standardized protocols of the techniques mentioned in the materials and methods section are summed up.

Supplemental protocol 1: Isolation of bacteria from the phyllosphere

1. Prepare a 10X sampling buffer. Heating can be necessary to dissolve all the components. Do not adjust pH (should be pH 6.65 ± 0.20). Sterilize the solution.
2. Before sampling, prepare 1X sampling buffer by diluting with sterile distilled water. Adjust to pH 7.00 with HCl.
3. Add 30 mL 1X sampling buffer to sterile 50 mL tubes (max. 28/run). Weigh the tubes.
4. In the field, transfer leaves into the tubes (2.0 ± 0.5 g leaf material/tube) using EtOH-sterilized forceps (take 70% EtOH solution with you). Back in the lab, weigh the tubes with leaves. Store at 4°C when not proceeding immediately with the following step.
5. Within 3 days from sampling, wash microbial cells from the leaf surface by shaking the tubes horizontally on an orbital shaker (200 rpm, 3 hours). (While shaking, sterilize, label and weigh all needed material if necessary.)
6. Centrifuge the 50 mL tubes for 10 min at 1000 rpm (RT).
7. Take out the leaves using sterilized forceps.
8. Centrifuge the 50 mL tubes for 10 min at 3000 rpm (RT)
9. Carefully remove the supernatant with a 50 mL pipette until approx. 3 mL supernatant and the pellet remain (do not agitate the pellet).
10. Resuspend and pool the suspensions (1000 µL pipette) from 4 tubes in sterile 15 mL tubes.
11. Centrifuge the 15 mL tubes for 10 min at 3000 rpm (RT)
12. Carefully remove the supernatant with a 1000 µL pipette until < 1 mL supernatant and the pellet remain (do not agitate the pellet). Resuspend and transfer the suspensions into sterilized, pre-weighed, 1.5 mL tubes.
13. Centrifuge the 1.5 mL tubes for 10 min at 3000 rpm (RT)
14. Completely remove the supernatant with a 1000 µL pipette. Weigh and store the 1.5 mL tubes at -20°C prior to DNA isolation.

Supplemental protocol 2: Bacterial screening

1. During the second last step of the 'isolation of microorganisms' protocol, take 10µL from the bacterial suspension and transfer to 120µL 10 mM $MgSO_4$.
2. Transfer 50µL of this suspension to 1/10 869 plates, use cell spreader.
3. Incubate plates at 30°C for 4 days (store them upside down).
4. After 4 days, count the number of colonies per plate for each species of bacteria.
5. Purification: see supplemental protocol 3
6. Glycerol stock: see supplemental protocol 3

Supplemental protocol 3: Purification of bacteria

1. Put 10 drops of 50 µl sterile 10 mM $MgSO_4$ on a sterile Petri dish and one big drop in the middle to cool the inoculations needle
2. Pick bacterial colonies with a sterile toothpick and dissolve in the drops of $MgSO_4$
3. Transfer this solution with the inoculation needle on a plate of rich medium (869 medium), divide the plate in 4
4. Incubate at 30°C for 5 days
5. Check if colonies are pure, if not repeat previous steps
6. Fill a 15ml falcon tube with 5 ml liquid 869 medium
7. Pick pure colonies of the plate with a sterile toothpick and break the toothpick in the 15ml tube
8. Incubate at 30°C for 3 days
9. Centrifuge tubes for 15 min at 4000rpm
10. Discard the supernatant and dissolve the pellet in 1.5 ml glycerol
11. Vortex and transfer suspension in a cryotube
12. Store in the -45°C freezer

Supplemental protocol 4: DNA isolation of bacteria from the phyllosphere with the NucleoSpin Soil kit

Before starting the preparation:

Check Lysis Buffer SL1 or SL2 for precipitated SDS. Dissolve any precipitate by incubating the buffer at 30–40 °C for 10 min and shaking the bottle every 2 min.

Procedure:

1. Prepare sample
   - Transfer 250–500 mg fresh sample material to a NucleoSpinR Bead Tube Type A containing the ceramic beads.
   - Important: Do not fill the tube higher than the 1 mL mark.
   - Add 700 µL Buffer SL1 or Buffer SL2.
   - Note for very dry material: If the sample material soaks up too much lysis buffer, fill the NucleoSpinR Bead Tube Type A up to the 1.5 mL mark with fresh lysis buffer.
   - Note for very wet material: Remove excess liquid before the addition of lysis buffer, if necessary after spinning down the sample.
2. Adjust lysis conditions
   - Add 150 µL Enhancer SX and close the cap.
   - Note: Enhancer SX ensures the highest possible DNA yield. It can, however, also promote the release of humic acids.
3. Sample lysis
   - Attach the NucleoSpin® Bead Tubes horizontally to a vortexer, for example, by taping or using a special adapter.
   - Vortex the samples at full speed and room temperature (18–25 °C) for 5 min.
4. Precipitate contaminants
   - Centrifuge for 2 min at 11,000 x *g* to eliminate the foam caused by the detergent.
   - Note: The clear supernatant can be transferred to a new collection tube (not provided) prior to the following precipitation. This might result in more consistent yields from prep to prep and is highly recommended for carbonate containing samples.

- Add 150 µL Buffer SL3 and vortex for 5 s.
- Incubate for 5 min at 0–4 °C.
- Centrifuge for 1 min at 11,000 x *g*.

5. Filter lysate
   - Place a NucleoSpinR Inhibitor Removal Column (red ring) in a Collection Tube (2 mL, lid).
   - Load up to 700 µL clear supernatant of step 4 onto the filter.
   - Centrifuge for 1 min at 11,000 x *g*.
   - Note: With very wet samples (e.g., sediments) the volume of clear supernatant of step 4 can exceed 700 µL significantly. In this case, transfer the NucleoSpinR Inhibitor Removal Column to a new collection tube (not provided) and load the remaining supernatant.
   - Centrifuge for 1 min at 11,000 x g.
   - Combine the flow-throughs.
   - Discard the NucleoSpin® Inhibitor Removal Column.
   - If a pellet is visible in the flow through, transfer the clear supernatant to a new collection tube (not provided).
6. Adjust binding conditions
   - Add 250 µL Buffer SB and close the lid.
   - Vortex for 5 s.
7. Bind DNA
   - Place a NucleoSpinR Soil Column (green ring) in a Collection Tube (2 mL).
   - Load 550 µL sample onto the column.
   - Centrifuge for 1 min at 11,000 x *g*.
   - Discard flow through and place the column back into the collection tube.
   - Load the remaining sample onto the column.
   - Centrifuge for 1 min at 11,000 x *g*.
   - Discard flow through and place the column back into the collection tube.
8. Wash and dry silica membrane

   First wash:
   - Add 500 µL Buffer SB to the NucleoSpin® Soil Column.
   - Centrifuge for 30 s at 11,000 x *g*.
   - Discard flow through and place the column back into the collection tube.

   Second wash:
   - Add 550 µL Buffer SW1 to the NucleoSpin® Soil Column.
   - Centrifuge for 30 s at 11,000 x *g*.
   - Discard flow through and place the column back into the collection tube.

   Third wash:
   - Add 700 µL Buffer SW2 to the NucleoSpin® Soil Column.
   - Close the lid and vortex for 2 s. Centrifuge for 30 s at 11,000 x *g*. Discard flow through and place the column back into the collection tube.

   Fourth wash:
   - Add 700 µL Buffer SW2 to the NucleoSpin® Soil Column.
   - Close the lid and vortex for 2 s. Centrifuge for 30 s at 11,000 x g. Discard flow through and place the column back into the collection tube.
   - Note: The same collection tube is used throughout the entire washing procedure to reduce plastic waste.
9. Dry silica membrane
   - Centrifuge for 2 min at 11,000 x *g*.
   - If for any reason, the liquid in the collection tube has touched the NucleoSpinR Soil Column after the drying step, discard flow through and centrifuge again.
10. Elute DNA

- Place the NucleoSpin® Soil Column into a new microcentrifuge tube (not provided).
- Add 30 µL (for high concentration), 50 µL (for medium concentration and yield), or 100 µL (for high yield) Buffer SE to the column.
- Do not close the lid and incubate for 1 min at room temperature (18–25 °C). Close the lid and centrifuge for 30 s at 11,000 x *g*.
- Note: Quantify DNA not only by UV-VIS but also run an agarose gel to verify yield and DNA quality.

Supplemental protocol 5: Sample library preparation with Nextera XT DNA Library Preparation and Index Set A kits

Amplicon PCR:

1. Set up the mastermix for amplicon PCR (Supplemental table 4).
2. Seal the plate and perform PCR in a thermal cycler using the program described in Supplemental table 5.
3. [Optional] Run 1 µl of the PCR product on a Bioanalyzer DNA 1000 chip to verify the size. Using the V3 and V4 primer pairs in the protocol, the expected size on a Bioanalyzer trace after the Amplicon PCR step is ~550 base pairs.

PCR clean-up:

1. Centrifuge the Amplicon PCR plate at 1,000 × g at 20°C for 1 minute to collect condensation, carefully remove the seal.
2. [Optional - for use with a shaker for mixing] Using a multichannel pipette set to 25 µl, transfer the entire Amplicon PCR product from the PCR plate to the MIDI plate. Change tips between samples.

NOTE
Transfer the sample to a 96-well MIDI plate if planning to use a shaker for mixing. If mixing by pipette, the sample can remain in the 96-well PCR plate.

3. Vortex the AMPure XP beads for 30 seconds to make sure that the beads are evenly dispersed. Add an appropriate volume of beads to a trough depending on the number of samples processing.
4. Using a multichannel pipette, add 20 µl of AMPure XP beads to each well of the Amplicon PCR plate. Change tips between columns.
5. Gently pipette entire volume up and down 10 times if using a 96-well PCR plate or seal plate and shake at 1800 rpm for 2 minutes if using a MIDI plate.
6. Incubate at room temperature without shaking for 5 minutes.
7. Place the plate on a magnetic stand for 2 minutes or until the supernatant has cleared.
8. With the Amplicon PCR plate on the magnetic stand, use a multichannel pipette to remove and discard the supernatant. Change tips between samples.
9. With the Amplicon PCR plate on the magnetic stand, wash the beads with freshly prepared 80% ethanol as follows:
    - Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.
    - Incubate the plate on the magnetic stand for 30 seconds.
    - Carefully remove and discard the supernatant.
10. With the Amplicon PCR plate on the magnetic stand, perform a second ethanol wash as follows:

- Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.
- Incubate the plate on the magnetic stand for 30 seconds.
- Carefully remove and discard the supernatant.
- Use a P20 multichannel pipette with fine pipette tips to remove excess ethanol.

11. With the Amplicon PCR plate still on the magnetic stand, allow the beads to air-dry for 10 minutes.
12. Remove the Amplicon PCR plate from the magnetic stand. Using a multichannel pipette, add 52.5 µl of 10 mM Tris pH 8.5 to each well of the Amplicon PCR plate.
13. Gently pipette mix up and down 10 times, changing tips after each column (or seal plate and shake at 1800 rpm for 2 minutes). Make sure that beads are fully resuspended.
14. Incubate at room temperature for 2 minutes.
15. Place the plate on the magnetic stand for 2 minutes or until the supernatant has cleared.
16. Using a multichannel pipette, carefully transfer 50 µl of the supernatant from the Amplicon PCR plate to a new 96-well PCR plate. Change tips between samples to avoid cross-contamination.

Index PCR:

1. Using a multichannel pipette, transfer 5 µl from each well to a new 96-well plate. The remaining 45 µl is not used in the protocol and can be stored for other uses.
2. Arrange the Index 1 and 2 primers in a rack (i.e. the TruSeq Index Plate Fixture) using the following arrangements as needed:
    - Arrange Index 2 primer tubes (white caps, clear solution) vertically, aligned with rows A through H.
    - Arrange Index 1 primer tubes (orange caps, yellow solution) horizontally, aligned with columns 1 through 12.
3. Place the 96-well PCR plate with the 5 µl of resuspended PCR product DNA in the TruSeq Index Plate Fixture.
4. Set up the mastermix for index PCR (Supplemental table 6).
5. Gently pipette up and down 10 times to mix.
6. Cover the plate with Microseal 'A'.
7. Centrifuge the plate at 1,000 × g at 20°C for 1 minute.
8. Perform PCR on a thermal cycler using the program described in Supplemental table 7.

PCR clean-up 2:

1. Centrifuge the Index PCR plate at 280 × g at 20°C for 1 minute to collect condensation.
2. [Optional - for use with shaker for mixing] Using a multichannel pipette set to 50 µl, transfer the entire Index PCR product from the PCR plate to the MIDI plate. Change tips between samples.

NOTE
Transfer the sample to a 96-well MIDI plate if planning to use a shaker for mixing. If mixing by pipette, the sample can remain in the 96-well PCR plate.

3. Vortex the AMPure XP beads for 30 seconds to make sure that the beads are evenly dispersed. Add an appropriate volume of beads to a trough.
4. Using a multichannel pipette, add 56 µl of AMPure XP beads to each well of the Index PCR plate.

5. Gently pipette mix up and down 10 times if using a 96-well PCR plate or seal plate and shake at 1800 rpm for 2 minutes if using a MIDI plate.
6. Incubate at room temperature without shaking for 5 minutes.
7. Place the plate on a magnetic stand for 2 minutes or until the supernatant has cleared.
8. With the Index PCR plate on the magnetic stand, use a multichannel pipette to remove and discard the supernatant. Change tips between samples.
9. With the Index PCR plate on the magnetic stand, wash the beads with freshly prepared 80% ethanol as follows:
   - Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.
   - Incubate the plate on the magnetic stand for 30 seconds.
   - Carefully remove and discard the supernatant.
10. With the Index PCR plate on the magnetic stand, perform a second ethanol wash as follows:
    - Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.
    - Incubate the plate on the magnetic stand for 30 seconds.
    - Carefully remove and discard the supernatant.
    - Use a P20 multichannel pipette with fine pipette tips to remove excess ethanol.
11. With the Index PCR plate still on the magnetic stand, allow the beads to air-dry for 10 minutes.
12. Remove the Index PCR plate from the magnetic stand. Using a multichannel pipette, add 27.5 µl of 10 mM Tris pH 8.5 to each well of the Index PCR plate.
13. If using a 96-well PCR plate, gently pipette mix up and down 10 times until beads are fully resuspended, changing tips after each column. If using a MIDI plate, seal plate and shake at 1800 rpm for 2 minutes.
14. Incubate at room temperature for 2 minutes.
15. Place the plate on the magnetic stand for 2 minutes or until the supernatant has cleared.
16. Using a multichannel pipette, carefully transfer 25 µl of the supernatant from the Index PCR plate to a new 96-well PCR plate. Change tips between samples to avoid cross-contamination.

[Optional] validate library:

Run 1 µl of a 1:50 dilution of the final library on a Bioanalyzer DNA 1000 chip to verify the size. Using the V3 and V4 primer pairs in the protocol, the expected size on a Bioanalyzer trace of the final library is ~630 bp.

Library quantification, normalization, and pooling:

Illumina recommends quantifying your libraries using a fluorometric quantification method that uses dsDNA binding dyes. Calculate DNA concentration in nM, based on the size of DNA amplicons as determined by an Agilent Technologies 2100 Bioanalyzer trace:

$$\frac{\text{Concentration in } \frac{\text{ng}}{\mu\text{l}}}{660\,\frac{\text{g}}{\text{mol}}\ \text{X average library size}}\ \text{X}\ 10^{\square} = \text{concentration in nM}$$

For example:

$$\frac{15\,\frac{\text{ng}}{\mu\text{l}}}{(660\,\frac{\text{g}}{\text{mol}}\ \text{X}\ 500)}\ \text{X}\ 10^{\square} =\ 45\ \text{nM}$$

Dilute concentrated final library using Resuspension Buffer (RSB) or 10 mM Tris pH 8.5 to 4 nM. Aliquot 5 µl of diluted DNA from each library and mix aliquots for pooling libraries with unique indices. Depending on coverage needs, up to 96 libraries can be pooled for one MiSeq run.

For metagenomics samples, >100,000 reads per sample is sufficient to fully survey the bacterial composition. This number of reads allows for sample pooling to the maximum level of 96 libraries, given the MiSeq output of > 20 million reads.

Supplemental protocol 6: 16S rRNA gene sequencing with Phix Control and MiSeq Reagent kits

Denature DNA:

1 Combine the following volumes of pooled final DNA library and freshly diluted 0.2 N NaOH in a microcentrifuge tube:
   - nM pooled library (5 µl)
   - 0.2 N NaOH (5 µl)
2 Set aside the remaining dilution of 0.2 N NaOH to prepare a PhiX control within the next 12 hours.
3 Vortex briefly to mix the sample solution, and then centrifuge the sample solution at 280 × g at 20°C for 1 minute.
4 Incubate for 5 minutes at room temperature to denature the DNA into single strands.
5 Add the following volume of pre-chilled HT1 to the tube containing denatured DNA:
   - Denatured DNA (10 µl)
   - Pre-chilled HT1 (990 µl)
   Adding the HT1 results in a 20 pM denatured library in 1 mM NaOH.
6 Place the denatured DNA on ice until you are ready to proceed to final dilution.

Dilute Denatured DNA:

1. Dilute the denatured DNA to the desired concentration using the following example:

NOTE
Illumina recommends targeting 800–1000 K/mm² raw cluster densities using MiSeq v3 reagents. It is suggested to start your first run using a 4 pM loading concentration and adjust subsequent runs appropriately.

Final concentration	2 pM	4 pM	6 pM	8 pM	10 pM
20 pM denaturated library	60 µl	120 µl	180 µl	240 µl	300 µl
Pre-chilled HT1	540 µl	480 µl	420 µl	360 µl	300 µl

2. Invert several times to mix and then pulse centrifuge the DNA solution.
3. Place the denatured and diluted DNA on ice.

Denature and Dilution of PhiX Control:

Use the following instructions to denature and dilute the 10 nM PhiX library to the same loading concentration as the Amplicon library. The final library mixture must contain at least 5% PhiX.

1. Combine the following volumes to dilute the PhiX library to 4 nM:
   - 10 nM PhiX library (2 µl)
   - 10 mM Tris pH 8.5 (3 µl)
2. Combine the following volumes of 4 nM PhiX and 0.2 N NaOH in a microcentrifuge tube:
   - 4 nM PhiX library (5 µl)
   - 0.2 N NaOH (5 µl)
3. Vortex briefly to mix the 2 nM PhiX library solution.
4. Incubate for 5 minutes at room temperature to denature the PhiX library into single strands.
5. Add the following volumes of pre-chilled HT1 to the tube containing denatured PhiX library to result in a 20 pM PhiX library:
   - Denatured PhiX library (10 µl)
   - Pre-chilled HT1 (990 µl)
6. Dilute the denatured 20 pM PhiX library to the same loading concentration as the Amplicon library as follows:

Final concentration	2 pM	4 pM	6 pM	8 pM	10 pM
20 pM denaturated library	60 µl	120 µl	180 µl	240 µl	300 µl
Pre-chilled HT1	540 µl	480 µl	420 µl	360 µl	300 µl

7. Invert several times to mix and then pulse centrifuge the DNA solution.
8. Place the denatured and diluted PhiX on ice.

Combine Amplicon Library and PhiX Control:

NOTE
The recommended PhiX control spike-in of ≥ 5% for low diversity libraries is possible with RTA v1.17.28 or later, which is bundled with MCS v2.2. For optimal performance, update to v3 software (MCS 2.3). If you are using an older version of the MiSeq software or sequencing these libraries on the GA or HiSeq, Illumina recommends using ≥ 25% PhiX control spike-in.

1. Combine the following volumes of denatured PhiX control library and your denatured amplicon library in a microcentrifuge tube:
   - Denatured and diluted PhiX control (30 µl)
   - Denatured and diluted amplicon library (570 µl)
2. Set the combined sample library and PhiX control aside on ice until you are ready to heat denature the mixture immediately before loading it onto the MiSeq v3 reagent cartridge.
3. Using a heat block, incubate the combined library and PhiX control tube at 96°C for 2 minutes.

4. After the incubation, invert the tube 1–2 times to mix and immediately place in the ice-water bath.
5. Keep the tube in the ice-water bath for 5 minutes.

NOTE
Perform the heat denaturation step immediately before loading the library into the MiSeq reagent cartridge to ensure efficient template loading on the MiSeq flow cell.

Supplemental protocol 7: Confocal microscopy of black carbon on the leaves

Preparations prior to measurements:

1 control leaf/participant > 3 positions per leaf > 6 Z-stacks per position ➔ 18 images per control leaf
±4 exposure leaves/participant > 3 positions per leaf > 6 z-stacks per position ➔ 72 images per exposure plant

- For each leaf (control and exposure) cut out 3 spots.
- Tape cut-outs on a coverslip, with the side that needs to be measured (backside of the leaf) facing towards the glass.

Measurements:

1. Put on the confocal microscope (LSM510 Meta) by switching on the remote control master (behind microscope).
2. Switch on both computers (switch is on the keyboard of the right computer, under the desk of the left computer) > LSM user
3. Check whether control system is set at 'LSM510' (B), the remote is on the floor behind the laser. If remote is on A, turn the key, then push in B and turn the key again.
4. Open Zen 2009 on the desktop (if correct, should say 'starting LSM510) > start system (if the correct control system is busy, a white light will emerge from the box on the table).
5. Acquisition > laser > MaiTai on
6. Laser > laser properties > gives settings of the laser → should be 810nm, mode-locked. If not, go to lightpad > laser > set at 730 > wait until modelocked > then set again at 810 > wait until modelocked

Before actual measurements, start with measuring the power + rhodamine (do this once in the morning).

Power:

1. In Zen2009: File > open > data> Urine > settings > powermeter > reuse settings of the laser
2. Turn out the 10x objective, place it carefully on the table on a tissue
3. Place the Thorlabs measuring device on the stage
4. Switch on the second computer, password Ellen
5. Open optical powermeter on the desktop
6. Choose 906
7. Settings: wavelength 810, meas config average 1, log config as high as possible
8. Reset/clear
9. Zen2009 Start experiment + Optical powermeter Start log
10. Wait until the line stabilizes (minimal fluctuation, laser is warming up) ➔ ± 60 mW
11. Write this all down in the logbook + save the image

Rhodamine:

1. In Zen2009: File > open > data > urine > settings > rhodamine > reuse settings of the laser
2. Turn the 10x objective back in
3. Stage at zero position: focus buttons on the right side of the microscope, lower button > Zen2009 > focus > show all > manually > Z-position = 0 > then put at 7508
4. Make sure the right objective is indicated: acquisition mode > objective > Plan-Neofluar 10x/0.3
5. 300µL rhodamine in a well > put on the stage
6. Start experiment > 10 images
7. Only use the green channel > histogram > put a rectangle on the picture > mean intensity should be ±1600 ±500 >change transmission to reach this (no higher than 13%, otherwise change the height of table)
8. Channel > 810 > the number is the transmission (% mW of the laser that is passed through)
9. Save the file of the rhodamine measurement, use rhodaminexx/xx/201X

Make sure to note each day:

- The power (±60 mW)
- The mean intensity of rhodamine (1600 ±500)
- The transmission (±7-13%)

Measurements:

1. Open all doors, put the laser up.
2. Choose the correct objective: 10x by clicking on 'Objective' on the right side of the microscope.
3. Place the sample under the microscope (10x/0·3 objective, Plan-Neofluar 10x/0·3, Carl Zeiss). You do NOT need to use immersion oil.
4. Resulting images have a field of view of 896.2µm x 896.2µm, 1.76 pixel size (Acquisition mode > image size, speed of the laser: 8 (not higher, otherwise the image will get blurry)
5. Position the sample under the microscope using the joy-stick
6. Look through the microscope and sharpen the image
7. Place black plate + black cloth against the microscope, so that no light can reach the sample (light will interfere with the laser).
8. Acquisition
   - Z stack
   - continuous
   - down until image is black (turn towards you) -> set first. Go deep enough to make sure you also image the particles who fell off the leaf and are on the glass.
   - up until image is black (turn away from you) -> set last
   - stop
   - click optimal (slices of 6.62µm)
   - start experiment

   ➔ do this for every spot (6 different spots per cut out).
9. Switch position by slightly tapping the joystick
10. Save the image correctly: for example, 530A_4_1.6 = ID 530A, leaf/coverslip number 4, spot number 1, Z-stack number 6
    - Leaf/coverslips: control and 1-4
    - Spot: 1-3
    - Z-stack: 1-6
11. Remove sample after 6 z-stacks per spot are made.
12. Store the sample in a Petri dish and keep away from light.

Attachment 2: Supplemental materials

In this section, the used materials (like medium, master mix, primers, etc.) are described.

Supplemental materials 1: Solutions

Supplemental table 1: *Components of the sampling buffer dissolved in distilled water.*

Components	**10X sampling buffer**	**1X sampling buffer**
NaCl	1370 mM	137 mM
$Na_2HPO_3 . 7H_2O$	100 mM	10 mM
KCl	27 mM	2.7 mM
KH_2PO_4	18 mM	1.8 mM
$Na_2EDTA . 2H_2O$	10 mM	1 mM
Tween 20	1%	0.1%
pH	6.65 ± 0.20	7.00 ± 0.01

EDTA: Ethylenediaminetetraacetic acid

Supplemental table 2: *Weights of components for 869 rich medium specified for each volume.*

		1L	**2L**	**3L**	**5L**	**10L**
Tryptone	(g)	10	20	30	50	100
Yeast extract powder	(g)	5	10	15	25	50
NaCl	(g)	5	10	15	25	50
D-(+)-glucose.H_2O	(g)	1	2	3	5	10
$CaCl_2.2H_2O$	(g)	0.345	0.690	1.035	1.725	3.450
Agar No.2 Bacteriological (if solid medium is desired)	(g)	15				

Supplemental table 3: *Quantities of components in the 15%$_w$ glycerol solution.*

Component	Glycerol (99%)	NaCl	Distilled water
Quantity	75 g	4.25 g	425 ml

Supplemental materials 2: PCR materials

Supplemental table 4: *Amplicon PCR unit for 11 reactions.*

	Volume (µl)
Mastermix	
µl Reaction Buffer (10X) with 18 mM MgCl2	27.5
µl dNTP-mix (10 mM each)	5.5
µl DMSO (100%)	13.75
µl Forward primer (10 µM)	5.5
µl Reverse primer (10 µM)	5.5
µl Enzyme Blend (5 U/µl)	2.2
µl RNase free water	204.05
Total	**264**
MM to distribute per Epp	24.0
DNA-sample	1.0
PCR unit	**25.0**

dNTP: deoxynucleotide; DMSO: dimethyl sulfoxide, MM: mastermix.

Supplemental table 5: *Amplicon PCR cycles*

	No. Cycles	Degrees (°C)	Time (sec)	
Initial denaturation	1	95	180	
Denaturation		94	30	
Annealing	32	55	45	
Extension		72	60	1 min per kb
Final extension	1	72	420	
		4	∞	

kb: kilobase.

Supplemental table 6: *Index PCR unit for 106 reactions.*

	Volume (µl)
Roche Mastermix	
Reaction Buffer (10X) with 18 mM MgCl2	530
dNTP-mix (10 mM each)	106.0
Enzyme Blend (5 U/µl)	42.4
RNase Free water	2596.0
Total	**3274.4**
PCR unit	
MM to distribute per Epp	35.0
i7-index per sample	5.0
i5-index per sample	5.0
Volume of DNA sample	5.0
Volume PCR unit	**50**

dNTP: deoxynucleotide; MM: mastermix.

Supplemental table 7: *Index PCR cycles*

	No. cycles	Degrees (°C)	Time	
Initial denaturation	1	95	2 min	
Denaturation		95	30 s	
Annealing	24	55	30 s	
Extension		72	1 min	1 min per kb
Final extension	1	72	6 min	
		4	∞	

kb: kilobase.

Supplemental table 8: *Primers used in the PCR and sequencing reactions.*

Primer	Primer Sequence (5' – 3')
Amplicon PCR	
515f with Illumina overhang adapter	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG - GTGCCAGCMGCCGCGGTAA
806r with Illumina overhang adapter	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG - GGACTACHVGGGTWTCTAAT
Index PCR	
Forward primer	AATGATACGGCGACCACCGAGATCTACAC - [i5 8-mer index] - TCGTCGGCAGCGTC
Reverse primer	CAAGCAGAAGACGGCATACGAGAT - [i7 8-mer index] - GTCTCGTGGGCTCGG
Sequencing	
Read 1 primer	ACACTCTTTCCCTACACGACGCTCTTCCGATCT
Index read primer	GATCGGAAGAGCACACGTCTGAACTCCAGTCAC
Read 2 primer	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

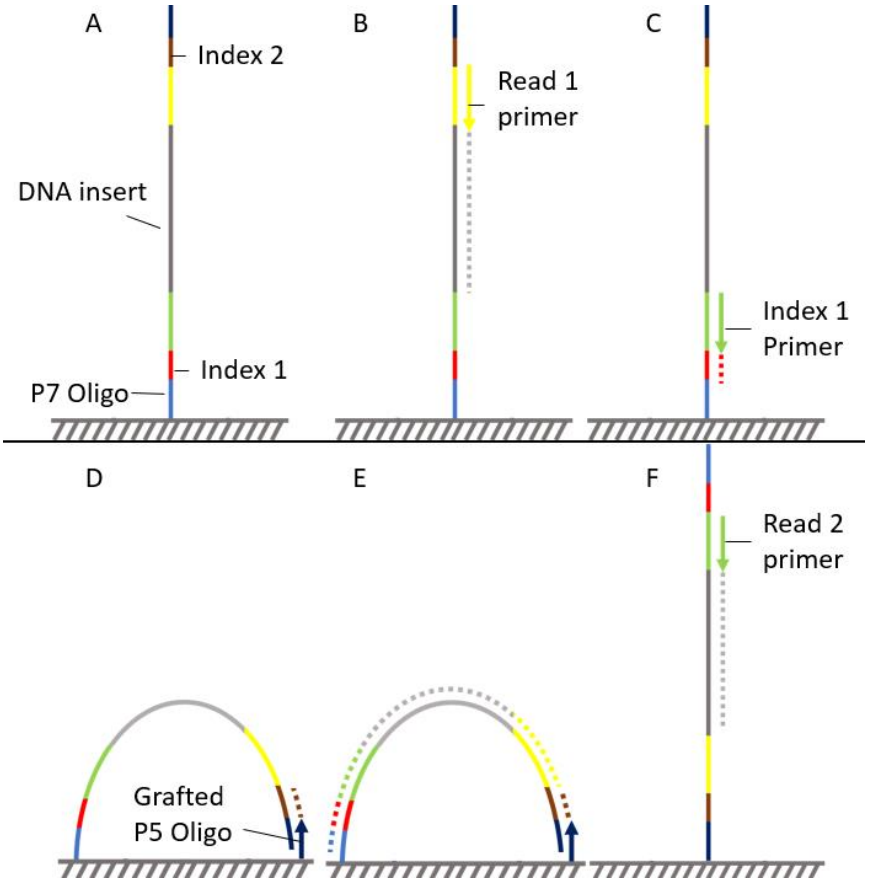
Supplemental table 9: *List of used indexes for the index PCR.*

Nextera XT index kit V2 set A			
i5 Index	**Sequence**	**i7 Index**	**Sequence**
S502	CTCTCTAT	N701	TCGCCTTA
S503	TATCCTCT	N702	CTAGTACG
S505	GTAAGGAG	N703	TTCTGCCT
S506	ACTGCATA	N704	GCTCAGGA
S507	AAGGAGTA	N705	AGGAGTCC
S508	CTAAGCCT	N706	CATGCCTA
S510	CGTCTAAT	N707	GTAGAGAG
S511	TCTCTCCG	N710	CAGCCTCG
		N711	TGCCTCTT
		N712	TCCTCTAC
		N714	TCATGAGC
		N715	CCTGAGAT

Attachment 3: Supplemental data

In this section, the raw data and additional figures and tables are displayed.

Supplemental data 1: Sequencing of 16S rRNA



Supplemental figure 1: Dual-indexed next-generation sequencing with the Illumina MiSeq*. (A) During the clustering step, the sample DNA is multiplied on oligos (P7 and P5) attached to the flow cell and the DNA attached to P5 oligos is removed. (B) The read 1 primer is annealed and sequencing of the DNA insert occurs. (C) The index 1 primer is annealed and index 1 is sequenced. (D) The sample DNA anneals to the grafted P5 oligo and index 2 is sequenced. (E) The sample DNA is copied and removed from the flow cell. (F) The read 2 primer anneals on the sample DNA copy and the DNA insert is sequenced.*

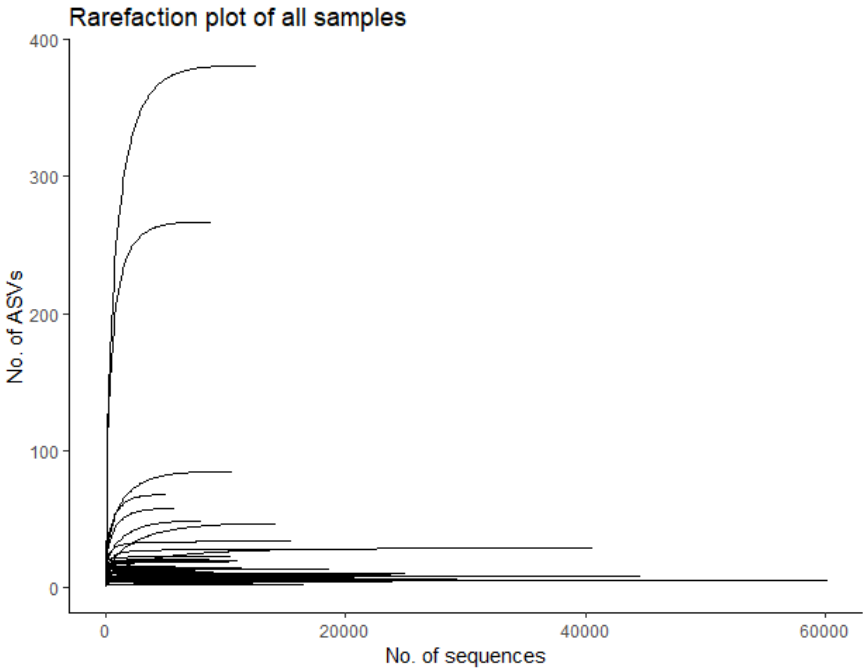
Supplemental data 2: Particle area covering the leaf

Supplemental table 10: The average BC particle area covering the leaf for each ID.

ID	Number of measurements	Average particle area (±SD)
95	72	221679 (±28149)
228	54	154411 (±30526)
252	54	60402 (±20098)
295	54	221415 (±60992)
332	72	225917 (±50122)
333	54	268049 (±92310)
342	54	265088 (±98474)
343	72	203625 (±51660)
364	54	120372 (±20688)
375	54	92394 (±45107)
434	54	86767 (±28904)
470	54	183057 (±32525)
518	72	368172 (±67799)
529	72	313133 (±46341)
530	72	259606 (±51429)
562	71	199738 (±76091)
567	72	155234 (±60177)
569	72	442087 (±51010)
575	72	261815 (±68470)
591	72	226760 (±74175)
613	72	381657 (±73546)
640	72	321516 (±58496)
651	72	157452 (±23067)
658	72	230275 (±47998)
671	72	89300 (±20337)
684	72	98378 (±22956)
687	72	239677 (±36431)
707	72	195292 (±10316)
710	72	473539 (±73137)
713	72	255856 (±64663)
715	72	494148 (±73650)
745	72	243442 (±57894)
754	72	189233 (±36377)
807	72	246322 (±255716)

SD: standard deviation.

Supplemental data 3: Sequencing data



Supplemental figure 2: The rarefaction plot of all sequenced samples.