Academic year 2015 – 2016

Faculty of Pharmaceutical, Biomedical and Veterinary Sciences

Biochemistry and Biotechnology

The structure of phyllospheric microbial communities and their role in phytoremediation of air pollution

Vincent Stevens

Thesis in partial fulfillment of the requirements for the degree

Master of Science in Biochemistry and Biotechnology

Promotor: prof. dr. Jaco Vangronsveld

Copromotor: dr. ir. Nele Weyens

Mentor: dr. Sofie Thijs

Centre for Environmental Sciences (CMK)

Hasselt University, Campus Diepenbeek

Agoralaan Building D

3590 Diepenbeek, Belgium

ABSTRACT

Air pollution constitutes a global threat. It is most pronounced in high-traffic environments comprising substantial amounts of airborne pollutants, such as particulate matter and volatile organic compounds. Many years of research have shown that plant-microbe interactions can be exploited to enhance phytoremediation of contaminated environments. Despite the complexity in composition of air pollution, phyllospheric microorganisms constitute promising candidates to detoxify the pollutants by means of degradation, transformation or sequestration. In this study, we employed complementary culture-independent techniques, 454 pyrosequencing and real-time PCR, to characterize the structure of phyllospheric microbial communities associated with common ivy (*Hedera helix*) growing at a polluted high-traffic urban area and a nonpolluted nature reserve in Belgium. Actinobacteria, Firmicutes, and Betaproteobacteria were more abundant in the polluted environment, while Bacteroidetes were less abundant. Moreover, our results indicated that the phyllosphere community of *H. helix* is similar to that of *Arabidopsis thaliana*, dominated by Actinobacteria, Bacteroidetes, Alphaproteobacteria and Gammaproteobacteria. Further, using a culture-dependent approach, we showed that air pollution increases the prevalence of bacterial functional traits affiliated to plant growth and phytoremediation of airborne pollutants. We evaluated different culture media for their potential to capture a substantial part of the total phyllospheric bacterial community, and YMAb 284 medium, an own creation, was found to be most representative of this community. In conclusion, we provide for the first time a detailed and integrated description of phyllospheric microbial communities associated with *H. helix*, including the identification of interesting bacteria with a significant potential for use in phytoremediation of air pollution.

SAMENVATTING

Luchtvervuiling vormt een wereldwijde bedreiging. Het is meest problematisch in omgevingen met veel verkeer, waar de omgevingslucht vaak grote hoeveelheden fijn stof en vluchtige organische stoffen bevat. Vele jaren onderzoek tonen aan dat plant-microbe interacties kunnen worden benut om fytoremediatie van vervuilde omgevingen te verbeteren. Ondanks de complexe samenstelling van luchtverontreiniging, zijn fyllosfeerbacteriën veelbelovende kandidaten om luchtpolluenten te detoxificeren via degradatie, transformatie of sequestratie. In deze studie hebben we gebruikt gemaakt van complementaire cultivatie-onafhankelijke technieken, 454 pyrosequencing en real-time PCR, om de structuur van microbiële fyllosfeergemeenschappen geassocieerd met klimop (*Hedera helix*), groeiende in een vervuild stedelijk gebied met veel verkeer en een niet vervuild natuurgebied in België, in kaart te brengen. Actinobacteria, Firmicutes en Betaproteobacteria waren meer aanwezig in de vervuilde omgeving, terwijl Bacteroidetes minder aanwezig was. Bovendien tonen onze resultaten aan dat de fyllosfeergemeenschap van *H. helix* vergelijkbaar is met die van *Arabidopsis thaliana*, gedomineerd door Actinobacteria, Bacteroidetes, Alphaproteobacteria en Gammaproteobacteria. Via een cultivatie-afhankelijke benadering toonden we ook aan dat luchtvervuiling de prevalentie van bacteriële functionele eigenschappen gerelateerd aan plantengroei en fytoremediatie van luchtpolluenten verhoogd. We evalueerden verschillende bacteriële kweekmedia op hun potentieel om een substantieel deel van de totale bacteriële fyllosfeergemeenschap te cultiveren, en YMAb 284 medium, een eigen creatie, bleek meest representatief te zijn voor deze gemeenschap. Ter conclusie presenteren wij met deze studie voor het eerst een gedetailleerde en geïntegreerde beschrijving van microbiële fyllosfeergemeenschappen geassocieerd met *H. helix,* inclusief de identificatie van interessante bacteriën met een aanzienlijk potentieel voor gebruik in fytoremediatie van luchtverontreiniging.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Doing my thesis work in the field of microbial ecology has been a fascinating and fulfilling experience. I thank my promotor, prof. dr. Jaco Vangronsveld, for taking me on as a student and for introducing me to a world of exciting, innovative and international research. Thank you for your trust and for letting me work with a lot of independence. Thank you for the research stay at the Warsaw University of Life Sciences. Visiting other research groups to broaden my knowledge and skills and to meet new people really constitutes a fundamental part of being a scientist to me. I would also like to thank my copromotor, dr. ir. Nele Weyens, for her feedback, support and encouragement.

My mentor, dr. Sofie Thijs, played an indispensable role during my stay at the Centre for Environmental Sciences. Sofie, you really are a highly competent scientist and teacher with a great personality. Thank you for everything.

I thank prof. dr. Tom Artois and prof. dr. Reinhart Ceulemans for their willingness to read and evaluate my written thesis.

I thank my lovely colleagues and friends for making life in and outside the laboratory enjoyable. Working at the Centre for Environmental Sciences really was the best part of my academic career so far.

LIST OF KEY ABBREVIATIONS

ACC	1-Aminocyclopropane-1-carboxylic acid
ANOSIM	Analysis of similarities
ARISA	Automated rRNA intergenic spacer analysis
BTEX	Benzene, toluene, ethylbenzene and xylene
DCPIP	2,6-Dichlorophenolindophenol
EDX	Energy-dispersive X-ray spectroscopy
emPCR	Emulsion-based PCR
EPFR	Environmentally persistent free radical
FISH	Fluorescence *in situ* hybridization
HPLC	High-performance liquid chromatography
IAA	Indole-3-acetic acid
LDA	Linear discriminative analysis
LEfSe	LDA with effect size
MID	Multiplex identifier
NGS	Next generation sequencing
NMDS	Nonmetric multidimensional scaling
OTU	Operational taxonomic unit
PAH	Polycyclic aromatic hydrocarbon
PCR	Polymerase chain reaction
PGP	Plant growth-promoting
PM	Particulate matter
PTFE	Polytetrafluoroethylene
qPCR	Real-time PCR
RA‰	Relative abundance per thousand
ROS	Reactive oxygen species
SEM	Scanning electron microscopy
UFP	Ultrafine particles
VOC	Volatile organic compound

INTRODUCTION

Air pollution – a global threat

Air pollution is a significant problem in many countries around the world, including Belgium. It is most pronounced in high-traffic environments like urban areas which comprise substantial amounts of airborne pollutants, such as particulate matter (PM) and volatile organic compounds (VOCs) (Zhang & Batterman, 2013). These pollutants are mainly generated through extensively used thermal processes such as the combustion of fuels (*e.g.* diesel fuel). Many epidemiologic studies have revealed associations between air pollution and adverse health effects (Saravia *et al.*, 2013; Saravia *et al.*, 2014). Moreover, a recent study published in Nature showed that more people now die from air pollution than malaria and human immunodeficiency virus (HIV) infection together (Lelieveld *et al.*, 2015). This includes 180,000 people per year in Europe.

Both PM and VOCs exert toxic effects in humans. Consensus exists that the level of PM toxicity depends on the chemical composition, particle size and shape (Kelly & Fussell, 2012). PM is composed of a relatively nonreactive part (*e.g.* carbon, calcium), to which biologically active chemicals such as trace metals (*e.g.* copper, zinc, cadmium), organic compounds (*e.g.* benzene, toluene) and environmentally persistent free radicals (EPFRs) can be adsorbed (Kelly & Fussell, 2012; Kim & Fergusson, 1994). The most dominant hypothesis is that ultrafine particles (UFP; *i.e.* PM of less than 100 nm in diameter) are most toxic and that this toxicity is caused by inducing the generation of reactive oxygen species (ROS) on their surface. Moreover, due to their small size, these particles are able to deeply penetrate lung tissue and subsequently enter systemic circulation, significantly increasing the range of effect (Kumar *et al.*, 2014; Saravia *et al.*, 2013). VOCs, of which some major molecules include benzene, toluene, ethylbenzene and xylene (BTEX), polycyclic aromatic hydrocarbons (PAHs) and formaldehyde, are also known to induce various short- and long-term adverse health effects in humans (Nurmatov *et al.*, 2015).

Although technological improvements and stricter emission regulations have contributed to a decrease in traffic-related emissions of airborne pollutants, this decrease has been offset by an increase in the number of vehicles, short trips and traffic congestions (Kelly & Fussell, 2012). To illustrate this, the latest annual report on air quality in Europe states that the EU limit and target values for PM continued to be exceeded in large parts of Europe in 2013 (European Environment Agency. Air quality in Europe – 2015 report). It is clear that current emission abatement strategies are not sufficient to meet environmental challenges and that alternative and additional, non-intruding, socially acceptable, multifunctional and cost-effective solutions need to be developed.

Getting help from plants and their microbes

Plants and their associated microorganisms are characterized by many complex interactions and are the subject of extensive research and diverse applications. For example, plant-associated bacteria often show plant growth-promoting (PGP) capabilities that can be exploited to obtain an enhanced growth of food and fodder crops, higher biomass production of energy crops and improved growth of plants on marginal soil (Weyens *et al.*, 2009b). Furthermore, many bacteria show a natural capacity to cope with various contaminants, making them useful in improving the efficiency of phytoremediation, *i.e.* the use of plants and their associated microorganisms to clean up contaminated environments (Weyens *et al.*, 2009b). Phytoremediation has already successfully been applied for soil and groundwater contaminated with trace metals and/or organic compounds (Pilon-Smits, 2005; Vangronsveld *et al.*, 2009). Plant-associated bacteria with the desired characteristics were enriched in plants by means of inoculation. After inoculation, an increased biomass, contaminant uptake and/or detoxification as well as a reduced phytotoxicity could be achieved (Barac *et al.*, 2004; Weyens *et al.*, 2009a).

Living on and in leaves

A diverse and abundant community of microorganisms naturally exists in the phyllosphere, *i.e.* the total above-ground portions of plants (Vorholt, 2012). It is one of the most prevalent microbial habitats on earth and bacteria are by far the most abundant and constant members, with a typical cell density of 10^6-10^7 cells/cm^2 (Bringel & Couée, 2015; Lindow & Brandl, 2003; Vorholt, 2012). Fungal communities are less numerous and archaea are rather non-abundant (Delmotte *et al.*, 2009; Knief *et al.*, 2012; Voriskova & Baldrian, 2013). Occasionally, also protozoa and nematodes can be found (Lindow & Brandl, 2003). Because in almost all plant species the majority of the surface area available for microbial colonization is located on the leaves, this is considered the dominant tissue of the phyllosphere (Vorholt, 2012).

Characterizing the structure and dynamics of the phyllosphere microbiome is a key step towards understanding how the microbial community may influence plant health and development. The composition of this community is dynamic, responding to environmental factors, both biotic and abiotic, such as leaf age and the co-presence of other microorganisms, flora, meso- and macrofauna, ultraviolet (UV) light exposure, (air) pollution, fertilization and water limitations (Müller & Ruppel, 2014). The composition has also a constant component, influenced by stable factors such as host plant species and geographical location (Vorholt, 2012). Host plant species can be the primary factor driving the composition of the phyllosphere microbiome (de Oliveira Costa *et al.*, 2012; Kim *et al.*, 2012), while in other cases geographic location can have the greatest influence on community composition (Finkel *et al.*, 2011; Rastogi *et al.*, 2012; Redford *et al.*, 2010). Perennial plants often have season-dependent microbial communities that are highly similar from year to year (Redford & Fierer, 2009), while other perennial plant-associated microbial communities change dramatically from one year to another and have high seasonal variability (Jackson & Denney, 2011).

In case of phytoremediation of air pollution, the surface of leaves is known to adsorb significant amounts of pollutants (Dzierżanowski *et al.*, 2011). Therefore, the phyllosphere microbiome is of high interest. Part of the adsorbed pollution is also finding its entry into the plant, making (particularly) leaf endophytes interesting to look at. Endophytic bacteria can be defined as bacteria colonizing the internal tissues of plants without causing symptoms of infection or other negative effects on their host (Reinhold-Hurek & Hurek, 2011). The primary site where endophytes gain entry into plants is via the roots (Germaine *et al.*, 2004). Once inside the plant, endophytes either reside in specific plant tissues like the root cortex or the xylem, or colonize the plant systematically by transport through the vascular system or the apoplast (Schmidt *et al.*, 2011). Leaf-endophytic and phyllospheric bacteria constitute promising candidates to detoxify part of the airborne pollutants by means of degradation, transformation or sequestration (Weyens *et al.*, 2015). Bacterial PGP traits such as the production of phytohormones and enzymes involved in phytohormone metabolism might positively affect pollutant accumulation through an increased (leaf) surface area (Sæbø *et al.*, 2012). Further, many members of phyllospheric bacterial communities are capable of nitrogen fixation, nitrification or methanol degradation and could potentially contribute to nutrient acquisition through the leaf (Delmotte *et al.*, 2009). Recently, the VOCs removal capacity of more than 100 plant species was summarized (Dela Cruz *et al.*, 2014). For the degradation of these VOCs plants significantly rely on their microbial community, mainly the bacterial members (Weyens *et al.*, 2015). A schematic overview of phytoremediation of airborne pollutants is presented in **FIGURE 1**.

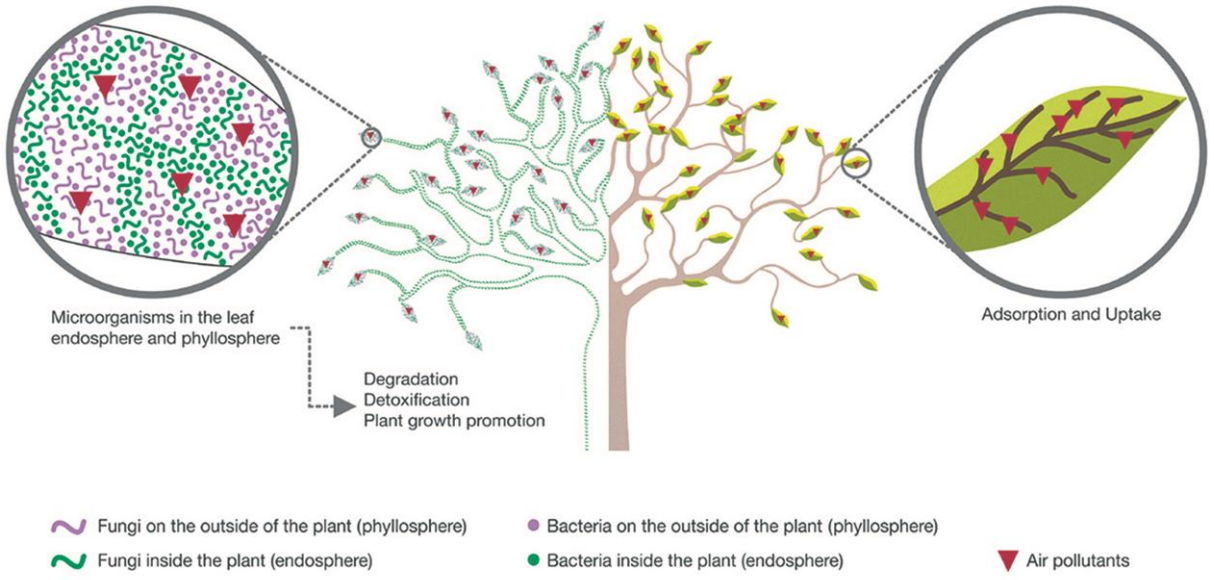

FIGURE 1 Schematic overview of phytoremediation of air pollution. Surfaces of leaves are known to adsorb significant amounts of airborne pollutants. Part of the adsorbed pollution is also finding its entry into the plant. Phyllospheric and leaf-endophytic microorganisms constitute promising candidates to detoxify these pollutants by means of degradation, transformation or sequestration. Adapted from Weyens *et al.* (2015).

From old-school microbiology to state-of-the-art microbial biotechnology

In order to investigate the potential of plant-microbe systems to decrease and detoxify airborne pollutants, a key step is to acquire knowledge about the structure and dynamics of leaf microbial communities. Consensus exists that 95-99.9% of bacteria from environmental samples has yet to be cultured (Ritz, 2007), which unavoidably separates this microbiological quest into a culture-dependent and culture-independent part.

To investigate the functional characteristics of bacterial members of microbial communities and to potentially exploit their beneficial effects by means of inoculation, these organisms need to be cultured under laboratory conditions. The first critical step in the culture-dependent approach is the isolation of bacteria. Different isolation procedures are described in literature, mainly dependent on the source of isolation (the phyllosphere, endosphere, rhizosphere, *etc.*) and the physical structure of the host plant species (Eevers *et al.*, 2015). The most crucial step however, and most likely the main reason why the majority of bacteria from environmental (plant) samples has yet to be cultured, is the choice of the culture medium (Ritz, 2007). Many different culture media are mentioned in literature, and a distinction can be made between complex, rich culture media that contain high amounts of rather undetermined nutrients and minimal culture media that contain significantly lower yet precise amounts of nutrients. The culture medium affects both the number and the diversity of bacteria that can be isolated from environmental samples and also the ultimate culturability of bacterial species altogether (Eevers *et al.*, 2015).

Once a collection of bacterial isolates is obtained and maintained in the laboratory, different phenotypic analyses can be conducted in order to investigate the capacity of these bacteria to cope with a specific (*e.g.* polluted) environment. This way, bacteria can be tested for their capacity to degrade diesel fuel, to cope with the PM-related trace metals copper, zinc and cadmium (Kim & Fergusson, 1994) and to improve pollutant adsorbance by producing the PGP hormones indole-3-acetic acid (IAA) and 3-hydroxy-2-butanone (acetoin), and the enzyme 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase. IAA is the most common phytohormone of the auxin class, and induces cell elongation and division with all subsequent results for plant growth and development (Zhao, 2010). The volatile phytohormone acetoin was shown to promote growth and induce systemic resistance in *Arabidopsis thaliana* (Ryu *et al.*, 2004; Ryu *et al.*, 2003). ACC deaminase catalyzes the hydrolysis of ACC, the immediate precursor of ethylene, resulting in the formation of ammonia and α-ketobutyrate and thereby reduces ethylene levels. Ethylene inhibits growth of roots and shoots; therefore lower levels of this phytohormone imply plant growth promotion (Glick, 2014).

Taxonomic identification of the phenotypically analyzed bacteria (who does what?) is essential to expand our knowledge of microbial biodiversity, and thus has become a cornerstone of microbial research. In this context, the 16S ribosomal RNA (rRNA) gene is the most commonly used molecular marker (Wang & Qian, 2009); it is universally present in all bacteria and contains regions with high variability and regions with low variability (Fox *et al.*, 1980). The regions with low variability can act as good primer targets across many phylogenetically diverse bacteria, designed to amplify the regions with high variability in

order to distinguish between closely related taxa (Chakravorty *et al.*, 2007; Maughan *et al.*, 2012). Using a culture-dependent approach, almost the whole 16S rRNA gene sequence of isolated bacteria can be determined using the Sanger sequencing method with an appropriate primer pair (Sanger & Coulson, 1975).

Culture-independent methods can capture an enormous greater richness and diversity of bacteria compared to culture-dependent approaches (Yashiro *et al.*, 2011). However, it is not possible to phenotypically investigate bacteria and to exploit their beneficial effects by means of inoculation. Nevertheless, culture-independent approaches are essential as these provide a lot of information about the structure (metagenomics) and function (metatranscriptomics) of microbiomes, providing also a guideline for culture-dependent research (Müller & Ruppel, 2014). DNA is directly isolated from environmental samples, without prior culturing of the microorganisms within the samples. Hence, there is no "culturing bias", and the whole microbial community is reflected much more realistically (Yashiro *et al.*, 2011). However, chloroplastal, mitochondrial and bacterial rRNA are highly homologous (Fox *et al.*, 1980). This homology makes amplification and sequencing of the bacterial 16S rRNA gene exclusively not straightforward, and optimization of the experimental setup (*e.g.* choosing an appropriate primer pair) is critical in order to obtain good data (Bulgarelli *et al.*, 2012). Further, a microbiome typically contains hundreds and potentially thousands of bacterial species, depending on the type and size of the environment being sampled. To identify every member of the microbial community and to accurately establish the relative abundances of these bacteria, thousands of sequences are required (Bokulich *et al.*, 2013). If this process were to be done using "old-school" Sanger sequencing, it would both be very time-consuming and expensive. Next generation sequencing (NGS) has had a dramatic impact on many fields of biology, including microbiology. NGS technologies sequence millions of DNA fragments simultaneously, at a relatively low cost, allowing whole communities of bacteria to be sequenced together (Metzker, 2010). In this context, bacterial community identification by NGS analysis based on partial 16S rRNA gene sequences is a widely used and effective method (Arenz *et al.*, 2015; Edwards *et al.*, 2015; Kõiv *et al.*, 2015; Lundberg *et al.*, 2013; Müller *et al.*, 2015; Shi *et al.*, 2014). One popular approach is to sequence the V3-V4 hypervariable regions of the bacterial 16S rRNA gene using 454 pyrosequencing technology. This technology is classified as "sequencing-by-synthesis" and exact sequence determination occurs by DNA polymerase-driven generation of inorganic pyrophosphate, resulting in the formation of ATP and ATP-dependent conversion of luciferin to oxyluciferin. The generation of oxyluciferin results in the emission of pulses of light, and the amplitude of each signal is directly related to the presence of one or more nucleosides (Petrosino *et al.*, 2009). Generally, 454 pyrosequencing technology is seen as advantageous because of the technical robustness of the chemistry and the relatively long sequences that are generated, allowing more frequent unambiguous mapping to complex targets (Metzker, 2010; Petrosino *et al.*, 2009).

Microbiome analysis – to see the wood for the trees

Microbiome research is becoming increasingly important for many fields, including health and disease, agriculture and bioremediation (Waldor *et al.*, 2015). The Human Microbiome Project (HMP) was established in 2008, with the mission of generating resources that would enable the comprehensive characterization of the human microbiome and analysis of its role in human health and disease (Peterson *et al.*, 2009). These standards are also being applied more recently to the Microbial Earth Project (MEP), the Earth Microbiome Project (EMP), the International Census of Marine Microbes (ICoMM) and many other environmentally based microbiome projects (Yilmaz *et al.*, 2011). Culture-independent studies of the leaf microbiome are still novel, and standardizing how the data is analyzed is crucial, since different methods can influence the obtained results substantially (Hamady & Knight, 2009; Vorholt, 2012).

Many software packages and tools have been (and are still being) developed in order to analyze microbiome data. Quantitative Insights Into Microbial Ecology (QIIME) is a popular tool for the analysis of such data (Caporaso *et al.*, 2010). This open-source bioinformatics pipeline is designed to take users from raw sequencing data generated on NGS platforms through "publication quality" graphics and statistics. This includes demultiplexing and quality filtering, operational taxonomic unit (OTU) picking, taxonomic assignment, phylogenetic reconstruction and diversity analyses. These analyses have been applied on a variety of data sets and prove to be reproducible and reliable (Gonzalez & Knight, 2012). Another popular software package designed for the analysis of microbiome data is mothur (Schloss *et al.*, 2009). Many of the algorithms are similar between QIIME and mothur, the largest difference being that all mothur scripts were developed in-house, rather than utilizing and combining algorithms from a range of sources. The purpose of both software packages is to aid in the analysis of microbial community data and develop a standardized methodology, allowing for better comparisons between data sets. Both software packages allow millions of sequences to be analyzed simultaneously and summarized in informative and integrated data displays.

A suitable host

In this study, common ivy (*Hedera helix*) is selected as host plant species. It is an evergreen plant with darkish green, broad and waxy leaves (**FIGURE 2**) that is known for its hardiness (*e.g.* high tolerance to salinity and drought) and wall climbing ability (Metcalfe, 2005). Furthermore, it has a high capacity to uptake and filter pollutants out of the air, a widespread occurrence and common presence in polluted city centres, which most likely provide an excellent source of (novel) bacteria capable of detoxifying airborne pollutants (Dzierżanowski *et al.*, 2011). Especially its common presence across much of the northern hemisphere makes *H. helix* very interesting regarding environmental and management implications in diverse urban settings, according to the "green architecture" concept (Sternberg *et al.*, 2010). Finally, it is also popular in indoor settings, further increasing its application range in the context of detoxifying and reducing airborne pollutants (Dela Cruz *et al.*, 2014).

FIGURE 2 Common ivy (*Hedera helix*), an evergreen plant with darkish green, broad and waxy leaves that is known for its hardiness and high capacity to uptake and filter pollutants out of the air.

Goals and objectives of this study

Since air pollution has become a major issue recognized by many people, the demand and public support to improve air quality has never been greater. This study contributes to providing a solution to this demand by exploring the use of plants and their associated microorganisms to clean airborne pollutants as a sustainable, cheap, appealing and socially acceptable phytoremediation technology. In order to exploit the selected plant-microbe system most effectively to decrease and detoxify airborne pollutants, an in-depth knowledge of the phyllosphere microbiome is necessary. Recent progress in sequencing technologies has already enhanced our understanding of the structure of phyllospheric microbial communities associated with *A. thaliana* (Vorholt, 2012). However, such data in the context of other plant-microbe systems is scarce. Moreover, the translation of this knowledge into phytoremediation applications is nonexistent. To address this goal and knowledge gap, we used complementary culture-dependent and culture-independent techniques to characterize the structure of phyllospheric microbial communities associated with *H. helix* growing at a polluted high-traffic urban area and a nonpolluted nature reserve. We combined the obtained phyllosphere microbiome data with an elaborate data set on cultured bacteria to (i) quantify and qualify the leaf-associated microbial biodiversity of *H. helix* at two environmentally distinct locations, (ii) identify potential interesting bacteria with regard to phytoremediation applications, (iii) understand relationships between microbial biodiversity, air pollution and traits of cultured bacterial isolates, and (iv) discuss their implications for our understanding of the use of plants and their bacteria to phytoremediate airborne pollutants.

MATERIALS AND METHODS

Sample collection and ambient air UFP measurements

H. helix leaves were collected at a high-traffic urban area in Hasselt, Belgium ("Polluted"; N 50°55'37" E 5°20'13") and nature reserve "De Maten" in Diepenbeek, Belgium ("Control"; N 50°56'24" E 5°26'19") on November 16, 2015 (n = 36; 18 "Polluted", 18 "Control"). Leaves were collected at shoulder height using sterile forceps, put separately in sterile tubes filled with 30 mL autoclaved phosphate buffer (6.33 g/L $NaH_2PO_4 \cdot H_2O$, 16.5 g/L $Na_2HPO_4 \cdot 7H_2O$, 100 µL/L Tween 80, pH 7.0) and immediately transferred to the laboratory. Leaf weight was determined gravimetrically. At the moment of sampling, ambient air pollution was assessed by measuring the UFP concentration and particle diameter for 20 min using an Aerasense NanoTracer (Philips Aerasense, Eindhoven, The Netherlands).

Isolation of phyllospheric microorganisms associated with *H. helix*

Microbial cells were washed from the surface of collected *H. helix* leaves using an optimized protocol. The tubes containing the leaves with 30 mL autoclaved phosphate buffer were vigorously shaken by hand (10 s), followed by sonication (160 W, 3 min), vortexing (1 min) and shaking on an orbital shaker (240 rpm, 15 min). Next, the leaf wash suspension containing phyllospheric microorganisms was centrifuged (4000 rpm, 15 min) and the pellet was resuspended in 3 mL remaining supernatant. The resuspended pellets of four samples were pooled resulting in two biological replicates per condition (n = 4; 2 "Polluted", 2 "Control"). 4 mL of all leaf wash samples was immediately stored at -80°C until DNA isolation. The remaining suspensions were stored overnight at 4°C, for culturing of phyllospheric bacteria.

Culturing of phyllospheric bacteria

100 µL of a ten-fold serial dilution ($0\text{-}10^{-1}\text{-}10^{-2}$) of all leaf wash samples (n = 4; 2 "Polluted", 2 "Control") was inoculated on six different solid culture media using 10-20 autoclaved glass beads and a shaker (170 rpm, 3 min). Plates were incubated at 30°C. All culture media were selected based on a literature review in order to cover a wide variety of bacterial taxa, including Proteobacteria (LB, 1/10 869), Actinobacteria (YECD, Flour 1) Firmicutes (LB, 1/10 869) and Rhizobiales (YMA) (Bergersen, 1961; Bertani, 1951; Coombs & Franco, 2003; Mergeay *et al.*, 1985). Gellan gum was used as polymer matrix in case of solid culture media (Eevers *et al.*, 2015). The compositions of the culture media are summarized in **TABLE 1**.

TABLE 1 Compositions of the culture media.

	LB (Bertani, 1951)	1/10 869 (Mergeay *et al.*, 1985)	YECD (Coombs & Franco, 2003)	YMA (Bergersen, 1961)	Flour 1 (Coombs & Franco, 2003)	YMAb 284[a] (This study)
$CaCl_2 \cdot 2H_2O$		0.035				
$CaCO_3$					0.3	
$CoCl_2$						190×10^{-6}
$CuCl_2$						17×10^{-6}
$Fe(III)NH_4$ citrate						4.8×10^{-3}
H_3BO_3						62×10^{-6}
K_2HPO_4			2	0.1		
KCl						
KH_2PO_4				0.4		
$MgCl_2 \cdot 6H_2O$						
$MgSO_4 \cdot 7H_2O$	1	1	0.5	0.2	0.5	0.2
$MnCl_2$						100×10^{-6}
$Na_2HPO_4 \cdot 2H_2O$						0.04
Na_2SO_4						
NaCl	10	0.5		0.1		0.1
$NaMoO_4$						36×10^{-6}
NH_4Cl						
$NiCl_2$						24×10^{-6}
Tris						6.06
$ZnSO_4 \cdot 7H_2O$						144×10^{-6}
D-fructose						0.54
D-glucose		0.1	0.3			0.52
Gluconate						0.66
Lactic acid						0.35
Mannitol				10		
Succinic acid						0.81
Sucrose					0.3	
Plain flour					6	
Tryptone	10	1				
Yeast extract	5	0.5	0.3	0.4	0.3	0.4
Gellan gum	7.5	7.5	7.5	7.5	7.5	7.5

Products are given in g/L distilled water. Gellan gum is not added in case of liquid culture media. Products marked in grey were filter-sterilized before being added to the other autoclaved components. Phosphate-containing components (underlined) were separately autoclaved to prevent the formation of growth-inhibiting compounds such as H_2O_2 (Tanaka *et al.*, 2014). All culture media have pH 7.0. [a]: combination of YMA and 284 media.

Establishing a collection of bacterial isolates

A collection of phyllospheric bacteria was obtained and maintained in culture during all subsequent analyses. Individual bacterial colonies (n = 192, 96 "Polluted", 96 "Control"), comprising a representative fraction, were picked from LB, 1/10 869, YECD and YMAb 284 plates (10^{-1} dilution) one week after culturing, and grown at 30°C on a shaker (150 rpm) in corresponding liquid culture media. After one week, all inoculated bacteria were spread on new solid culture media to check for purity, and repurified if needed. All bacteria were also flash-frozen in a solution of 20% (v/v) glycerol and 0.15 M NaCl, and stored at -45°C.

Taxonomic identification of culturable and total phyllospheric bacterial communities

454 pyrosequencing technology was used to taxonomically identify the phyllospheric bacteria present on solid culture media. First, bacterial cells were harvested from plates two weeks after inoculation. 10 mL autoclaved 10 mM $MgSO_4$ solution together with 10-20 autoclaved glass beads was added to LB, 1/10 869, YMA, Flour 1 and YMAb 284 plates (undiluted; n = 20, 4 per culture medium; 10 "Polluted", 10 "Control"), and plates were put on an orbital shaker (80 rpm, overnight). Next, the suspensions were collected in sterile tubes and centrifuged (4000 rpm, 15 min) to obtain bacterial pellets, which were immediately stored at -80°C until DNA isolation. Genomic DNA was isolated from the pellets using the DNeasy Blood and Tissue Kit (QIAGEN, Venlo, The Netherlands). The quantity and quality of the isolated DNA was determined spectrophotometrically using a NanoDrop ND-1000 UV-Vis Spectrophotometer (Isogen Life Science, De Meern, The Netherlands). The V3-V4 hypervariable regions of the bacterial 16S rRNA genes were PCR-amplified using 341F (5'-CCTACGGGNGGCWGCAG-3') and 785R (5'-GACTACHVGGGTATCTAATCC-3') primers with attached GS FLX Titanium adaptors for annealing the emulsion-based PCR (emPCR) and sequencing primers, the sequencing key TCAG, and a sample-specific multiplex identifier (MID) (Andersson *et al.*, 2008; Klindworth *et al.*, 2013). The 50 µL PCR reaction consisted of 1 x PCR buffer, 1.5 mM $MgSO_4$, 0.2 µM dNTP mix, 0.2 µM of each primer, 1.25 U Taq polymerase (FastStart High Fidelity PCR System, Roche Diagnostics, Vilvoorde, Belgium), 1 µL DNA template, and autoclaved molecular grade water. Amplification conditions were: 95°C for 3 min, 30 cycles of 94°C for 20 s, 53°C for 40 s, 72°C for 40 s, followed by a 7 min extension at 72°C. Following PCR, 5 µL of the 514 bp PCR product was run on a 1.5% agarose gel for evaluation. PCR products of all samples were purified using the UltraClean PCR Clean-Up Kit (MO BIO Laboratories, Carlsbad, California, USA).

For taxonomic identification of total phyllospheric bacterial communities using 454 pyrosequencing technology, 4 mL of all leaf wash samples (n = 4; 2 "Polluted", 2 "Control") was centrifuged (13200 rpm, 20 min) at 4°C and genomic DNA was isolated from the pellets using the DNeasy Blood and Tissue Kit (QIAGEN, Venlo, The Netherlands). Similar as described above, the V3-V4 hypervariable regions of the bacterial 16S rRNA genes were PCR-amplified using a two-step PCR. After the second PCR, 160 µL PCR product per sample was loaded on a 1.5% agarose gel. The 514 bp bacterial amplicon was excised from the gel and purified using the UltraClean GelSpin DNA Extraction Kit (MO BIO Laboratories, Carlsbad, California, USA).

All samples were brought to equimolar concentrations (10^{10} molecules/µL) using Quant-iT PicoGreen (Life Technologies, Carlsbad, California, USA). Two libraries were made, one for culturable and one for total phyllospheric bacterial communities. Correct amplicon size and integrity were checked by analyzing 1 µL of each library on an Agilent 2100 Bioanalyzer system (Agilent Technologies, Diegem, Belgium). Each library was sequenced using GS FLX Titanium chemistry (Macrogen Korea, Seoul, South Korea).

Sanger sequencing of bacterial isolates

Genomic DNA of the cultured phyllospheric bacteria (n = 192, 96 "Polluted", 96 "Control") was isolated using the MagMAX DNA Multi-Sample Kit (Life Technologies, Carlsbad, California, USA) and a MagMAX Express-96 Deep Well Magnetic Particle Processor (Life Technologies, Carlsbad, California, USA). The quantity and quality of the isolated DNA was determined spectrophotometrically using a NanoDrop ND-1000 UV-Vis Spectrophotometer (Isogen Life Science, De Meern, The Netherlands). The bacterial 16S rRNA gene was partially PCR-amplified using 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1492R (5'-TACGGYTACCTTGTTACGACTT-3') primers. The 25 µL PCR reaction consisted of 1 x PCR buffer, 1.5 mM $MgSO_4$, 0.2 µM dNTP mix, 0.2 µM of each primer, 1.25 U Taq polymerase (FastStart High Fidelity PCR System, Roche Diagnostics, Vilvoorde, Belgium), 1 µL DNA template, and autoclaved molecular grade water. Amplification conditions were: 95°C for 5 min, 30 cycles of 94°C for 1 min, 57°C for 30 s, 72°C for 3.5 min, followed by a 10 min extension at 72°C. Following PCR, 5 µL of the PCR product was run on a 1.5% agarose gel for evaluation and 20 µL was used for unidirectional Sanger sequencing using the 27F primer (Macrogen Europe, Amsterdam, The Netherlands).

Sanger sequencing data was processed using Geneious v.4.8.5 (Kearse *et al.*, 2012). First, 5'- and 3'-ends of all sequences were trimmed using an Error Probability Limit of 0.05. Next, sequences containing less than 100 bp and/or a quality score (HQ%) of less than 50% were removed. All remaining sequences were analyzed with the Sequence Match tool of the Ribosomal Database Project (RDP; Release 11; May 26, 2015) for taxonomic identification of the cultured phyllospheric bacteria.

qPCR quantification

Selected taxa were quantified in total phyllospheric microbial communities using real-time PCR (qPCR). First, standard curves were made for all taxa in order to allow for absolute quantification. Relevant sequences were obtained by PCR amplification, with selected primer pairs, of genomic DNA isolated using the DNeasy Blood and Tissue Kit (QIAGEN, Venlo, The Netherlands) from following organisms: *Penicillium canescens* (kingdom Fungi), *Escherichia coli* (domain Bacteria), *Okibacterium fritillariae* (phylum Actinobacteria), *Chryseobacterium vrystaatense* (phylum Bacteroidetes), *Bacillus mycoides* (phylum Firmicutes), *Novosphingobium barchaimii* (class Alphaproteobacteria), *Burkholderia sediminicola* (class Betaproteobacteria) and *Pseudomonas flavescens* (genus *Pseudomonas*). All primers were selected based on a literature review and synthesized by Integrated DNA Technologies (Leuven, Belgium); an overview is presented in **TABLE 2**. PCR products were run on a 1.5% agarose gel to confirm the specificity of the amplification, cloned in a 3000 bp pGEM-T Vector (Promega Benelux, Leiden, the Netherlands), and incorporated in competent *E. coli* JM109 cells (Promega Benelux, Leiden, the Netherlands). Plasmids were isolated from successfully transformed *E. coli* cells using the UltraClean Standard Mini Plasmid Prep Kit (MO BIO Laboratories, Carlsbad, California, USA) and sent for bidirectional Sanger sequencing (Macrogen Europe, Amsterdam, The Netherlands) to confirm the correct inserts.

DNA concentrations of the plasmid suspensions were determined using Quant-iT PicoGreen (Life Technologies, Carlsbad, California, USA). Standard curves with plasmid DNA concentrations ranging from 10^3 to 10^8 copies/µL (ten-fold serial dilution) and a no template control (NTC) were made in triplicate for all taxa.

The 10 µL qPCR reaction consisted of 5.0 µL QuantiTect SYBR Green PCR Master Mix (QIAGEN, Venlo, The Netherlands), 0.2-0.8 µL of each primer (100-400 nM; depending on taxon, **TABLE 2**), 2 µL DNA template (0.5 ng/µL), and autoclaved molecular grade water. DNA of all leaf wash samples (n = 4; 2 "Polluted", 2 "Control") was obtained by centrifuging 4 mL (13200 rpm, 20 min) at 4°C and subsequently isolating genomic DNA from the pellets using the DNeasy Blood and Tissue Kit (QIAGEN, Venlo, The Netherlands). Amplification conditions were: 95°C for 15 min, 40 cycles of 94°C for 15 s, 30 s at the annealing temperature, 72°C for 30 s. Annealing temperatures (**TABLE 2**) were experimentally optimized to maximize the specificity of amplification. Each reaction was set up in triplicate per leaf wash sample (n = 12; 6 "Polluted", 6 "Control"). At the end of every qPCR reaction, a melting curve analysis was performed with following conditions: 95°C for 15 s, 60 s at the annealing temperature, 95°C for 15 s, 60°C for 15 s. All qPCR results were expressed as the log number of bacterial 16S rRNA or fungal 18S rRNA gene copies per µg leaf material. The qPCR setup for all standard curves was optimized in order to obtain a primer efficiency between 80-120% with linear resolution (R^2 > 0.99) over six orders of dynamic range (**FIGURE A1**, **APPENDIX I – REAL-TIME PCR STANDARD CURVES**). Melting curves were evaluated to confirm that the detected fluorescence originated from specific products and not from primer dimers or other amplification artifacts.

TABLE 2 Primers used for real-time PCR quantification of selected taxa.

Taxon	Primer	Sequence (5'→3')	Primer conc.	Annealing temp.	Reference
Fungi	Euk345F	AAGGAAGGCAGCAGGCG	400 nM	50°C	Zhu *et al.* (2005)
	Euk499R	CACCAGACTTGCCCTCYAAT			
Bacteria	Eub338F	ACTCCTACGGGAGGCAGCAG	400 nM	50°C	Fierer *et al.* (2005)
	Eub518R	ATTACCGCGGCTGCTGG			
Bacteria	Bact1369F	CGGTGAATACGTTCYCGG	400 nM	50°C	Suzuki *et al.* (2000)
	Prok1492R	GGWTACCTTGTTACGACTT			
Actinobacteria	Acti1154F	GADACYGCCGGGGTYAACT	100 nM	55°C	Pfeiffer *et al.* (2014)
	Acti1339R	TCWGCGATTACTAGCGAC			
Bacteroidetes	Bdet107F	GCACGGGTGMGTAACRCGTAT	100 nM	55°C	Pfeiffer *et al.* (2014)
	Bdet309R	GTRTCTCAGTDCCARTGTGGG			
Firmicutes	Firm352F	CAGCAGTAGGGAATCTTC	100 nM	55°C	Pfeiffer *et al.* (2014)
	Firm525R	ACCTACGTATTACCGCGG			
Alphaproteobacteria	aProt528F	CGGTAATACGRAGGGRGYT	400 nM	55°C	Pfeiffer *et al.* (2014)
	aProt689R	CBAATATCTACGAATTYCACCT			
Betaproteobacteria	bProt972F	CGAARAACCTTACCYACC	300 nM	55°C	Pfeiffer *et al.* (2014)
	bProt1221	GTATGACGTGTGWAGCC			
Pseudomonas	Pse435F	ACTTTAAGTTGGGAGGAAGGG	100 nM	55°C	Bergmark *et al.* (2012)
	Pse686R	ACACAGGAAATTCCACCACCC			

For all bacterial taxa (domain Bacteria, phyla Actinobacteria, Bacteroidetes and Firmicutes, classes Alphaproteobacteria and Betaproteobacteria, and genus *Pseudomonas*), the primers target a part of the bacterial 16S rRNA gene. For the fungal taxon (kingdom Fungi), the primers target a part of the fungal 18S rRNA gene.

Analyses of plant growth promotion

Cultured phyllospheric bacteria (n = 192, 96 "Polluted", 96 "Control") were grown for 24 hours in liquid 869 medium (Mergeay *et al.*, 1985) at 30°C on a shaker (150 rpm), washed and resuspended in 2 mL autoclaved 10 mM $MgSO_4$ solution to obtain suspensions containing bacteria in mid-exponential phase ($OD_{600\ nm}$ = 0.4; OD: optical density). Next, 20 µL of this bacterial suspension was used for the detection of IAA production using the Salkowski's reagent method (Patten & Glick, 2002), of acetoin production using the Voges-Proskauer test (Romick & Fleming, 1998), and for assessing ACC deaminase activity by monitoring the amount of α-ketobutyrate that was generated by the enzymatic hydrolysis of ACC, as described by Thijs *et al.* (2014a), adapted from Belimov *et al.* (2005). For all analyses of PGP traits, bacteria were evaluated as either positive or negative for the respective analysis and the corresponding PGP trait. In each case, negative controls (without bacteria and with bacteria which are known to test negatively) and positive controls (with bacteria which are known to test positively) were included.

Evaluation of trace metal resistance

Cultured phyllospheric bacteria (n = 192, 96 "Polluted", 96 "Control") were grown overnight in liquid 869 medium at 30°C on a shaker (150 rpm), and then inoculated on solid 284 media (Schlegel *et al.*, 1961) enriched with either 0.4 mM $CuSO_4 \cdot 7H_2O$, 0.6 mM $ZnSO_4 \cdot 7H_2O$ or 0.4 mM $CdSO_4 \cdot 8H_2O$. All plates were incubated for one week at 30°C. Bacteria growing on these plates were subsequently inoculated on solid 284 media containing higher concentrations of the respective trace metals (0.8 mM $CuSO_4 \cdot 7H_2O$, 1.0 mM $ZnSO_4 \cdot 7H_2O$ or 0.8 mM $CdSO_4 \cdot 8H_2O$). Bacteria that were still growing on these plates were evaluated as being copper, zinc, or cadmium resistant, respectively.

Diesel fuel degradation assay

Cultured phyllospheric bacteria (n = 192, 96 "Polluted", 96 "Control") were grown overnight in liquid 869 medium at 30°C on a shaker (150 rpm), and washed three times with 0.01 M phosphate-buffered saline (PBS) buffer. Bacterial cells were left in PBS buffer overnight at 30°C on a shaker (150 rpm). Next, 80 µL of the bacterial suspension was transferred to 750 µL W minimal medium (Bučková *et al.*, 2013) with the addition of 50 µL filter-sterilized 150 µg/mL 2,6-dichlorophenolindophenol (DCPIP), 50 µL of filter-sterilized 150 µg/mL $FeCl_3 \cdot 6H_2O$, and 5 µL of polytetrafluoroethylene (PTFE) filter-sterilized 0.45 µM diesel fuel, as the sole carbon source. After incubation in the dark for 92 hours at 30°C on a shaker (150 rpm), suspensions showing a visible change from blue to colorless (due to reduction of DCPIP, indicating diesel fuel respiration) were evaluated as containing bacteria capable of diesel fuel degradation. Negative controls (without bacteria and with bacteria which are known to test negatively) and positive controls (with bacteria which are known to test positively) were included (Gkorezis *et al.*, 2015).

SEM and EDX of *H. helix* leaf surfaces

Square cuts (approx. 1 x 1 cm) were made from the centre of collected *H. helix* leaves (n = 20; 10 "Polluted", 10 "Control"). As described by Jiménez *et al.* (2012), leaf cuts were fixed in 10 mL Karnovsky's fixative (5 mL 0.2 M sodium cacodylate buffer, 2 mL 10% paraformaldehyde, 1 mL 25% glutaraldehyde, 3 mL distilled water) for 2 hours at 4°C, and then washed in 10 mL sodium phosphate buffer (0.95 mL 0.2 M $NaH_2PO_4 \cdot 2H_2O$, 4.05 mL 0.2 M $Na_2HPO_4 \cdot 2H_2O$, 5 mL distilled water) for 30 min. Subsequently, the leaf cuts were dehydrated in an ethanol series (30%, 50%, 70% for 15 min; 80%, 90%, 95%, 100% for 30 min), rinsed with isobutanol (2 x 10 min), and left overnight in isobutanol at -20°C. The next day, the leaf cuts were dried by sublimation for a minimum of 24 hours, before being mounted on a stub for scanning electron microscopy (SEM).

SEM was performed using a FEI Quanta 200 FEG-SEM (FEI Company, Eindhoven, The Netherlands) under low vacuum. Elemental identification was done by energy-dispersive X-ray spectroscopy (EDX), a chemical microanalysis technique used in conjunction with SEM. SEM and EDX was performed at the Institute for Materials Research (IMO) of Hasselt University (Hasselt, Belgium).

Bioinformatic processing of 454 pyrosequencing data

454 pyrosequencing data was processed using mothur v.1.36.1 (Schloss *et al.*, 2011; Schloss *et al.*, 2009); *cf.* **APPENDIX II – 454 PYROSEQUENCING DATA PROCESSING** for a complete protocol with each step/algorithm explained concisely. Briefly, all flowgram data was first partitioned by sample based on the sample-specific multiplex identifier (MID). Flowgrams outside the bounds of 360-720 flows, containing 8-mer and longer homopolymers, and having too many mismatches within primers/MIDs were removed. Next, using an expectation-maximization algorithm, the flowgrams were corrected to identify the idealized form of each flowgram and hereby translated to DNA (fasta) sequences. Sequences outside the bounds of 400-450 bp, barcodes and primers were removed. All data was aligned to the SILVA rRNA reference database (Release 123; July 23, 2015) and sequences that did not overlap in the same alignment space were removed. Sequences within 1 bp of a more abundant sequence were merged. Further, chimeras and sequences classified as "Chloroplast", "Mitochondria", or "Unknown" as well as archaeal and eukaryotic 16S/18S rRNA gene sequences were identified and removed. In order to bin sequences into OTUs, a distance matrix was generated and all sequences were clustered *de novo* into OTUs at sequence similarity ≥ 97%. Finally, an OTU table was generated and taxonomic information was linked to each OTU. Also using mothur, rarefaction and alpha/beta diversity analyses were performed. Three alpha diversity indices (the observed amount of OTUs, the nonparametric estimation of Shannon's diversity index and the inverse of Simpson's diversity index) were calculated at an equal sampling depth (Chao & Shen, 2003; Simpson, 1949). For beta diversity analyses, the Yue & Clayton measure of dissimilarity was calculated (Yue & Clayton, 2005).

Statistical analyses and data visualization

Before statistical analyses, OTU tables were normalized by dividing the OTU counts for each sample by the sum of all OTU counts within that sample and multiplying by 1000, resulting in relative abundance per thousand (RA‰). Next, a transformation of the data was performed ($T = \log_2(RA‰ + 1)$) and an inclusion threshold of $T > 2$ was defined such that OTUs have to be above this threshold in at least one sample (Paulson *et al.*, 2013).

R v3.3.0 (R Core Team) was used for performing nonparametric Wilcoxon rank-sum and Kruskal-Wallis tests, nonmetric multidimensional scaling (NMDS) and analysis of similarities (ANOSIM). Significant differentially abundant taxa in total and culturable phyllospheric bacterial communities between the polluted and control environment (culturable/total), and between the investigated culture media (culturable), were identified by linear discriminative analysis (LDA) with effect size (LEfSe) (Segata *et al.*, 2011). In the first step, LEfSe uses the Kruskal-Wallis test to detect taxa with significant differential abundance between the indicated classes, and in the second step a set of pairwise tests among subclasses is performed using the Wilcoxon rank-sum test. Finally, LEfSe uses LDA to estimate the effect size of each differentially abundant taxon (Segata *et al.*, 2011).

Graphs for visualizing rarefaction curves, and the qPCR quantification and standard curves of the selected taxa were made using Microsoft Excel 2010 (Microsoft Corporation, Redmond, Washington, USA). Taxonomic plots of total and culturable phyllospheric bacterial communities were realized using Explicet v.2.10.5 (Robertson *et al.*, 2013). To visualize exclusive and shared OTUs between groups, Venn diagrams were constructed using Venny v.2.1.0 (Oliveros, 2015). Graphical tree representations of differentially abundant taxa were generated using GraPhIan v1.0.0 (Asnicar *et al.*, 2015).

RESULTS

Higher UFP concentration and smaller particle diameter in an air-polluted environment

At the moment of sampling *H. helix* leaves, ambient air pollution was assessed by measuring the UFP concentration and particle diameter for 20 min. At the high-traffic urban area, the average UFP concentration was significantly higher (Wilcoxon rank-sum test, $p < 0.01$) and the average particle diameter was significantly smaller (Wilcoxon rank-sum test, $p < 0.05$) compared to the nature reserve (**TABLE 3**).

TABLE 3 Ambient air ultrafine particles (UFP) measurements at the moment of sampling.

Sampling location	Condition	Avg. concentration[a]	Avg. diameter[a]
High-traffic urban area Hasselt, Belgium (N 50°55'37" E 5°20'13")	Polluted	23599 UFP/cm^3 **	49 nm *
Nature reserve "De Maten" Diepenbeek, Belgium (N 50°56'24" E 5°26'19")	Control	3697 UFP/cm^3 **	55 nm *

[a]: the average of 120 measurements during a time period of 20 min. *: $p < 0.05$, **: $p < 0.01$; Wilcoxon rank-sum test.

Air pollution increases the amount of particulates and trace metals on leaves

SEM was performed on collected *H. helix* leaves (n = 20; 10 "Polluted", 10 "Control"). More particulates and other (mostly undetermined) components could be observed on polluted adaxial leaf surfaces, seen as light gray areas (**FIGURE 3a**). Moreover, elemental identification using EDX confirmed a higher amount of a range of (trace) metals on these leaf surfaces, seen as bright white spots, including (but not limited to) iron, silver, bismuth, copper and zinc.

Microorganisms, mainly bacteria and fungi, were observed readily dispersed on all inspected leaves, on both adaxial and abaxial leaf surfaces (**FIGURE 3a, b, c, d**). This identification was done mainly based on cell shape and size. Most rod-shaped and filamentous bacteria are between 1-10 µm in length, while the diameter of most spherical bacteria is between 0.5-2.0 µm. Fungal (cellular) structures are typically larger. Also, multicellular fungi tend to grow as hyphae. No remarkable differences in microbial cell density and distribution could be noticed on leaves between the polluted and control environment. Interestingly, it could be observed that microorganisms clustered in epidermal grooves, near stomata and on trichomes. Mainly bacteria were found around stomata (**FIGURE 4a**), while fungi dominated on trichomes (**FIGURE 4b**).

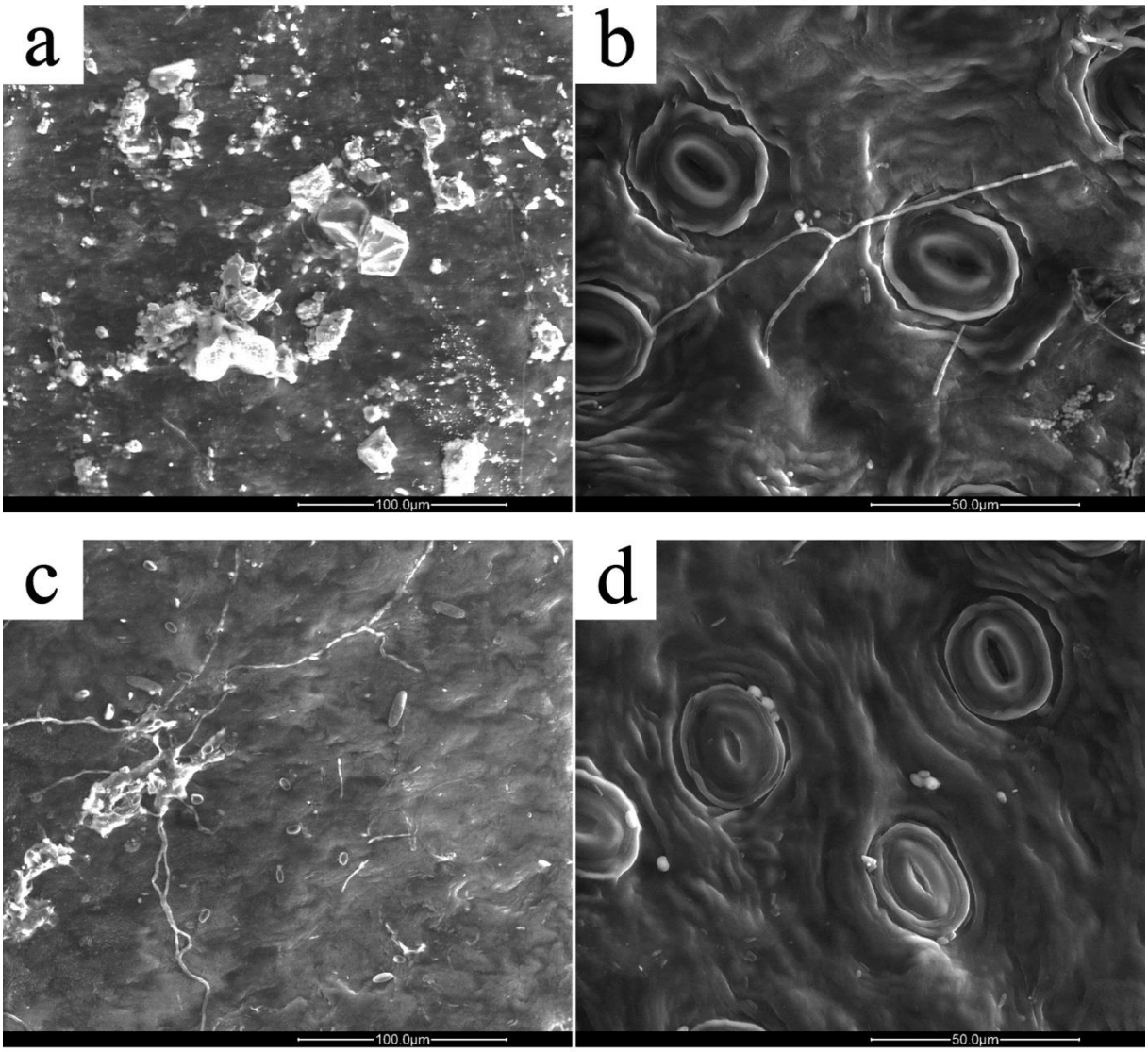

FIGURE 3 Representative scanning electron microscopy (SEM) images of adaxial (**a**, **c**) and abaxial (**b**, **d**) surfaces of *H. helix* leaves from the polluted (**a**, **b**) and control (**c**, **d**) environment. More particulates and trace metals were present on polluted adaxial leaf surfaces (**a**). Microorganisms, mainly bacteria and fungi, were found readily dispersed on adaxial and abaxial leaf surfaces from both sampling locations (**a**, **b**, **c**, **d**).

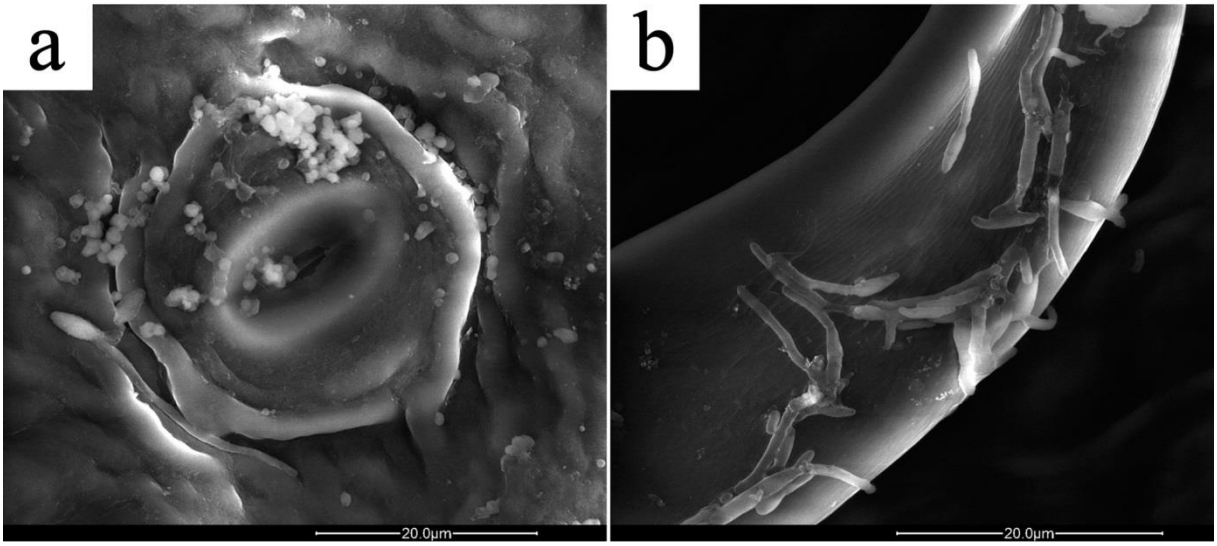

FIGURE 4 Representative scanning electron microscopy (SEM) images of a stoma (**a**) and the lower part of a trichome (**b**). Microorganisms clustered near stomata and on trichomes; mainly bacteria were found around stomata, while fungi dominated on trichomes.

Total and culturable phyllospheric microbial communities associated with *H. helix* – a taxonomic survey using 454 pyrosequencing and qPCR

454 pyrosequencing to taxonomically identify total phyllospheric bacterial communities, resulted in a total of 151,599 raw sequences with an average length of 405 bp. Using the mothur pipeline as previously described, a total of 21,157 high-quality sequences were recovered distributed over 3629 OTUs (273 different OTUs). For culturable phyllospheric bacterial communities, a total of 112,561 raw sequences with an average length of 435 bp were obtained; 37,137 high-quality sequences distributed over 966 OTUs (103 different OTUs) could be recovered. Both OTU tables were normalized and transformed as previously described, and an inclusion threshold was defined for statistical analyses.

Bacterial diversity measured as OTU richness was estimated for total and culturable phyllospheric bacterial communities by rarefaction analyses. Both rarefaction plots indicate that sequencing was adequate, as almost all curves show approximate saturation using a sequence similarity cut-off ≥ 97% to define OTUs (**FIGURE 5a, b**). Bacterial diversity of both total and culturable phyllospheric bacterial communities was further estimated by alpha diversity analyses. Three alpha diversity indices (the observed amount of OTUs, the nonparametric estimation of Shannon's diversity index and the inverse of Simpson's diversity index) were calculated at an equal sampling depth: 3200 and 1200 sequences per sample for total and culturable phyllospheric bacterial communities, respectively (**TABLE 4**). No significant effect of the condition ("Polluted" *vs.* "Control") could be revealed for any of the alpha diversity indices (Wilcoxon rank-sum test, $p < 0.05$), for both total and culturable phyllospheric bacterial communities. For culturable phyllospheric bacterial communities, alpha diversity (all three indices) was significantly higher in 1/10 869 samples and lower in Flour 1 samples (Kruskal-Wallis test, $p < 0.05$).

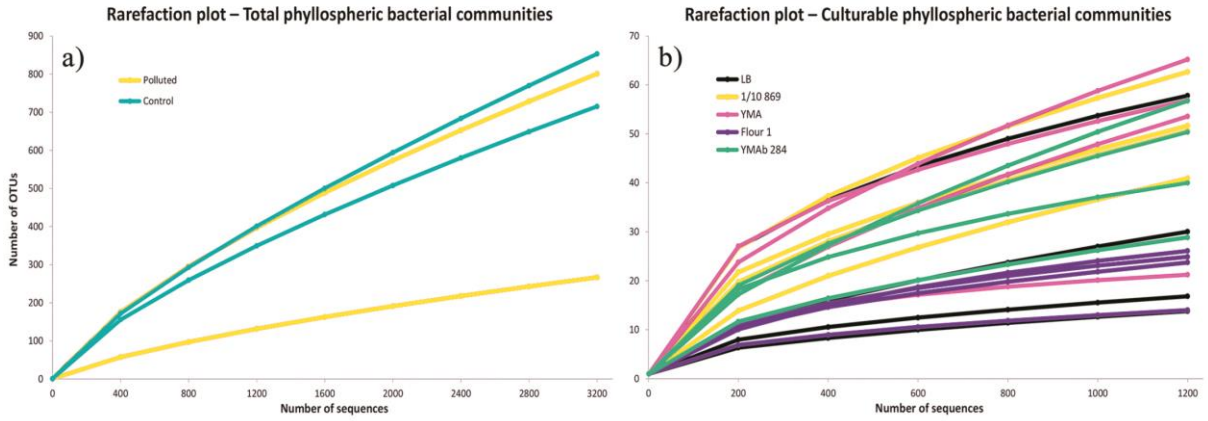

FIGURE 5 Rarefaction analyses. Rarefaction plots indicating the average number of operational taxonomic units (OTUs; sequence similarity ≥ 97%) for each sample, for total (**a**; n = 4) and culturable (**b**; n = 20) phyllospheric bacterial communities. In both cases, 454 pyrosequencing was adequate, as almost all curves show approximate saturation.

TABLE 4 Alpha diversity analyses.

Sample ID	Culture medium	Condition	Obs. OTUs	NpShannon	InvSimpson
Total phyllospheric bacterial communities					
P1	N/A	Polluted	805.0 ± 0.0	5.36 ± 0.00	37.53 ± 0.00
P2	N/A	Polluted	268.4 ± 9.3	2.08 ± 0.04	2.05 ± 0.03
C1	N/A	Control	858.5 ± 14.1	5.34 ± 0.03	29.50 ± 0.96
C2	N/A	Control	719.7 ± 14.0	5.07 ± 0.03	25.71 ± 0.86
Culturable phyllospheric bacterial communities					
P.Ia	LB	Polluted	57.8 ± 1.7	2.49 ± 0.02	6.55 ± 0.09
P.Ib	LB	Polluted	30.1 ± 2.9	1.13 ± 0.04	1.76 ± 0.04
C.Ia	LB	Control	16.9 ± 1.7	0.54 ± 0.03	1.23 ± 0.02
C.Ib	LB	Control	13.8 ± 1.5	0.84 ± 0.02	1.72 ± 0.03
P.IIa	1/10 868	Polluted	40.8 ± 3.5 *	0.81 ± 0.04 *	1.29 ± 0.02 *
P.IIb	1/10 869	Polluted	62.7 ± 2.7 *	2.46 ± 0.03 *	6.35 ± 0.11 *
C.IIa	1/10 869	Control	50.7 ± 2.6 *	1.68 ± 0.03 *	2.30 ± 0.05 *
C.IIb	1/10 869	Control	51.9 ± 2.5 *	2.19 ± 0.03 *	4.38 ± 0.09 *
P.IIIa	YMA	Polluted	65.3 ± 1.3	2.35 ± 0.01	5.10 ± 0.05
P.IIIb	YMA	Polluted	56.9 ± 2.9	2.49 ± 0.03	6.77 ± 0.14
C.IIIa	YMA	Control	53.6 ± 2.2	1.66 ± 0.02	2.58 ± 0.04
C.IIIb	YMA	Control	21.2 ± 1.5	0.72 ± 0.03	1.31 ± 0.02
P.IVa	Flour 1	Polluted	26.2 ± 2.2 *	0.99 ± 0.03 *	1.83 ± 0.04 *
P.IVb	Flour 1	Polluted	14.0 ± 1.6 *	1.02 ± 0.02 *	2.25 ± 0.03 *
C.IVa	Flour 1	Control	23.8 ± 2.1 *	0.60 ± 0.03 *	1.22 ± 0.01 *
C.IVb	Flour 1	Control	25.0 ± 1.5 *	0.77 ± 0.02 *	1.38 ± 0.02 *
P.Va	YMAb 284	Polluted	56.9 ± 3.1	1.76 ± 0.03	3.22 ± 0.05
P.Vb	YMAb 284	Polluted	50.5 ± 3.2	1.50 ± 0.04	2.01 ± 0.04
C.Va	YMAb 284	Control	28.8 ± 2.3	1.17 ± 0.03	2.06 ± 0.04
C.Vb	YMAb 284	Control	40.1 ± 2.6	1.60 ± 0.04	2.22 ± 0.05

Three alpha diversity indices (mean values ± SD; Obs. OTUs: observed amount of operational taxonomic units (OTUs), NpShannon: nonparametric estimation of Shannon's diversity index, InvSimpson: inverse of Simpson's diversity index) were calculated at an equal sampling depth for total (n = 4; 3200 sequences per sample) and culturable (n = 20; 1200 sequences per sample) phyllospheric bacterial communities. *: $p < 0.05$; Kruskal-Wallis test.

The Yue & Clayton measure of dissimilarity was calculated for both total and culturable phyllospheric bacterial communities. NMDS was performed and in either case, no significant effect of the condition ("Polluted" *vs.* "Control") and/or the culture medium could be revealed (ANOSIM, $p < 0.05$).

Comparison of exclusive and shared OTUs between the total and culturable fraction showed that 82 out of 272 OTUs (30.1%) of total phyllospheric bacterial communities are also present in the culturable fraction (**FIGURE 6a**). Moreover, for both fractions, about 50% of all OTUs are shared between the polluted and control environment, indicating an equal important effect of the host plant species and the environment on phyllospheric bacterial community structure (**FIGURE 6b, c**).

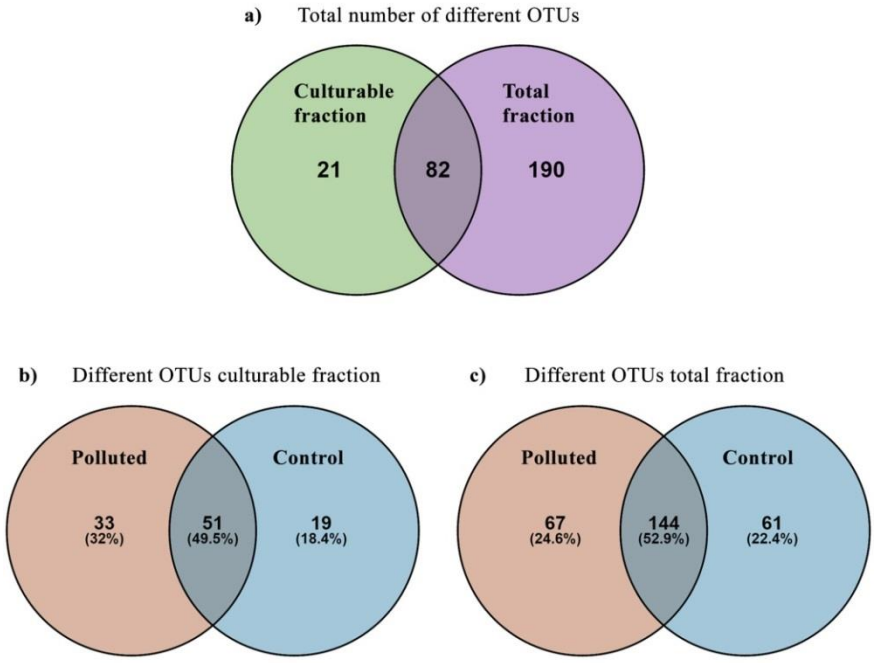

FIGURE 6 Venn diagrams of exclusive and shared operational taxonomic units (OTUs) between total and culturable phyllospheric bacterial communities (**a**) and between the polluted and control environment for the culturable (**b**) and total (**c**) fraction.

The top 10 phyla/classes of total phyllospheric bacterial communities (both the polluted and control environment) are Alphaproteobacteria (39.2%), Actinobacteria (13.1%), Bacteroidetes (12.8%), unclassified bacteria (8.9%), Gammaproteobacteria (7.4%), unclassified Proteobacteria (3.8%), Firmicutes (3.3%), Betaproteobacteria (2.7%), Acidobacteria (2.6%) and Planctomycetes (2.0%) (**FIGURE 7a**). At the genus level, *Hymenobacter* (4.5%), *Sphingomonas* (2.6%), *Methylobacterium* (1.5%) and *Spirosoma* (1.2%) are included in the top 20 OTUs (**FIGURE 9a**).

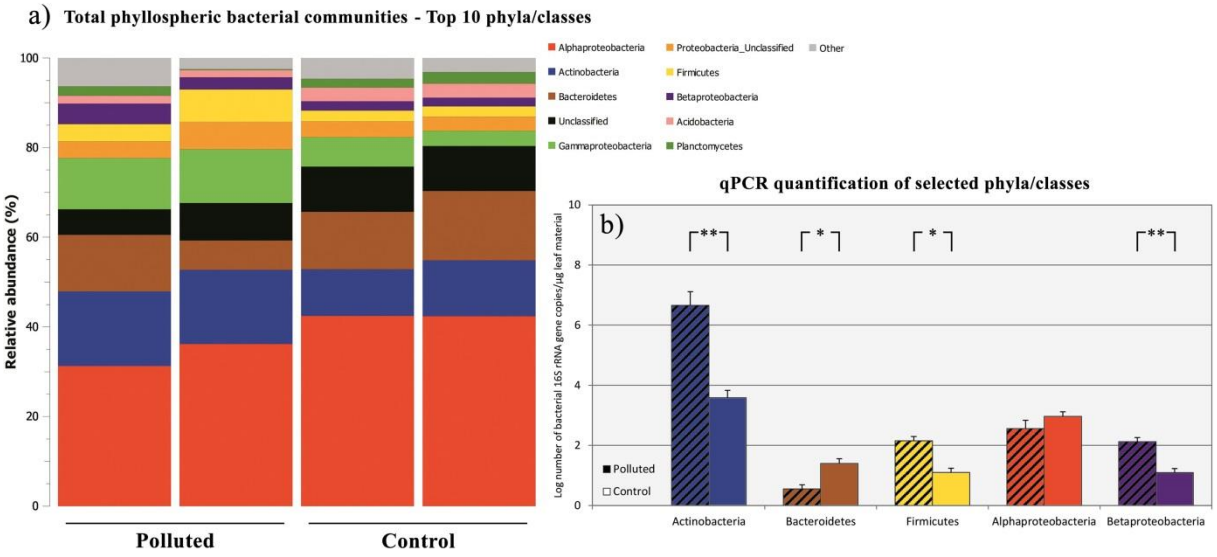

FIGURE 7 Top 10 phyla/classes as determined by 454 pyrosequencing (**a**; n = 4) and real-time PCR (qPCR) quantification of selected taxa (**b**; n = 12; mean values ± SD) in total phyllospheric bacterial communities. The legend (**a**) is constructed in order of overall taxon abundance (top: highest overall abundance, bottom: lowest overall abundance; from left to right). *: $p < 0.05$, **: $p < 0.01$; Wilcoxon rank-sum test.

Culturable phyllospheric bacterial communities (all culture media; both the polluted and control environment) were classified into four known phyla: (i) Proteobacteria (49.6%), with classes Gammaproteobacteria (20.6%), Alphaproteobacteria (20.5%) and Betaproteobacteria (8.5%), (ii) Actinobacteria (23.2%), (iii) Firmicutes (17.6%) and (iv) Bacteroidetes (4.2%) (**FIGURE 8**). At the genus level, *Bacillus* (4.9%), *Pseudomonas* (3.8%), *Burkholderia* (3.6%), *Stenotrophomonas* (2.6%), *Paenibacillus* (2.3%), *Sphingomonas* (2.1%), *Rhizobium* (2.0%), *Roseomonas* (2.0%) and *Lysinibacillus* (1.6%) are included in the top 20 OTUs (**FIGURE 9b**). In total, 5.4% of the cultured bacteria could not be classified. The relative abundances of the cultured phyla for all investigated culture media are summarized in **TABLE 5**. These results indicate that culturing using YMAb 284 medium is most representative of the bacterial community composition in total phyllospheric bacterial communities.

TABLE 5 Relative abundances of the cultured phyla for all investigated culture media.

Phylum	Relative abundance (%)				
	LB	1/10 869	YMA	Flour 1	YMAb 284
Actinobacteria	23.8	17.0	28.1	31.6	20.9
Bacteroidetes	1.4	1.5	7.2	2.6	7.1
Firmicutes	29.9	25.9	10.0	3.5	14.7
Proteobacteria	37.4	51.4	50.2	58.8	50.2
Alphaproteobacteria	10.2	22.8	17.2	25.4	25.3
Betaproteobacteria	2.0	10.8	11.3	16.7	3.1
Gammaproteobacteria	25.2	17.8	21.7	16.7	21.8
Unclassified	7.5	4.2	4.5	3.5	7.1

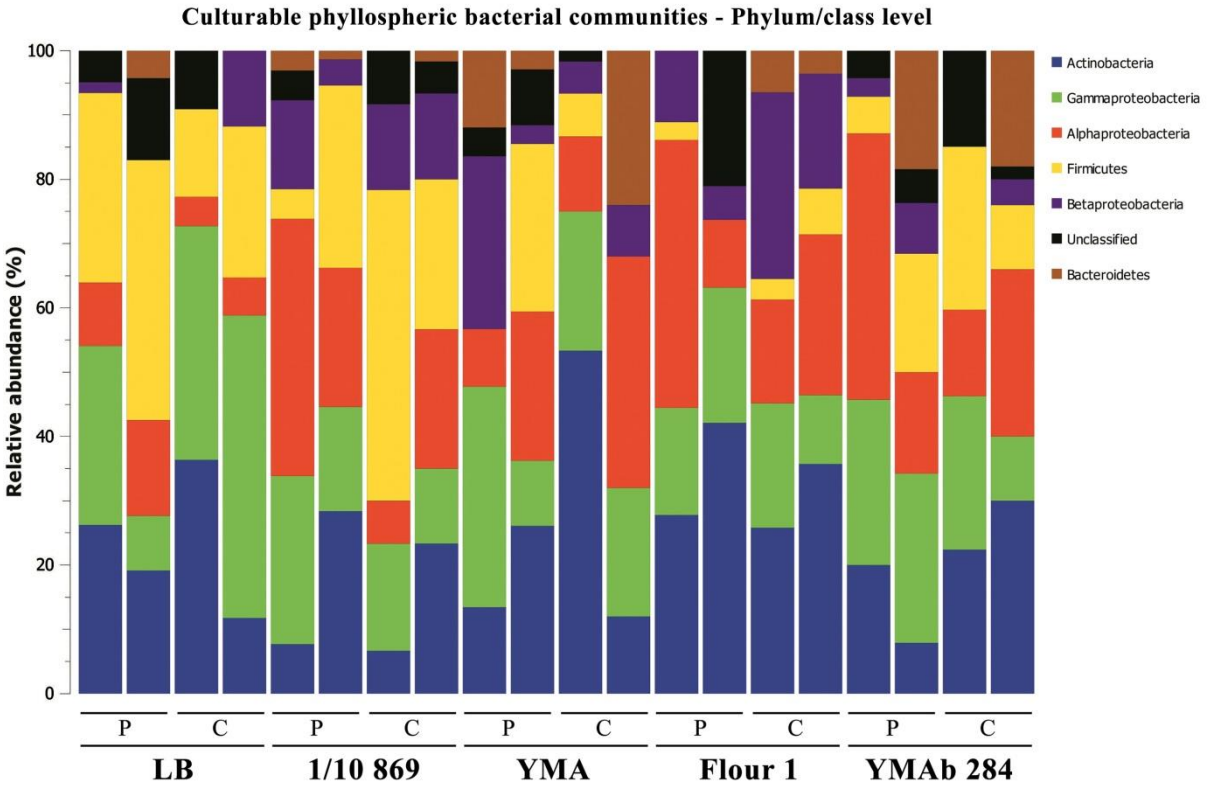

FIGURE 8 454 pyrosequencing to taxonomically identify culturable phyllospheric bacterial communities (n = 20). Cultured bacteria belonged to four known phyla: (i) Actinobacteria, (ii) Bacteroidetes, (iii) Firmicutes and (iv) Proteobacteria, with classes Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria. Not all cultured bacteria could be classified. P: polluted, C: control.

Bacteria and fungi were significantly more abundant in total phyllospheric microbial communities of the polluted environment, as determined by qPCR (Wilcoxon rank-sum test, $p < 0.05$; **FIGURE 10a, b**). For bacteria, this discrepancy was consistent for both used primer pairs targeting different regions of the 16S rRNA gene: Eub338F/Eub518R and Bact1369F/Prok1492R. However, qPCR quantification using the Eub338F/Eub518R primer pair indicated about a five-fold higher abundance of bacteria compared to the Bact1369F/Prok1492R primer pair (**FIGURE 10b**).

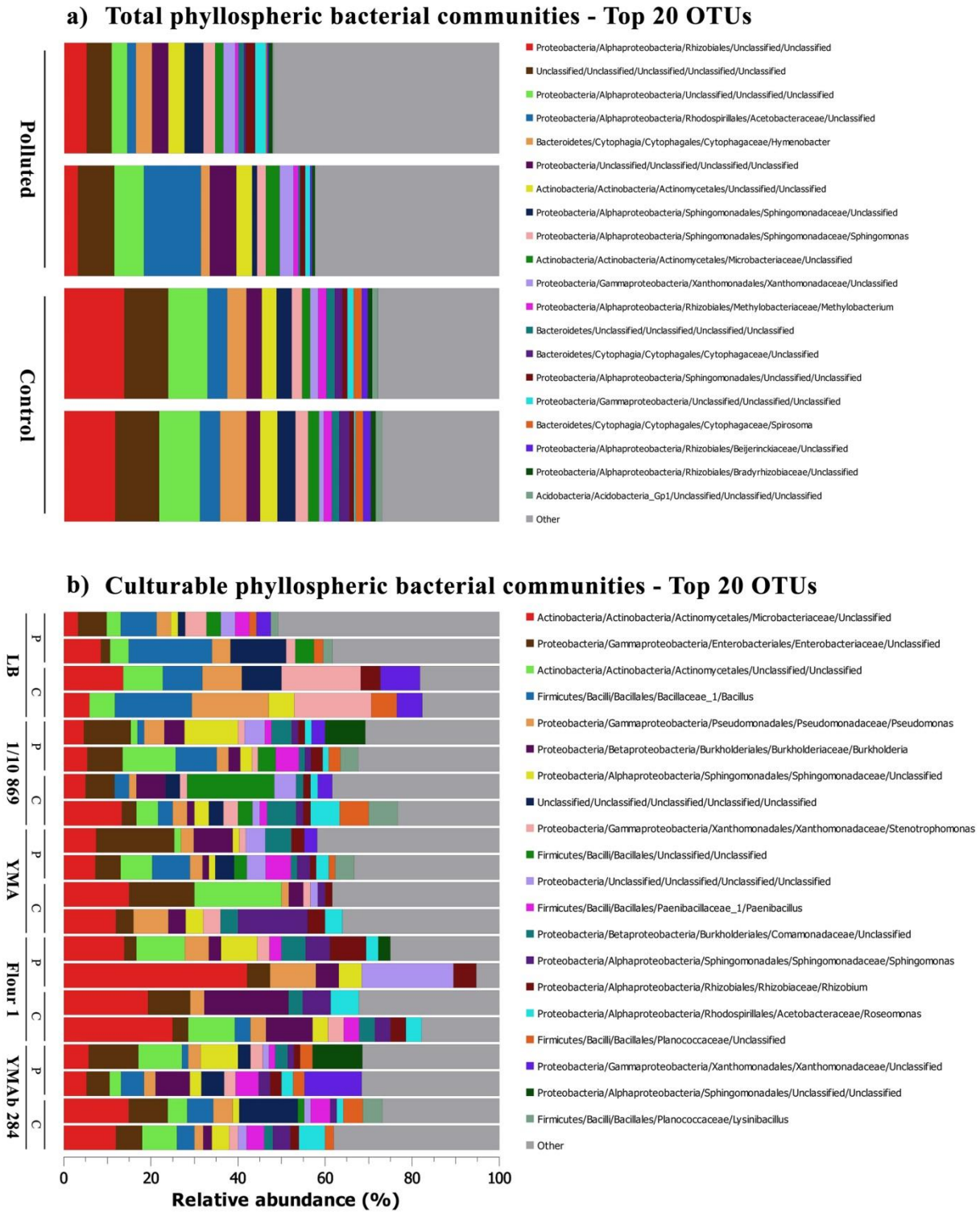

FIGURE 9 Top 20 operational taxonomic units (OTUs), as determined by 454 pyrosequencing, for total (**a**; n = 4) and culturable (**b**; n =20) phyllospheric bacterial communities. The legend is constructed in order of overall OTU abundance (top: highest overall abundance, bottom: lowest overall abundance). P: polluted, C: control.

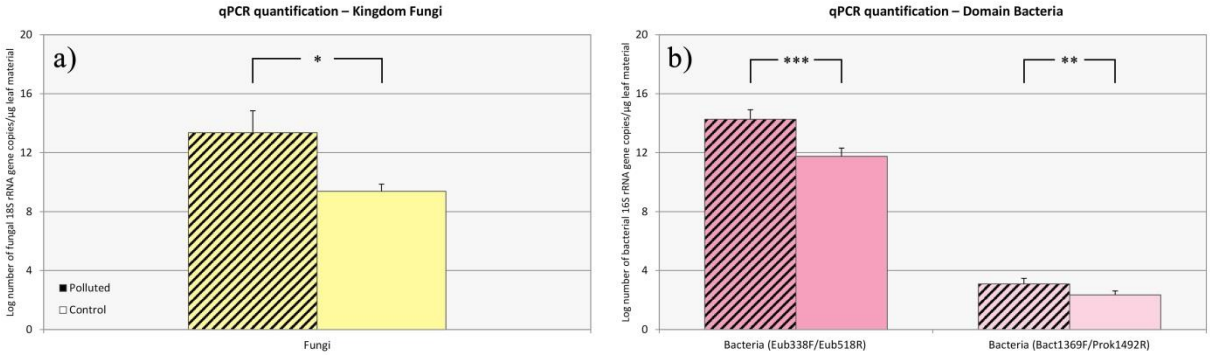

FIGURE 10 Real-time PCR (qPCR) quantification (mean values ± SD) of fungi (**a**; n = 12) and bacteria (**b**; Eub338F/Eub518R and Bact1369F/Prok1492R primer pairs; n = 12) in total phyllospheric microbial communities. Both taxa were more abundant in the polluted environment. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$; Wilcoxon rank-sum test.

For total bacterial communities, the relative abundances of Actinobacteria, Firmicutes and Betaproteobacteria tend to be higher in the polluted environment as determined by 454 pyrosequencing (**FIGURE 7a**), but this could not be statistically validated (LEfSe, $p < 0.05$; Wilcoxon rank-sum test, $p < 0.05$). However, qPCR quantification of the respective taxa indicated these differences are significant (Wilcoxon rank-sum test, $p < 0.05$; **FIGURE 7b**). The same is true for *Pseudomonas*; results of 454 pyrosequencing are indicative of a higher relative abundance in the polluted environment (**FIGURE 12a**). While this could not be statistically validated (LEfSe, $p < 0.05$; Wilcoxon rank-sum test, $p < 0.05$), qPCR quantification showed that *Pseudomonas* indeed is significantly more present in the polluted environment (Wilcoxon rank-sum test, $p < 0.01$; **FIGURE 12b**). 454 pyrosequencing also indicated phylum Bacteroidetes is relatively less abundant in the polluted environment (**FIGURE 7a**), but again this could not be statistically validated (LEfSe, $p < 0.05$; Wilcoxon rank-sum test, $p < 0.05$), while qPCR quantification showed this difference is significant (Wilcoxon rank-sum test, $p < 0.05$; **FIGURE 7b**). Further, 454 pyrosequencing indicated that Alphaproteobacteria are relatively most abundant in total bacterial communities (both the polluted and control environment; 39.2%), while qPCR quantification indicated that Actinobacteria are more abundant than Alphaproteobacteria (**FIGURE 7a, b**).

Significant differentially abundant taxa in culturable bacterial communities between the investigated culture media, and between the polluted and control environment, were identified. Betaproteobacteria were significantly more abundant on 1/10 869 media, *Bacillus* and Bacillaceae on LB media, and *Burkholderia*, Burkholderiaceae and unclassified genera of Microbacteriaceae on Flour 1 media (LEfSe, $p < 0.05$; **FIGURE 11a**). Further, Caulobacteraceae, Caulobacterales, *Bosea*, Bradyrhizobiaceae, *Achromobacter* and unclassified genera, families and orders of Alphaproteobacteria were significantly more cultured with samples from the polluted environment; for the control environment the same is true for *Variovorax*, *Staphylococcus* and Staphylococcaceae (LEfSe, $p < 0.05$; **FIGURE 11b**).

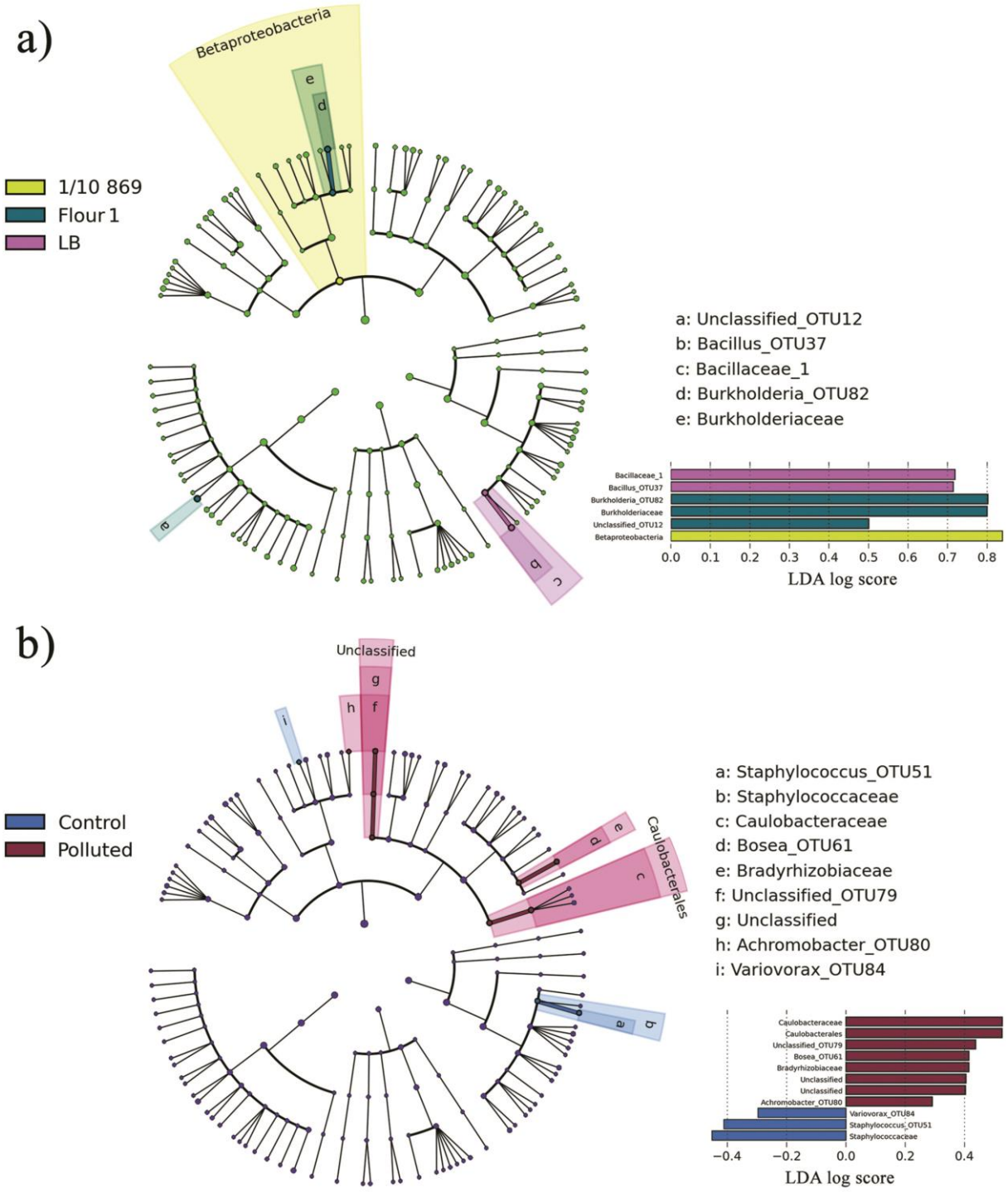

FIGURE 11 Graphical tree representations of significant differentially abundant taxa in culturable bacterial communities (n = 20) between the investigated culture media (**a**), and between the polluted and control environment (**b**), as identified by linear discriminative analysis (LDA) with effect size (LEfSe, $p < 0.05$). At the bottom right of each graphical tree, the LDA log score is presented for all significant differentially abundant taxa. The higher the absolute value of the LDA log score, the more differentially abundant the respective taxon.

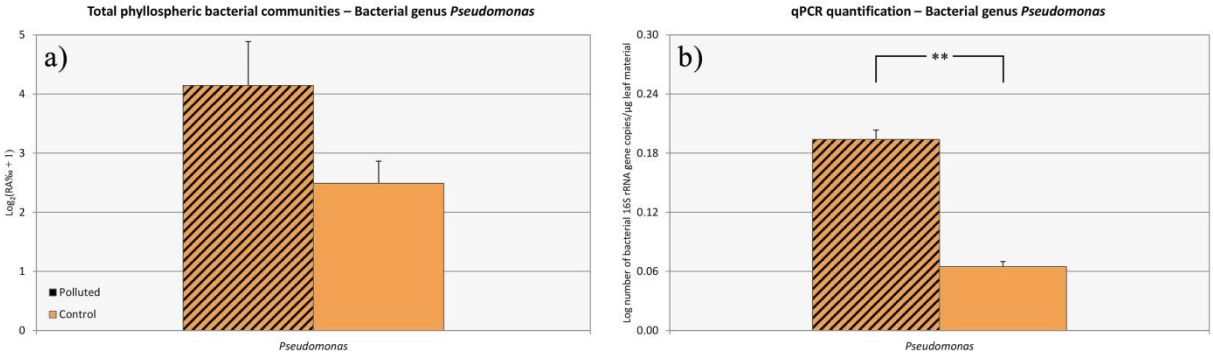

FIGURE 12 Relative abundance of *Pseudomonas* as determined by 454 pyrosequencing (**a**; n = 4) and real-time PCR (qPCR) quantification (**b**; n = 12; mean values ± SD) in total phyllospheric bacterial communities. 454 pyrosequencing results are indicative of *Pseudomonas* being more abundant in the polluted environment, as confirmed by qPCR quantification. **: $p < 0.01$; Wilcoxon rank-sum test.

Air pollution increases the prevalence of bacterial functional traits affiliated to plant growth and phytoremediation of airborne pollutants

Cultured phyllospheric bacteria (n = 192, 96 "Polluted", 96 "Control") were phenotypically investigated by performing analyses of PGP traits, trace metal (copper, zinc, cadmium) resistance and diesel fuel degradation capacity; the results are summarized in **TABLE 6**. A complete list of the phenotypically analyzed bacteria is presented in **TABLE A1**, **APPENDIX III – PHENOTYPICALLY ANALYZED BACTERIA**.

TABLE 6 Summary of the phenotypically analyzed phyllospheric bacteria.

Condition	Culture medium	Total number of isolates	Percentage of isolates that tested positively for the respective analysis						
			Cu res.	Zn res.	Cd res.	IAA	Acet.	ACCd	Diesel degr.
Polluted	LB	24	21	4	0	83	71	21	13
	1/10 869	24	0	0	0	75	29	50	4
	YECD	24	0	0	0	75	8	58	0
	YMAb 284	24	0	0	0	88	17	54	0
	Total	**96**	**5**	**1**	**0**	**80**	**31**	**46**	**4**
Control	LB	24	0	0	0	54	4	8	0
	1/10 869	24	0	0	0	63	17	25	0
	YECD	24	0	0	0	88	4	17	0
	YMAb 284	24	0	0	0	92	4	54	0
	Total	**96**	**0**	**0**	**0**	**74**	**7**	**26**	**0**

Cu, Zn, Cd res.: copper (0.8 mM), zinc (1.0 mM) and cadmium (0.8 mM) resistance, respectively. IAA: indole-3-acetic acid (IAA) production. Acet.: 3-hydroxy-2-butanone (acetoin) production. ACCd: 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase activity. Diesel degr.: diesel fuel degradation.

For the majority of the conducted analyses, more cultured phyllospheric bacteria from the polluted environment tested positively. Regarding the analyses of PGP traits, this difference is especially distinct for the assessment of acetoin production (31% *vs.* 7%) and ACC deaminase activity (46% *vs.* 26%), while IAA production was comparable (80% *vs.* 74%). Further, only bacteria from the polluted environment and isolated using LB and/or 1/10 869 media tested positively for copper and zinc resistance (5% and 1%, respectively; LB) and diesel fuel degradation (4%; LB and 1/10 869). Copper resistant bacteria belonged to the genera *Bacillus* (PhB4) and *Pantoea* (PhB10, PhB11, PhB19 and PhB20), while the zinc resistant bacterium belonged to the genus *Bacillus* (PhB4). Bacteria testing positively for diesel fuel degradation belonged to the genera *Bacillus* (PhB4, PhB12 and PhB51) and *Rhodococcus* (PhB6). PhB4 (*Bacillus*) had the most promising performance; this bacterium tested positively for 5 out of 7 analyses, while all other bacteria tested positively for maximum 3 out of 7 analyses.

Cultured phyllospheric bacteria were taxonomically identified by Sanger sequencing a large part of the 16S rRNA gene. After quality filtering of the sequences as previously described, 133 high-quality were recovered and a total of 26 different genera could be identified: *Bacillus*, *Brevibacterium*, *Caryophanon*, *Cellulomonas*, *Curtobacterium*, *Erwinia*, *Frigoribacterium*, *Frondihabitans*, *Leifsonia*, *Lysinibacillus*, *Methylobacterium*, *Microbacterium*, *Mucilaginibacter*, *Nocardioides*, *Novosphingobium*, *Paenibacillus*, *Pantoea*, *Patulibacter*, *Pedobacter*, *Pseudomonas*, *Rathayibacter*, *Rhodococcus*, *Roseomonas*, *Sphingomonas*, *Variovorax* and *Xanthomonas*. Interestingly, 44 out of 133 (33.1%) taxonomically identified bacteria were not previously cultured under laboratory conditions. Further, the V3-V4 hypervariable regions of the bacterial 16S rRNA gene of 112 out of 133 (84.2%) sequences matched perfectly (100% sequence identity) with sequences obtained by 454 pyrosequencing of total phyllospheric bacterial communities. In the case of culturable phyllospheric bacterial communities, the same was true for 94 out of 133 (70.7%) sequences.

DISCUSSION

In this study, we used complementary culture-dependent and culture-independent techniques to characterize the structure of phyllospheric microbial communities associated with *H. helix* growing at a polluted high-traffic urban area and a nonpolluted nature reserve. We combined the obtained phyllosphere microbiome data with an elaborate data set on cultured bacteria to (i) quantify and qualify the leaf-associated microbial biodiversity of *H. helix* at two environmentally distinct locations, (ii) identify potential interesting bacteria with regard to phytoremediation applications, (iii) understand relationships between microbial biodiversity, air pollution and traits of cultured bacterial isolates, and (iv) discuss their implications for our understanding of the use of plants and their bacteria to phytoremediate airborne pollutants.

The sampling locations, a polluted high-traffic urban area and a nonpolluted nature reserve, were carefully selected in order to be relevant in the context of phytoremediation of air pollution. While the specific environment of the locations is already indicative of this relevance, we consolidated this by measuring the UFP concentration and particle diameter in the ambient air and by SEM combined with EDX of collected leaves. Not unexpected, the UFP concentration was higher and the particle diameter was smaller at the high-traffic urban area. These observations can be mainly explained by the much higher prevalence of traffic-related combustion of fuels, such as diesel fuel (Kumar *et al.*, 2014). Moreover, a smaller particle diameter implies a higher toxicity in humans; this is because it increases the efficiency of lung tissue penetration and subsequently the entry in systemic circulation, significantly increasing the range of effect (Saravia *et al.*, 2013). SEM combined with EDX revealed the presence of more particulates and trace metals, including copper and zinc, on polluted adaxial leaf surfaces; the surfaces facing the surrounding environment. The fact that a higher UFP concentration was measured at the polluted environment, combined with the fact that *H. helix* has a high capacity to uptake particulates out of the air (Dzierżanowski *et al.*, 2011), could explain the higher amount of particulates on leaf surfaces of the polluted environment compared to the control environment. Moreover, a higher UFP concentration is also indicative of higher concentrations of bigger-sized particulates (Kumar *et al.*, 2014). The higher presence of copper and zinc can be linked to tire rubber abrasion and diesel fuel combustion at high-traffic areas, respectively, as this was shown to be causally related by Ozaki *et al.* (2004).

Microorganisms, mainly bacteria and fungi, were found readily dispersed on adaxial and abaxial surfaces of all 20 leaves inspected with SEM. No differences in microbial cell density and distribution between the polluted and control environment were found. However, to provide us of more solid answers on this topic, an increase of the sample size, a more standardized approach including appropriate statistical analyses, and the use of complementary techniques such as fluorescence *in situ* hybridization (FISH) is recommended (Amann & Fuchs, 2008; Cardinale, 2014). Further, it was repeatedly observed that microorganisms clustered near stomata and on trichomes, clearly indicating a trend. This is supported by similar observations in *A. thaliana* (Remus-Emsermann *et al.*, 2014), and can be explained by the fact that the microenvironment of stomata and surface appendages such as

trichomes typically show an increased nutrient availability (Leveau & Lindow, 2001). While fungi dominated on trichomes, we observed mainly bacteria around stomata; the morphology of these bacteria was indicative of bacteria from the genus *Methylobacterium*, well-known to utilize methanol emitted by stomata (Gourion *et al.*, 2006; Nemecek-Marshall *et al.*, 1995; Peyraud *et al.*, 2012). At the time of writing, FISH experiments using taxon-specific probes to confirm the presence of *Methylobacterium* around stomata are ongoing.

Total phyllospheric microbial communities associated with *H. helix* were characterized using 454 pyrosequencing and qPCR, two complementary culture-independent techniques. We found no significant differences in phyllosphere community diversity (alpha and beta diversity) between the polluted and control environment, as determined by 454 pyrosequencing, suggesting both environments have a comparable phyllosphere community diversity. However, these results could be undermined by the fact that a total sample size of only 4 was used for 454 pyrosequencing. Both environments can be considered highly dynamic. Except for the level of air pollution which was reasonably well identified, other important parameters influencing phyllosphere community diversity such as the co-presence of other microorganisms, flora, meso- and macrofauna, UV light exposure, (air) pollution, fertilization and water limitations were not characterized and likely differ in both environments (Müller & Ruppel, 2014). This makes the generation of hypotheses in this context somewhat difficult. However, qPCR quantification showed that bacteria and fungi are more abundant in the polluted environment. Quantity and diversity of microorganisms in a given community are not always positively correlated, but this could be indicative of a higher phyllosphere community diversity in the polluted environment. Further investigation with an increased sample size could provide us with more coherent answers on this topic.

Taxonomic analysis using 454 pyrosequencing revealed that total phyllospheric bacterial communities (both the polluted and control environment) are dominated by Alphaproteobacteria (39.2%), Actinobacteria (13.1%), Bacteroidetes (12.8%) and Gammaproteobacteria (7.4%), together comprising 72.5% of the total phyllospheric bacterial community. The relative abundances of these dominant taxa in the phyllosphere community of *H. helix* are similar to those of *A. thaliana*, as reviewed by Vorholt (2012). However, the phyllosphere community composition of *H. helix* markedly differs from the total phyllospheric bacterial community of *Trifolium repens* (white clover) and *Glycine max* (soybean), where Alphaproteobacteria are extremely abundant (up to 67% of the total phyllosphere community), and from *Oryza sativa* (rice), with a phyllomicrobiome dominated by Actinobacteria (up to 43% of the total phyllosphere community) (Delmotte *et al.*, 2009; Knief *et al.*, 2012). At the genus level, *Hymenobacter*, *Methylobacterium*, *Sphingomonas*, and *Spirosoma* are highly abundant. It is interesting to have *Methylobacterium* in this list, as it strengthens our previously stated hypothesis that *Methylobacterium* is abundantly present on leaf surfaces, presumably around stomata, as was observed with SEM. Comparing total phyllospheric bacterial communities between the polluted high-traffic urban area and the nonpolluted nature reserve, no significant differentially abundant taxa could be identified in case of 454 pyrosequencing. Two statistical methods were employed: the rather simplistic nonparametric Wilcoxon rank-sum test and the more sophisticated LEfSe method, specially

designed for providing statistically valid answers in this context. While it could be the case that there really are no differentially abundant taxa between these environments, the results of qPCR quantification indicate this is likely not the case. This culture-independent technique complementary to 454 pyrosequencing showed that phyla Actinobacteria and Firmicutes, class Betaproteobacteria and genus *Pseudomonas* are more abundant in the polluted environment, while for phylum Bacteroidetes the opposite is true. Most probably, it is the difference in sample size that underlies this discrepancy between the results of 454 pyrosequencing and qPCR; while for qPCR the total sample size was 12, it was only 4 in case of 454 pyrosequencing. Further investigation using appropriate sample sizes could abate this discrepancy. However, this statement is maybe somewhat too optimistic, as for example 454 pyrosequencing indicated that Alphaproteobacteria are relatively most abundant in total bacterial communities (both the polluted and control environment), while qPCR quantification indicated that Actinobacteria are most abundant. Further research is necessary in order to tackle such incoherencies. In this context, other molecular markers than the 16S rRNA gene could be used; an elaborate list of molecular markers used in microbial studies is reviewed by Liu *et al.* (2012). Also, other complementary approaches can be employed. At the time of writing, automated rRNA intergenic spacer analysis (ARISA) experiments using taxon-specific primers to investigate the abundance of selected taxa between the polluted and control environment are ongoing.

454 pyrosequencing technology was also employed to taxonomically identify the phyllospheric bacteria present on five different solid culture media (LB, 1/10 869, YMA, Flour 1 and YMAb 284), in order to evaluate these culture media for their potential to capture a substantial part (in terms of diversity) of the total phyllospheric bacterial community. Phyllosphere community richness as estimated by alpha diversity analyses was significantly higher on 1/10 869 medium and lower on Flour 1 medium, while no significant differences in beta diversity were observed between the investigated culture media or between the polluted and control environment. 1/10 869 medium can be considered a complex, rich culture medium that contains high amounts of a range of nutrients, while Flour 1 medium contains less diverse and lower amounts of nutrients. The richness (in terms of diversity of nutrients) of the culture medium is often positively correlated with the diversity of bacteria that can be isolated from environmental samples and also the ultimate culturability of bacterial species altogether (Eevers *et al.*, 2015), which could explain our findings. However, when using this as an argument it is difficult to explain the fact that community richness was not significantly higher on LB medium, even a richer culture medium compared to 1/10 869 medium.

Comparison of exclusive and shared OTUs between the total and culturable fraction showed that 30.1% of OTUs in total phyllospheric bacterial communities are also present in the culturable fraction. Although five different culture media were used, selected to cover a wide variety of bacterial taxa, and phyllosphere communities are typically much less diverse compared to for example rhizosphere communities (Vorholt, 2012), it is reasonable to state that we cultured a substantial part of the total phyllospheric bacterial community given the fact that consensus exists that 95-99.9% of bacteria from environmental samples has yet to be cultured (Ritz, 2007). Moreover, for both fractions, about 50% of all OTUs are shared

between the polluted and control environment, indicating an equal important effect of the host plant species and the environment on phyllospheric bacterial community structure. This finding is in partial agreement with the studies of de Oliveira Costa *et al.* (2012) and Kim *et al.* (2012), were the authors concluded that host plant species is the primary factor driving the composition of the phyllosphere microbiome. Further, taxonomic analysis showed that culturable bacterial communities (all culture media; both the polluted and control environment) could be classified into four known phyla: (i) Actinobacteria, (ii) Bacteroidetes, (iii) Firmicutes and (iv) Proteobacteria, with classes Alphaproteobacteria, Betaproteobacteria and Gammaproteobacteria. These are considered the four most commonly cultured phyla (Müller & Ruppel, 2014; Ritz, 2007; Tanaka *et al.*, 2014). At the genus level, the commonly cultured genera *Bacillus*, *Burkholderia*, *Lysinibacillus*, *Paenibacillus*, *Pseudomonas*, *Rhizobium*, *Roseomonas*, *Sphingomonas* and *Stenotrophomonas* were abundantly present on our media plates. Interestingly, not all cultured bacteria could be taxonomically classified; this number was the highest on LB media and the lowest on Flour 1 media, indicating that LB medium is a good choice in order to identify new phyllospheric bacteria. Culturing with YMAb 284 medium, a combination of YMA and 284 medium that was used for the first time in this study, was most representative of the total phyllospheric bacterial community composition, which makes this medium an interesting choice for future research. Further, LEfSe indicated Betaproteobacteria were significantly more abundant on 1/10 869 media, *Bacillus* and Bacillaceae on LB media, and *Burkholderia*, Burkholderiaceae and unclassified genera of Microbacteriaceae on Flour 1 media. Hence, for specific research objectives involving one of these taxa, the respective medium would be a good choice. Finally, Caulobacteraceae, Caulobacterales, *Bosea*, Bradyrhizobiaceae, *Achromobacter* and unclassified genera, families and orders of Alphaproteobacteria were significantly more cultured with samples from the polluted environment; for the control environment the same is true for *Variovorax*, *Staphylococcus* and Staphylococcaceae.

A collection of 192 cultured and taxonomically identified phyllospheric bacteria was phenotypically investigated by performing analyses of PGP traits (IAA production, acetoin production and ACC deaminase activity), trace metal (copper, zinc, cadmium) resistance and diesel fuel degradation capacity. A total of 26 different genera were isolated, and 33.1% of the bacteria were not previously cultured under laboratory conditions. Further, most of the sequences obtained by Sanger sequencing matched perfectly with sequences obtained by 454 pyrosequencing of total and culturable phyllospheric bacterial communities, indicating the sequencing efforts were adequate.

For all three conducted PGP trait analyses, more cultured phyllospheric bacteria from the polluted environment tested positively. This is supportive of the hypothesis that plants in a polluted environment attract microbial symbionts which enhance their ability to grow and survive in a situation that is rather challenging compared to a nonpolluted environment (Berendsen *et al.*, 2012; de Souza *et al.*, 2015). Copper and zinc resistant bacteria, and bacteria with diesel fuel degradation capacity were isolated only from the polluted environment. This could be explained by the much higher prevalence of copper, zinc and traffic-related combustion of diesel fuel in a high-traffic urban area (Kumar *et al.*, 2014;

Ozaki *et al.*, 2004). Copper and zinc resistance, and the ability to use diesel fuel as a carbon source likely constitute advantageous traits in such an environment, providing a platform for the process of natural selection. PhB4, a *Bacillus* bacterium, had the most promising performance, testing positively for 5 out of 7 conducted analyses, and therefore is an excellent candidate for further experiments in order to explore its beneficial effects in the context of phytoremediation of airborne pollutants. Testing of such interesting bacteria using high-performance liquid chromatography (HPLC) to check for degradation products of selected pollutants will be the next step. Also, whole genome shotgun sequencing using the Ion Torrent platform (Ion PGM Sequencer) and related data analyses as described by Thijs *et al.* (2014b) could provide us with fundamental insights into the genomic basis of pollutant-detoxifying properties.

In conclusion, we used complementary culture-independent techniques, 454 pyrosequencing and qPCR, to characterize the structure of phyllospheric microbial communities associated with *H. helix* growing at a polluted high-traffic urban area and a nonpolluted nature reserve. Using on-site ambient air measurements and SEM combined with EDX of collected leaves, we found a higher UFP concentration and smaller particle diameter in the polluted environment, and leaves in this environment had an increased amount of particulates and trace metals at their surfaces. While no significant differentially abundant taxa could be identified in case of 454 pyrosequencing, qPCR quantification indicated that phyla Actinobacteria and Firmicutes, class Betaproteobacteria and genus *Pseudomonas* are higher abundant in the polluted environment, while for phylum Bacteroidetes the opposite is true. For both the polluted and control environment, we found that the phyllosphere community of *H. helix* is similar to that of *A. thaliana*, dominated by Actinobacteria, Bacteroidetes, Alphaproteobacteria and Gammaproteobacteria. Further, using a culture-dependent approach, five different culture media were evaluated for their potential to capture a substantial part (in terms of diversity) of the total phyllospheric bacterial community. LB medium was found to be a good choice in order to identify new phyllospheric bacteria. YMAb 284 medium, a combination of YMA and 284 medium that was used for the first time in this study, constitutes an interesting choice for future research, as it was most representative of the total phyllospheric bacterial community composition. A collection of 192 cultured and taxonomically identified phyllospheric bacteria was phenotypically investigated by performing analyses of PGP traits, trace metal resistance and diesel fuel degradation capacity. Our results indicated that air pollution increases the prevalence of these bacterial functional traits affiliated to plant growth and phytoremediation of airborne pollutants. To the best of our knowledge, this study presents for the first time a detailed and integrated description of phyllospheric microbial communities associated with *H. helix* in the context of phytoremediation of air pollution.

REFERENCES

- Amann R., and Fuchs B.M. (2008). Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat. Rev. Microbiol.* 6, 339-348.
- Andersson A.F., Lindberg M., Jakobsson H., Bäckhed F., Nyrén P., and Engstrand L. (2008). Comparative analysis of human gut microbiota by barcoded pyrosequencing. *PLOS ONE* 3, e2836.
- Arenz B.E., Schlatter D.C., Bradeen J.M., and Kinkel L.L. (2015). Blocking primers reduce co-amplification of plant DNA when studying bacterial endophyte communities. *J. Microbiol. Methods* 117, 1-3.
- Asnicar F., Weingart G., Tickle T.L., Huttenhower C., and Segata N. (2015). Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ* 3, e1029.
- Barac T., Taghavi S., Borremans B., Provoost A., Oeyen L., Colpaert J.V., Vangronsveld J., and van der Lelie D. (2004). Engineered endophytic bacteria improve phytoremediation of water-soluble, volatile, organic pollutants. *Nat. Biotechnol.* 22, 583-588.
- Belimov A.A., Hontzeas N., Safronova V.I., Demchinskaya S.V., Piluzza G., Bullitta S., and Glick B.R. (2005). Cadmium-tolerant plant growth-promoting bacteria associated with the roots of Indian mustard (*Brassica juncea* L. Czern.). *Soil Biol. Biochem.* 37, 241-250.
- Berendsen R.L., Pieterse C.M., and Bakker P.A. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17, 478-486.
- Bergersen F.J. (1961). The growth of *Rhizobium* in synthetic media. *Aust. J. Biol. Sci.* 14, 349-360.
- Bergmark L., Poulsen P.H., Al-Soud W.A., Norman A., Hansen L.H., and Sørensen S.J. (2012). Assessment of the specificity of *Burkholderia* and *Pseudomonas* qPCR assays for detection of these genera in soil using 454 pyrosequencing. *FEMS Microbiol. Lett.* 333, 77-84.
- Bertani G. (1951). Studies on lysogenesis. I. The mode of phage liberation by lysogenic *Escherichia coli*. *J. Bacteriol.* 62, 293-300.
- Bokulich N.A., Subramanian S., Faith J.J., Gevers D., Gordon J.I., Knight R., Mills D.A., and Caporaso J.G. (2013). Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* 10, 57-59.
- Bringel F., and Couée I. (2015). Pivotal roles of phyllosphere microorganisms at the interface between plant functioning and atmospheric trace gas dynamics. *Front. Microbiol.* 6, 486.
- Bučková M., Puškarová A., Chovanová K., Kraková L., Ferianc P., and Pangallo D. (2013). A simple strategy for investigating the diversity and hydrocarbon degradation abilities of cultivable bacteria from contaminated soil. *World J. Microbiol. Biotechnol.* 29, 1085-1098.

- Bulgarelli D., Rott M., Schlaeppi K., Ver Loren van Themaat E., Ahmadinejad N., Assenza F., Rauf P., Huettel B., Reinhardt R., Schmelzer E., Peplies J., Gloeckner F.O., Amann R., Eickhorst T., and Schulze-Lefert P. (2012). Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature* 488, 91-95.
- Caporaso J.G., Kuczynski J., Stombaugh J., Bittinger K., Bushman F.D., Costello E.K., Fierer N., Pena A.G., Goodrich J.K., Gordon J.I., Huttley G.A., Kelley S.T., Knights D., Koenig J.E., Ley R.E., Lozupone C.A., McDonald D., Muegge B.D., Pirrung M., Reeder J., Sevinsky J.R., Turnbaugh P.J., Walters W.A., Widmann J., Yatsunenko T., Zaneveld J., and Knight R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335-336.
- Cardinale M. (2014). Scanning a microhabitat: plant-microbe interactions revealed by confocal laser microscopy. *Front. Microbiol.* 5, 94.
- Chakravorty S., Helb D., Burday M., Connell N., and Alland D. (2007). A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* 69, 330-339.
- Chao A., and Shen T.J. (2003). Nonparametric estimation of Shannon's index of diversity when there are unseen species in sample. *Environ. Ecol. Stat.* 10, 429-443.
- Coombs J.T., and Franco C.M. (2003). Isolation and identification of Actinobacteria from surface-sterilized wheat roots. *Appl. Environ. Microbiol.* 69, 5603-5608.
- de Oliveira Costa L.E., de Queiroz M.V., Borges A.C., de Moraes C.A., and de Araújo E.F. (2012). Isolation and characterization of endophytic bacteria isolated from the leaves of the common bean (*Phaseolus vulgaris*). *Braz. J. Microbiol.* 43, 1562-1575.
- de Souza R., Ambrosini A., and Passaglia L.M. (2015). Plant growth-promoting bacteria as inoculants in agricultural soils. *Genet. Mol. Biol.* 38, 401-419.
- Dela Cruz M., Christensen J.H., Thomsen J.D., and Müller R. (2014). Can ornamental potted plants remove volatile organic compounds from indoor air? — a review. *Environ. Sci. Pollut. Res. Int.* 21, 13909-13928.
- Delmotte N., Knief C., Chaffron S., Innerebner G., Roschitzki B., Schlapbach R., von Mering C., and Vorholt J.A. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 106, 16428-16433.
- Dzierżanowski K., Popek R., Gawrońska H., Sæbø A., and Gawroński S.W. (2011). Deposition of particulate matter of different size fractions on leaf surfaces and in waxes of urban forest species. *Int. J. Phytoremediation* 13, 1037-1046.
- Edwards J., Johnson C., Santos-Medellín C., Lurie E., Podishetty N.K., Bhatnagar S., Eisen J.A., and Sundaresan V. (2015). Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl. Acad. Sci. U.S.A.* 112, E911-920.
- Eevers N., Gielen M., Sánchez-López A., Jaspers S., White J.C., Vangronsveld J., and Weyens N. (2015). Optimization of isolation and cultivation of bacterial endophytes through addition of plant extract to nutrient media. *Microb. Biotechnol.* 8, 707-715.
- Fierer N., Jackson J.A., Vilgalys R., and Jackson R.B. (2005). Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl. Environ. Microbiol.* 71, 4117-4120.

- Finkel O.M., Burch A.Y., Lindow S.E., Post A.F., and Belkin S. (2011). Geographical location determines the population structure in phyllosphere microbial communities of a salt-excreting desert tree. *Appl. Environ. Microbiol.* 77, 7647-7655.
- Fox G.E., Stackebrandt E., Hespell R.B., Gibson J., Maniloff J., Dyer T.A., Wolfe R.S., Balch W.E., Tanner R.S., Magrum L.J., Zablen L.B., Blakemore R., Gupta R., Bonen L., Lewis B.J., Stahl D.A., Luehrsen K.R., Chen K.N., and Woese C.R. (1980). The phylogeny of prokaryotes. *Science* 209, 457-463.
- Germaine K., Keogh E., Garcia-Cabellos G., Borremans B., van der Lelie D., Barac T., Oeyen L., Vangronsveld J., Moore F.P., Moore E.R., Campbell C.D., Ryan D., and Dowling D.N. (2004). Colonisation of poplar trees by *gfp* expressing bacterial endophytes. *FEMS Microbiol. Ecol.* 48, 109-118.
- Gkorezis P., Rineau F., Van Hamme J., Franzetti A., Daghio M., Thijs S., Weyens N., and Vangronsveld J. (2015). Draft genome sequence of *Acinetobacter oleivorans* PF1, a diesel-degrading and plant-growth-promoting endophytic strain isolated from poplar trees growing on a diesel-contaminated plume. *Genome Announc.* 3, e01430-01414.
- Glick B.R. (2014). Bacteria with ACC deaminase can promote plant growth and help to feed the world. *Microbiol. Res.* 169, 30-39.
- Gonzalez A., and Knight R. (2012). Advancing analytical algorithms and pipelines for billions of microbial sequences. *Curr. Opin. Biotechnol.* 23, 64-71.
- Gourion B., Rossignol M., and Vorholt J.A. (2006). A proteomic study of *Methylobacterium extorquens* reveals a response regulator essential for epiphytic growth. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13186-13191.
- Hamady M., and Knight R. (2009). Microbial community profiling for human microbiome projects: Tools, techniques, and challenges. *Genome Res.* 19, 1141-1152.
- Jackson C.R., and Denney W.C. (2011). Annual and seasonal variation in the phyllosphere bacterial community associated with leaves of the southern Magnolia (*Magnolia grandiflora*). *Microb. Ecol.* 61, 113-122.
- Jiménez M., Gatica A., Sánchez E., and Valdez M. (2012). Ultrastructural analysis of the ontogenetic development of shoot induced from embryonic axes of Costa Rican bean varieties (*Phaseolus vulgaris* L.) under *in vitro* conditions by scanning electronic microscopy. *Am. J. Plant Sci.* 3, 489-494.
- Kearse M., Moir R., Wilson A., Stones-Havas S., Cheung M., Sturrock S., Buxton S., Cooper A., Markowitz S., Duran C., Thierer T., Ashton B., Meintjes P., and Drummond A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647-1649.
- Kelly F.J., and Fussell J.C. (2012). Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmos. Environ.* 60, 504-526.
- Kim M., Singh D., Lai-Hoe A., Go R., Abdul Rahim R., Ainuddin A.N., Chun J., and Adams J.M. (2012). Distinctive phyllosphere bacterial communities in tropical trees. *Microb. Ecol.* 63, 674-681.
- Kim N.D., and Fergusson J.E. (1994). The concentrations, distribution and sources of cadmium, copper, lead and zinc in the atmosphere of an urban environment. *Sci. Total Environ.* 144, 179-189.

- Klindworth A., Pruesse E., Schweer T., Peplies J., Quast C., Horn M., and Glöckner F.O. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, e1.
- Knief C., Delmotte N., Chaffron S., Stark M., Innerebner G., Wassmann R., von Mering C., and Vorholt J.A. (2012). Metaproteogenomic analysis of microbial communities in the phyllosphere and rhizosphere of rice. *ISME J.* 6, 1378-1390.
- Kõiv V., Roosaare M., Vedler E., Kivistik P.A., Toppi K., Schryer D.W., Remm M., Tenson T., and Mäe A. (2015). Microbial population dynamics in response to *Pectobacterium atrosepticum* infection in potato tubers. *Sci. Rep.* 5, 11606.
- Kumar P., Morawska L., Birmili W., Paasonen P., Hu M., Kulmala M., Harrison R.M., Norford L., and Britter R. (2014). Ultrafine particles in cities. *Environ. Int.* 66, 1-10.
- Lelieveld J., Evans J.S., Fnais M., Giannadaki D., and Pozzer A. (2015). The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* 525, 367-371.
- Leveau J.H., and Lindow S.E. (2001). Appetite of an epiphyte: quantitative monitoring of bacterial sugar consumption in the phyllosphere. *Proc. Natl. Acad. Sci. U.S.A.* 98, 3446-3453.
- Lindow S.E., and Brandl M.T. (2003). Microbiology of the phyllosphere. *Appl. Environ. Microbiol.* 69, 1875-1883.
- Liu W., Li L., Khan M.A., and Zhu F. (2012). Popular molecular markers in bacteria. *Mol. Genet. Microbiol. Virol.* 27, 103-107.
- Lundberg D.S., Yourstone S., Mieczkowski P., Jones C.D., and Dangl J.L. (2013). Practical innovations for high-throughput amplicon sequencing. *Nat. Methods* 10, 999-1002.
- Maughan H., Wang P.W., Diaz Caballero J., Fung P., Gong Y., Donaldson S.L., Yuan L., Keshavjee S., Zhang Y., Yau Y.C., Waters V.J., Tullis D.E., Hwang D.M., and Guttman D.S. (2012). Analysis of the cystic fibrosis lung microbiota via serial Illumina sequencing of bacterial 16S rRNA hypervariable regions. *PLOS ONE* 7, e45791.
- Mergeay M., Nies D., Schlegel H.G., Gerits J., Charles P., and Van Gijsegem F. (1985). *Alcaligenes eutrophus* CH34 is a facultative chemolithotroph with plasmid-bound resistance to heavy metals. *J. Bacteriol.* 162, 328-334.
- Metcalfe D.J. (2005). *Hedera helix* L. *J. Ecol.* 93, 632-648.
- Metzker M.L. (2010). Sequencing technologies - the next generation. *Nat. Rev. Genet.* 11, 31-46.
- Müller H., Berg C., Landa B.B., Auerbach A., Moissl-Eichinger C., and Berg G. (2015). Plant genotype-specific archaeal and bacterial endophytes but similar *Bacillus* antagonists colonize Mediterranean olive trees. *Front. Microbiol.* 6, 138.
- Müller T., and Ruppel S. (2014). Progress in cultivation-independent phyllosphere microbiology. *FEMS Microbiol. Ecol.* 87, 2-17.

- Nemecek-Marshall M., MacDonald R.C., Franzen J.J., Wojciechowski C.L., and Fall R. (1995). Methanol emission from leaves (enzymatic detection of gas-phase methanol and relation of methanol fluxes to stomatal conductance and leaf development). *Plant Physiol.* 108, 1359-1368.
- Nurmatov U.B., Tagiyeva N., Semple S., Devereux G., and Sheikh A. (2015). Volatile organic compounds and risk of asthma and allergy: a systematic review. *Eur. Respir. Rev.* 24, 92-101.
- Oliveros J.C. (2015). Venny. An interactive tool for comparing lists with Venn's diagrams.
- Ozaki H., Watanabe I., and Kuno K. (2004). Investigation of the heavy metal sources in relation to automobiles. *Water Air Soil Pollut.* 157, 209-223.
- Patten C.L., and Glick B.R. (2002). Role of *Pseudomonas putida* indoleacetic acid in development of the host plant root system. *Appl. Environ. Microbiol.* 68, 3795-3801.
- Paulson J.N., Stine O.C., Bravo H.C., and Pop M. (2013). Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* 10, 1200-1202.
- Peterson J., Garges S., Giovanni M., McInnes P., Wang L., Schloss J.A., Bonazzi V., McEwen J.E., Wetterstrand K.A., Deal C., Baker C.C., Di Francesco V., Howcroft T.K., Karp R.W., Lunsford R.D., Wellington C.R., Belachew T., Wright M., Giblin C., David H., Mills M., Salomon R., Mullins C., Akolkar B., Begg L., Davis C., Grandison L., Humble M., Khalsa J., Little A.R., Peavy H., Pontzer C., Portnoy M., Sayre M.H., Starke-Reed P., Zakhari S., Read J., Watson B., and Guyer M. (2009). The NIH Human Microbiome Project. *Genome Res.* 19, 2317-2323.
- Petrosino J.F., Highlander S., Luna R.A., Gibbs R.A., and Versalovic J. (2009). Metagenomic pyrosequencing and microbial identification. *Clin. Chem.* 55, 856-866.
- Peyraud R., Kiefer P., Christen P., Portais J.C., and Vorholt J.A. (2012). Co-consumption of methanol and succinate by *Methylobacterium extorquens* AM1. *PLOS ONE* 7, e48271.
- Pfeiffer S., Pastar M., Mitter B., Lippert K., Hackl E., Lojan P., Oswald A., and Sessitsch A. (2014). Improved group-specific primers based on the full SILVA 16S rRNA gene reference database. *Environ. Microbiol.* 16, 2389-2407.
- Pilon-Smits E. (2005). Phytoremediation. *Annu. Rev. Plant. Biol.* 56, 15-39.
- Rastogi G., Sbodio A., Tech J.J., Suslow T.V., Coaker G.L., and Leveau J.H. (2012). Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. *ISME J.* 6, 1812-1822.
- Redford A.J., Bowers R.M., Knight R., Linhart Y., and Fierer N. (2010). The ecology of the phyllosphere: geographic and phylogenetic variability in the distribution of bacteria on tree leaves. *Environ. Microbiol.* 12, 2885-2593.
- Redford A.J., and Fierer N. (2009). Bacterial succession on the leaf surface: a novel system for studying successional dynamics. *Microb. Ecol.* 58, 189-198.
- Reinhold-Hurek B., and Hurek T. (2011). Living inside plants: bacterial endophytes. *Curr. Opin. Plant Biol.* 14, 435-443.

- Remus-Emsermann M.N., Lücker S., Müller D.B., Potthoff E., Daims H., and Vorholt J.A. (2014). Spatial distribution analyses of natural phyllosphere-colonizing bacteria on *Arabidopsis thaliana* revealed by fluorescence *in situ* hybridization. *Environ. Microbiol.* 16, 2329-2340.
- Ritz K. (2007). The plate debate: cultivable communities have no utility in contemporary environmental microbial ecology. *FEMS Microbiol. Ecol.* 60, 358-362.
- Robertson C.E., Harris J.K., Wagner B.D., Granger D., Browne K., Tatem B., Feazel L.M., Park K., Pace N.R., and Frank D.N. (2013). Explicet: graphical user interface software for metadata-driven management, analysis and visualization of microbiome data. *Bioinformatics* 29, 3100-3101.
- Romick T.L., and Fleming H.P. (1998). Acetoin production as an indicator of growth and metabolic inhibition of *Listeria monocytogenes*. *J. Appl. Microbiol.* 84, 18-24.
- Ryu C.M., Farag M.A., Hu C.H., Reddy M.S., Kloepper J.W., and Pare P.W. (2004). Bacterial volatiles induce systemic resistance in *Arabidopsis*. *Plant Physiol.* 134, 1017-1026.
- Ryu C.M., Farag M.A., Hu C.H., Reddy M.S., Wei H.X., Pare P.W., and Kloepper J.W. (2003). Bacterial volatiles promote growth in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4927-4932.
- Sæbø A., Popek R., Nawrot B., Hanslin H.M., Gawrońska H., and Gawroński S.W. (2012). Plant species differences in particulate matter accumulation on leaf surfaces. *Sci. Total. Environ.* 427-428, 347-354.
- Sanger F., and Coulson A.R. (1975). A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 94, 441-448.
- Saravia J., Lee G.I., Lomnicki S.M., Dellinger B., and Cormier S.A. (2013). Particulate matter containing environmentally persistent free radicals and adverse infant respiratory health effects: a review. *J. Biochem. Mol. Toxicol.* 27, 56-68.
- Saravia J., You D., Thevenot P., Lee G.I., Shrestha B., Lomnicki S.M., and Cormier S.A. (2014). Early-life exposure to combustion-derived particulate matter causes pulmonary immunosuppression. *Mucosal Immunol.* 7, 694-704.
- Schlegel H.G., Kaltwasser H., and Gottschalk G. (1961). A submersion method for culture of hydrogen-oxidizing bacteria: growth physiological studies. *Arch. Microbiol.* 38, 209-222.
- Schloss P.D., Gevers D., and Westcott S.L. (2011). Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLOS ONE* 6, e27310.
- Schloss P.D., Westcott S.L., Ryabin T., Hall J.R., Hartmann M., Hollister E.B., Lesniewski R.A., Oakley B.B., Parks D.H., Robinson C.J., Sahl J.W., Stres B., Thallinger G.G., Van Horn D.J., and Weber C.F. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537-7541.
- Schmidt M.A., Souza E.M., Baura V., Wassem R., Yates M.G., Pedrosa F.O., and Monteiro R.A. (2011). Evidence for the endophytic colonization of *Phaseolus vulgaris* (common bean) roots by the diazotroph *Herbaspirillum seropedicae*. *Braz. J. Med. Biol. Res.* 44, 182-185.

- Segata N., Izard J., Waldron L., Gevers D., Miropolsky L., Garrett W.S., and Huttenhower C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12, R60.
- Shi Y., Yang H., Zhang T., Sun J., and Lou K. (2014). Illumina-based analysis of endophytic bacterial diversity and space-time dynamics in sugar beet on the north slope of Tianshan mountain. *Appl. Microbiol. Biotechnol.* 98, 6375-6385.
- Simpson E.H. (1949). Measurement of diversity. *Nature* 163, 688.
- Sternberg T., Viles H., Cathersides A., and Edwards M. (2010). Dust particulate absorption by ivy (*Hedera helix* L.) on historic walls in urban environments. *Sci. Total Environ.* 409, 162-168.
- Suzuki M.T., Taylor L.T., and DeLong E.F. (2000). Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl. Environ. Microbiol.* 66, 4605-4614.
- Tanaka T., Kawasaki K., Daimon S., Kitagawa W., Yamamoto K., Tamaki H., Tanaka M., Nakatsu C.H., and Kamagata Y. (2014). A hidden pitfall in the preparation of agar media undermines microorganism cultivability. *Appl. Environ. Microbiol.* 80, 7659-7666.
- Thijs S., Van Dillewijn P., Sillen W., Truyens S., Holtappels M., D'Haen J., Carleer R., Weyens N., Ameloot M., Ramos J.L., and Vangronsveld J. (2014a). Exploring the rhizospheric and endophytic bacterial communities of *Acer pseudoplatanus* growing on a TNT-contaminated soil: towards the development of a rhizocompetent TNT-detoxifying plant growth promoting consortium. *Plant Soil* 385, 15-36.
- Thijs S., Van Hamme J., Gkorezis P., Rineau F., Weyens N., and Vangronsveld J. (2014b). Draft genome sequence of *Raoultella ornithinolytica* TNT, a trinitrotoluene-denitrating and plant growth-promoting strain isolated from explosive-contaminated soil. *Genome Announc.* 2, e00491-00414.
- Vangronsveld J., Herzig R., Weyens N., Boulet J., Adriaensen K., Ruttens A., Thewys T., Vassilev A., Meers E., Nehnevajova E., van der Lelie D., and Mench M. (2009). Phytoremediation of contaminated soils and groundwater: lessons from the field. *Environ. Sci. Pollut. Res. Int.* 16, 765-794.
- Vorholt J.A. (2012). Microbial life in the phyllosphere. *Nat. Rev. Microbiol.* 10, 828-840.
- Voriskova J., and Baldrian P. (2013). Fungal community on decomposing leaf litter undergoes rapid successional changes. *ISME J.* 7, 477-486.
- Waldor M.K., Tyson G., Borenstein E., Ochman H., Moeller A., Finlay B.B., Kong H.H., Gordon J.I., Nelson K.E., Dabbagh K., and Smith H. (2015). Where next for microbiome research? *PLOS Biol.* 13, e1002050.
- Wang Y., and Qian P.Y. (2009). Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLOS ONE* 4, e7401.
- Weyens N., Thijs S., Popek R., Witters N., Przybysz A., Espenshade J., Gawrońska H., Vangronsveld J., and Gawroński S.W. (2015). The role of plant-microbe interactions and their exploitation for phytoremediation of air pollutants. *Int. J. Mol. Sci.* 16, 25576-25604.

- Weyens N., van der Lelie D., Artois T., Smeets K., Taghavi S., Newman L., Carleer R., and Vangronsveld J. (2009a). Bioaugmentation with engineered endophytic bacteria improves contaminant fate in phytoremediation. *Environ. Sci. Technol.* 43, 9413-9418.
- Weyens N., van der Lelie D., Taghavi S., Newman L., and Vangronsveld J. (2009b). Exploiting plant-microbe partnerships to improve biomass production and remediation. *Trends Biotechnol.* 27, 591-598.
- Yashiro E., Spear R.N., and McManus P.S. (2011). Culture-dependent and culture-independent assessment of bacteria in the apple phyllosphere. *J. Appl. Microbiol.* 110, 1284-1296.
- Yilmaz P., Gilbert J.A., Knight R., Amaral-Zettler L., Karsch-Mizrachi I., Cochrane G., Nakamura Y., Sansone S.A., Glockner F.O., and Field D. (2011). The genomic standards consortium: bringing standards to life for microbial ecology. *ISME J.* 5, 1565-1567.
- Yue J.C., and Clayton M.K. (2005). A similarity measure based on species proportions. *Commun. Stat. Theor. M.* 34, 2123-2131.
- Zhang K., and Batterman S. (2013). Air pollution and health risks due to vehicle traffic. *Sci. Total Environ.* 450-451, 307-316.
- Zhao Y. (2010). Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* 61, 49-64.
- Zhu F., Massana R., Not F., Marie D., and Vaulot D. (2005). Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52, 79-92.

APPENDIX I – REAL-TIME PCR STANDARD CURVES

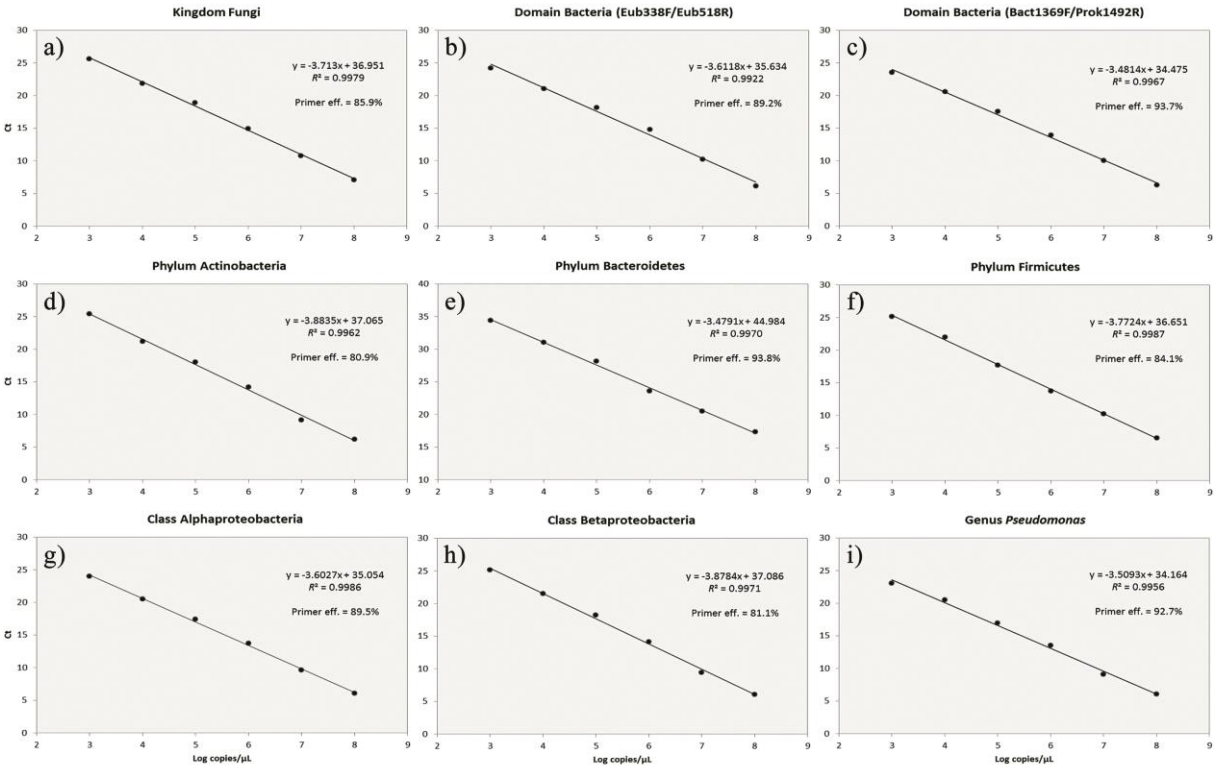

FIGURE A1 Standard curves of selected taxa (**a**: kingdom Fungi; **b**, **c**: domain Bacteria; **d**, **e**, **f**: bacterial phyla Actinobacteria, Bacteroidetes, Firmicutes; **g**, **h**: bacterial classes Alphaproteobacteria, Betaproteobacteria; **i**: bacterial genus *Pseudomonas*) were made with plasmid DNA concentrations ranging from 10^3 to 10^8 copies/µL in order to allow for absolute quantification using real-time PCR (qPCR). The qPCR setup for all standard curves was optimized in order to obtain a primer efficiency between 80-120% with linear resolution ($R^2 > 0.99$) over six orders of dynamic range. Ct: number of cycles to reach the threshold.

APPENDIX II – 454 PYROSEQUENCING DATA PROCESSING

This protocol was made by me, Vincent, and aims to help you to process your raw 454 pyrosequencing data (GS FLX Titanium) with mothur (v.1.36.1) to an OTU table that can be used for further analyses. Each step/algorithm is explained concisely. Although this can be considered a complete protocol, I strongly recommend to explore the different parameters, and give them your own flavor. Many roads lead to Rome.

1. Set the file directory where you put your sff data file (should be provided by the sequencing company) and where mothur will make/use files.

   ```
   set.dir(input=C:\Users\User\Desktop\FOLDERNAME)
   set.dir(output=C:\Users\User\Desktop\FOLDERNAME)
   ```

2. Extract the fasta, qual, and flow data from the sff data.

   ```
   sffinfo(sff=FILENAME.sff,flow=T)
   ```

3. Let's see what your sequences look like.

   ```
   summary.seqs(fasta=FILENAME.fasta,processors=2)
   ```

4. Partition your data by sample based on the sample-specific MIDs, trim the flows to a specified length range, remove sequences that have too many mismatches within primers/MIDs, *etc.* (*cf.* the different parameters).
   NOTE: The oligos file provides the MIDs and primers you used to mothur. You have to make this file yourself.

   ```
   trim.flows(flow=FILENAME.flow,oligos=FILENAME.oligos,pdiffs=2,bdiffs=1,minflows=360,maxflows=720,maxhomop=8,processors=2)
   ```

5. Using an expectation-maximization algorithm, the flowgrams are corrected to identify the idealized form of each flowgram and hereby translated to DNA (fasta) sequences.
   NOTE: This could take many hours/days, even with good hardware. Also, don't forget you need to download the "LookUp_Titanium.pat" file.

   ```
   shhh.flows(file=FILENAME.flow.files,processors=2)
   ```

6. The next thing you likely want to do is to remove MID/primer sequences from each idealized fasta sequence and generate a group file like you would if you had just used the raw fasta data.

   ```
   trim.seqs(fasta=FILENAME.shhh.fasta,name=FILENAME.shhh.names,oligos=FILENAME.oligos,pdiffs=2,bdiffs=1,minlength=400,maxlength=450,maxhomop=8,flip=T,processors=2)
   ```

7. Let's see what your new sequences look like.

   ```
   summary.seqs(fasta=FILENAME.shhh.trim.fasta,name=FILENAME.shhh.trim.names,processors=2)
   ```

8. The next step is to simplify the dataset by working with only the unique sequences.

   ```
   unique.seqs(fasta=FILENAME.shhh.trim.fasta,name=FILENAME.shhh.trim.names)
   ```

9. Take a look at your fasta data.

```
summary.seqs(fasta=FILENAME.shhh.trim.unique.fasta,name=FILENAME.shhh.trim.unique.names,processors=2)
```

10. Next, you need to align your data to the SILVA reference database.
NOTE: I recommend using the SILVA seed alignment. Download the latest version of the reference file and rename it "silva.align" (for some reason mothur crashes if you don't rename it).

```
align.seqs(fasta=FILENAME.shhh.trim.unique.fasta,reference=silva.align,flip=T,processors=2)
```

11. Take a look at the newly aligned sequences.

```
summary.seqs(fasta=FILENAME.shhh.trim.unique.align,name=FILENAME.shhh.trim.unique.names,processors=2)
```

12. To make sure that all of the sequences overlap in the same alignment space you need to remove those sequences that are outside the desired range.

```
screen.seqs(fasta=FILENAME.shhh.trim.unique.align,name=FILENAME.shhh.trim.unique.names,group=FILENAME.shhh.groups,start=XXX,end=YYY,processors=2)
```

13. Check the result.

```
summary.seqs(fasta=FILENAME.shhh.trim.unique.good.align,name=FILENAME.shhh.trim.unique.good.names,processors=2)
```

14. Next, you need to filter your alignment so that all of your sequences only overlap in the same region and to remove any columns in the alignment that don't contain data.

```
filter.seqs(fasta=FILENAME.shhh.trim.unique.good.align,trump=.,processors=2)
```

15. Let's further simplify the dataset again by working with only the unique sequences.

```
unique.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.fasta,name=FILENAME.shhh.trim.unique.good.names)
```

16. Let's see.

```
summary.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.fasta,name=FILENAME.shhh.trim.unique.good.filter.names,processors=2)
```

17. The final step you can take to reduce the sequencing error is to merge sequence counts that are within 1 bp (or more, you decide) of a more abundant sequence.
NOTE: This could take a while (hours/days), especially if your sequences per sample ratio is high.

```
pre.cluster(fasta=FILENAME.shhh.trim.unique.good.filter.unique.fasta,name=FILENAME.shhh.trim.unique.good.filter.names,group=FILENAME.shhh.good.groups,diffs=1,processors=2)
```

18. What is left now?

```
summary.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.names,processors=2)
```

19. Next to removing sequencing errors, you need to identify and remove chimeras.

    chimera.uchime(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.names,group=FILENAME.shhh.good.groups,processors=2)
    remove.seqs(accnos=FILENAME.shhh.trim.unique.good.filter.unique.precluster.denovo.uchime.accnos,fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.names,group=FILENAME.shhh.good.groups)

20. Check.

    summary.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.names,processors=2)

21. Now let's remove those sequences that are classified as “Chloroplast”, “Mitochondria”, or “Unknown” as well as archaeal and eukaryotic 16S/18S rRNA gene sequences.
    NOTE: You need to download the latest versions of the 16S rRNA gene reference (PDS) files.

    classify.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.names,group=FILENAME.shhh.good.pick.groups,template=trainset14_032015.pds.fasta,taxonomy=trainset14_032015.pds.tax,cutoff=80,processors=2)
    remove.lineage(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.names,group=FILENAME.shhh.good.pick.groups,taxonomy=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pds.wang.taxonomy,taxon=Mitochondria-Chloroplast-Archaea-Eukaryota-Unknown)

22. Check!

    summary.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.fasta,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.names,processors=2)

23. There are different approaches to analyze 16S rRNA gene sequences. Let’s analyze sequences binned into OTUs. First, generate a distance matrix.

    dist.seqs(fasta=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.fasta,cutoff=0.15,processors=2)

24. Next, you want to cluster these sequences into OTUs based on the newly created distance matrix and the latest names file.

    cluster(column=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.dist,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.names)

25. Now you can create a table that indicates the number of times an OTU shows up in each sample (*i.e.* an OTU table).

    make.shared(list=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.an.list,group=FILENAME.shhh.good.pick.pick.groups,label=0.03)

26. In this final step, you can get the taxonomic information for each of your OTUs.

    classify.otu(list=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.an.list,name=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pick.names,taxonomy=FILENAME.shhh.trim.unique.good.filter.unique.precluster.pick.pds.wang.pick.taxonomy,group=FILENAME.shhh.good.pick.pick.groups,label=0.03,cutoff=80)

APPENDIX III – PHENOTYPICALLY ANALYZED BACTERIA

TABLE A1 Complete list of the phenotypically analyzed phyllospheric bacteria. N/A: not assigned.

ID	Condition	Culture medium	Genus	Previously cultured	Cu res.	Zn res.	Cd res.	IAA prod.	Acetoin prod.	ACC deam. activity	Diesel fuel degr.
PhB1	Polluted	LB	N/A	N/A	–	–	–	+	–	+	–
PhB2	Polluted	LB	*Bacillus*	Yes	–	–	–	+	+	–	–
PhB3	Polluted	LB	N/A	N/A	–	–	–	–	+	–	–
PhB4	Polluted	LB	*Bacillus*	Yes	+	+	–	+	+	–	+
PhB5	Polluted	LB	*Novosphingobium*	Yes	–	–	–	+	–	+	–
PhB6	Polluted	LB	*Rhodococcus*	Yes	–	–	–	+	–	–	+
PhB7	Polluted	LB	N/A	N/A	–	–	–	+	–	–	–
PhB8	Polluted	LB	*Pantoea*	No	–	–	–	+	+	–	–
PhB9	Polluted	LB	*Microbacterium*	Yes	–	–	–	+	–	–	–
PhB10	Polluted	LB	*Pantoea*	No	+	–	–	+	+	–	–
PhB11	Polluted	LB	*Pantoea*	No	+	–	–	+	+	–	–
PhB12	Polluted	LB	*Bacillus*	No	–	–	–	+	+	–	+
PhB13	Polluted	LB	*Curtobacterium*	Yes	–	–	–	–	+	–	–
PhB14	Polluted	LB	*Bacillus*	No	–	–	–	–	+	–	–
PhB15	Polluted	LB	*Microbacterium*	No	–	–	–	–	+	–	–
PhB16	Polluted	LB	*Brevibacterium*	Yes	–	–	–	+	+	–	–
PhB17	Polluted	LB	N/A	N/A	–	–	–	+	–	+	–
PhB18	Polluted	LB	*Xanthomonas*	Yes	–	–	–	+	+	+	–
PhB19	Polluted	LB	*Pantoea*	No	+	–	–	+	+	–	–
PhB20	Polluted	LB	*Pantoea*	No	+	–	–	+	+	–	–
PhB21	Polluted	LB	*Novosphingobium*	Yes	–	–	–	+	–	–	–
PhB22	Polluted	LB	*Leifsonia*	Yes	–	–	–	+	+	–	–
PhB23	Polluted	LB	*Leifsonia*	Yes	–	–	–	+	+	–	–
PhB24	Polluted	LB	*Frigoribacterium*	Yes	–	–	–	+	+	+	–
PhB25	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	+	–
PhB26	Control	LB	*Curtobacterium*	Yes	–	–	–	–	+	–	–
PhB27	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB28	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB29	Control	LB	*Microbacterium*	Yes	–	–	–	+	–	–	–
PhB30	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB31	Control	LB	N/A	N/A	–	–	–	–	–	–	–
PhB32	Control	LB	N/A	N/A	–	–	–	+	–	–	–
PhB33	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB34	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB35	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB36	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB37	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB38	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB39	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	+	–
PhB40	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB41	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB42	Control	LB	*Frigoribacterium*	Yes	–	–	–	+	–	–	–
PhB43	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB44	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB45	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB46	Control	LB	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB47	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB48	Control	LB	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB49	Polluted	1/10 869	*Pedobacter*	No	–	–	–	+	–	+	–
PhB50	Polluted	1/10 869	*Microbacterium*	Yes	–	–	–	+	+	–	–
PhB51	Polluted	1/10 869	*Bacillus*	Yes	–	–	–	–	+	–	+
PhB52	Polluted	1/10 869	*Sphingomonas*	No	–	–	–	–	+	–	–
PhB53	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB54	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB55	Polluted	1/10 869	*Novosphingobium*	Yes	–	–	–	+	–	–	–
PhB56	Polluted	1/10 869	N/A	N/A	–	–	–	+	+	–	–
PhB57	Polluted	1/10 869	*Novosphingobium*	Yes	–	–	–	+	–	–	–
PhB58	Polluted	1/10 869	*Cellulomonas*	Yes	–	–	–	+	+	–	–

PhB59	Polluted	1/10 869	*Variovorax*	No	–	–	–	–	–	+	–
PhB60	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB61	Polluted	1/10 869	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB62	Polluted	1/10 869	*Pseudomonas*	No	–	–	–	+	–	+	–
PhB63	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB64	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	–	–
PhB65	Polluted	1/10 869	*Bacillus*	Yes	–	–	–	–	+	–	–
PhB66	Polluted	1/10 869	N/A	N/A	–	–	–	–	–	+	–
PhB67	Polluted	1/10 869	*Pedobacter*	Yes	–	–	–	+	–	+	–
PhB68	Polluted	1/10 869	N/A	N/A	–	–	–	+	+	–	–
PhB69	Polluted	1/10 869	*Methylobacterium*	No	–	–	–	+	–	+	–
PhB70	Polluted	1/10 869	*Cellulomonas*	No	–	–	–	–	–	+	–
PhB71	Polluted	1/10 869	*Methylobacterium*	No	–	–	–	+	–	–	–
PhB72	Polluted	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB73	Control	1/10 869	*Methylobacterium*	No	–	–	–	–	–	+	–
PhB74	Control	1/10 869	*Methylobacterium*	Yes	–	–	–	–	+	–	–
PhB75	Control	1/10 869	*Methylobacterium*	No	–	–	–	+	–	+	–
PhB76	Control	1/10 869	*Sphingomonas*	No	–	–	–	–	+	–	–
PhB77	Control	1/10 869	*Sphingomonas*	No	–	–	–	+	–	–	–
PhB78	Control	1/10 869	*Curtobacterium*	No	–	–	–	+	–	–	–
PhB79	Control	1/10 869	N/A	N/A	–	–	–	–	–	–	–
PhB80	Control	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB81	Control	1/10 869	*Sphingomonas*	Yes	–	–	–	+	–	–	–
PhB82	Control	1/10 869	N/A	N/A	–	–	–	–	+	–	–
PhB83	Control	1/10 869	N/A	N/A	–	–	–	–	–	–	–
PhB84	Control	1/10 869	*Curtobacterium*	Yes	–	–	–	–	+	–	–
PhB85	Control	1/10 869	*Methylobacterium*	No	–	–	–	+	–	+	–
PhB86	Control	1/10 869	*Caryophanon*	No	–	–	–	+	–	–	–
PhB87	Control	1/10 869	N/A	N/A	–	–	–	+	–	–	–
PhB88	Control	1/10 869	*Sphingomonas*	No	–	–	–	+	–	–	–
PhB89	Control	1/10 869	*Rathayibacter*	No	–	–	–	+	–	–	–
PhB90	Control	1/10 869	*Sphingomonas*	Yes	–	–	–	–	–	–	–
PhB91	Control	1/10 869	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB92	Control	1/10 869	N/A	N/A	–	–	–	+	–	–	–
PhB93	Control	1/10 869	*Rathayibacter*	No	–	–	–	+	–	+	–
PhB94	Control	1/10 869	*Mucilaginibacter*	No	–	–	–	–	–	–	–
PhB95	Control	1/10 869	N/A	N/A	–	–	–	+	–	+	–
PhB96	Control	1/10 869	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB97	Polluted	YECD	*Erwinia*	Yes	–	–	–	+	+	+	–
PhB98	Polluted	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB99	Polluted	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB100	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB101	Polluted	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB102	Polluted	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB103	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB104	Polluted	YECD	N/A	N/A	–	–	–	–	–	+	–
PhB105	Polluted	YECD	*Frigoribacterium*	No	–	–	–	+	+	+	–
PhB106	Polluted	YECD	N/A	N/A	–	–	–	–	–	–	–
PhB107	Polluted	YECD	*Frigoribacterium*	Yes	–	–	–	–	–	–	–
PhB108	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB109	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB110	Polluted	YECD	*Nocardioides*	Yes	–	–	–	+	–	+	–
PhB111	Polluted	YECD	N/A	N/A	–	–	–	–	–	–	–
PhB112	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB113	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB114	Polluted	YECD	*Frigoribacterium*	Yes	–	–	–	–	–	–	–
PhB115	Polluted	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB116	Polluted	YECD	*Frigoribacterium*	No	–	–	–	+	–	+	–
PhB117	Polluted	YECD	*Frigoribacterium*	Yes	–	–	–	+	–	+	–
PhB118	Polluted	YECD	*Frigoribacterium*	Yes	–	–	–	+	–	+	–
PhB119	Polluted	YECD	N/A	N/A	–	–	–	–	–	–	–
PhB120	Polluted	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB121	Control	YECD	*Paenibacillus*	Yes	–	–	–	+	–	–	–
PhB122	Control	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB123	Control	YECD	*Paenibacillus*	No	–	–	–	+	–	–	–
PhB124	Control	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB125	Control	YECD	N/A	N/A	–	–	–	+	–	–	–

PhB126	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	+	–
PhB127	Control	YECD	*Rathayibacter*	No	–	–	–	–	–	–	–
PhB128	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB129	Control	YECD	*Methylobacterium*	No	–	–	–	+	–	–	–
PhB130	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB131	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB132	Control	YECD	N/A	N/A	–	–	–	+	–	+	–
PhB133	Control	YECD	*Roseomonas*	Yes	–	–	–	–	–	–	–
PhB134	Control	YECD	*Cellulomonas*	No	–	–	–	+	–	–	–
PhB135	Control	YECD	*Sphingomonas*	Yes	–	–	–	+	–	–	–
PhB136	Control	YECD	*Curtobacterium*	Yes	–	–	–	–	–	+	–
PhB137	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	+	+	–
PhB138	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB139	Control	YECD	*Nocardioides*	Yes	–	–	–	+	–	–	–
PhB140	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB141	Control	YECD	*Curtobacterium*	No	–	–	–	+	–	–	–
PhB142	Control	YECD	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB143	Control	YECD	*Cellulomonas*	Yes	–	–	–	+	–	–	–
PhB144	Control	YECD	N/A	N/A	–	–	–	+	–	–	–
PhB145	Polluted	YMAb 284	*Methylobacterium*	Yes	–	–	–	+	–	+	–
PhB146	Polluted	YMAb 284	*Curtobacterium*	Yes	–	–	–	–	–	–	–
PhB147	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB148	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB149	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB150	Polluted	YMAb 284	*Cellulomonas*	No	–	–	–	+	–	–	–
PhB151	Polluted	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	+	–
PhB152	Polluted	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	+	–
PhB153	Polluted	YMAb 284	*Frondihabitans*	Yes	–	–	–	+	–	+	–
PhB154	Polluted	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB155	Polluted	YMAb 284	N/A	N/A	–	–	–	–	+	–	–
PhB156	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB157	Polluted	YMAb 284	*Lysinibacillus*	No	–	–	–	+	–	+	–
PhB158	Polluted	YMAb 284	*Frigoribacterium*	Yes	–	–	–	+	–	+	–
PhB159	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB160	Polluted	YMAb 284	*Frigoribacterium*	No	–	–	–	+	–	+	–
PhB161	Polluted	YMAb 284	*Frondihabitans*	Yes	–	–	–	+	+	+	–
PhB162	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB163	Polluted	YMAb 284	N/A	N/A	–	–	–	–	+	–	–
PhB164	Polluted	YMAb 284	*Bacillus*	Yes	–	–	–	+	+	–	–
PhB165	Polluted	YMAb 284	*Frondihabitans*	Yes	–	–	–	+	–	–	–
PhB166	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB167	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB168	Polluted	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB169	Control	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB170	Control	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	+	–
PhB171	Control	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB172	Control	YMAb 284	*Frigoribacterium*	Yes	–	–	–	+	–	–	–
PhB173	Control	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB174	Control	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	–	–
PhB175	Control	YMAb 284	*Methylobacterium*	No	–	–	–	–	–	–	–
PhB176	Control	YMAb 284	N/A	N/A	–	–	–	+	+	+	–
PhB177	Control	YMAb 284	*Methylobacterium*	Yes	–	–	–	+	–	–	–
PhB178	Control	YMAb 284	*Methylobacterium*	No	–	–	–	+	–	+	–
PhB179	Control	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	+	–
PhB180	Control	YMAb 284	*Methylobacterium*	Yes	–	–	–	+	–	+	–
PhB181	Control	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB182	Control	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB183	Control	YMAb 284	N/A	N/A	–	–	–	+	–	–	–
PhB184	Control	YMAb 284	*Mucilaginibacter*	Yes	–	–	–	–	–	–	–
PhB185	Control	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	+	–
PhB186	Control	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	–	–
PhB187	Control	YMAb 284	*Patulibacter*	Yes	–	–	–	+	–	–	–
PhB188	Control	YMAb 284	*Frondihabitans*	Yes	–	–	–	+	–	+	–
PhB189	Control	YMAb 284	N/A	N/A	–	–	–	+	–	+	–
PhB190	Control	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	+	–
PhB191	Control	YMAb 284	*Curtobacterium*	Yes	–	–	–	+	–	+	–
PhB192	Control	YMAb 284	*Rathayibacter*	No	–	–	–	+	–	+	–