Green infrastructure and atmospheric pollution shape diversity and composition of phyllosphere bacterial communities in an urban landscape

Green infrastructure and atmospheric pollution shape diversity and composition of phyllosphere bacterial communities in an urban landscape

Karen WUYTS[1,*], Wenke SMETS[a], Sarah LEBEER[2], Roeland SAMSON[b]

Environmental Ecology and Applied Microbiology (ENdEMIC), Department of Bioscience Engineering, University of Antwerp, Groenenborgerlaan 171, 2020 Antwerp, Belgium

[a] KW and WS should be considered joint first author

[b] SL and RS should be considered joint senior author

[*]To whom correspondence should be addressed. Tel: +32 3265 3452; email: karen.wuyts@uantwerpen.be; ORCID iD: 0000-0002-3371-6828

Published in FEMS Microbiology Ecology 96 (2020) doi: 10.1093/femsec/fiz173

[1] KW and WS should be considered joint first author

[2] SL and RS should be considered joint senior author

Abstract

The microbial habitat on leaf surfaces, also called the phyllosphere, is a selective environment for bacteria, harbouring specific phyllosphere bacterial communities (PBCs). These communities influence plant health, plant-community diversity, ecosystem functioning and ecosystem services. Host plants in an urban environment accommodate different PBCs than those in non-urban environments, but previous studies did not address individual urban factors. In this study, the PBC composition and diversity of 55 London plane (*Platanus* x *acerifolia*) trees throughout an urban landscape (Antwerp, Belgium) was determined using 16S rRNA amplicon sequencing. An increasing proportion of green infrastructure in the surrounding of the trees, and subsequently decreasing proportion of anthropogenic land use, was linked with taxa loss, expressed in lower phyllosphere alpha diversity and higher abundances of typical phyllosphere bacteria such as *Hymenobacter*, *Pseudomonas* and *Beijerinckia*. Although air pollution exposure, as assessed by leaf magnetic analysis, did not link with alpha diversity, it correlated with shifts in PBC composition in form of turnover, an equilibrium of taxa gain and taxa loss. We found that both urban landscape composition and air pollution exposure - each in their own unique way – influence bacterial communities in the urban tree phyllosphere.

Introduction

Where plant surfaces are exposed to the atmosphere, a unique microbial habitat exists, which is termed the phyllosphere (Last, 1955; Lindow and Leveau, 2002). The phyllosphere harbours specific microbial communities of which the diversity and composition vary across plant species and with spatial and temporal factors. These variations are of great interest as phyllosphere bacterial communities (PBC) directly and indirectly influence plant diversity, ecosystem functioning and ecosystem services (Bradley et al., 2008; Laforest-Lapointe et al., 2017b).

We know that urbanization drives global and local environmental change, decreases local biodiversity and accelerates phenotypic changes in plants and animals (Alberti et al., 2017; Grimm et al., 2008; McKinney, 2008). Mechanisms involved include loss and fragmentation of native habitat, habitat modification (e.g. urban heat-island effect), increase in habitat heterogeneity, change in biotic interactions (e.g. introduction of exotic species) and novel disturbances (e.g. air and soil pollutants). It is therefore not surprising that important significant differences have been found between PBC of urban sites compared to non-urban sites (Brighigna et al., 2000; Khanna, 1986; Smets et al., 2016b). These studies propose air pollution as the driving environmental factor, however, land use could also affect conditions in the phyllosphere and limit dispersal of phyllosphere bacteria through the air. Indeed, the proportion of vegetated area in an urban area affects the airborne bacterial communities (Mhuireach et al., 2016). Nevertheless, up to now, it is unclear to what extent different urban disturbance factors among which air pollution, habitat change and altered plant host characteristics are responsible for shifts in urban PBC.

Following a landscape ecological approach (Turrini & Knop, 2015), we investigated to what extent the inter-tree variation in diversity and composition of bacterial communities in the phyllosphere of trees (*Platanus* x *acerifolia*) in an urban environment is related to exposure to air pollution, to habitat change (land use/cover composition), habitat fragmentation and heterogeneity (land use/cover diversity and isolation metrics), and to plant traits. Our hypothesis is that intra-urban

shifts in PBC composition are induced by a combination of (a) a lower richness in phyllosphere bacteria in areas with less (and further from) green infrastructure due to reduced dispersal from these sources of bacteria, (b) an increase in bacterial diversity in areas with high land use/cover diversity and (c) and the favouring of some taxa and the reduction of others due to the influence of air pollution either as a resource or as a stressor.

Materials and methods

Experimental layout & sampling

In summer 2015, London plane tree (*Platanus* × *acerifolia* Willd.) leaves were sampled at 55 locations (< 11 km apart) in the city of Antwerp, in the northern part of Belgium (Fig. 1). The city houses 516 000 inhabitants (2015) on 205 km^2. The most densely built-up part of the city resides on the right bank of the river Scheldt. The old city centre is a residential and commercial area, with small parks, and is surrounded by a heavily trafficked ring motorway of six to ten lanes. Outside the ring road, neighbourhoods are quieter and encompass larger green areas. The ring road, the motorways connected to it and the local roads entering the city daily suffer serious traffic congestions. The Port of Antwerp, Europe's second largest container port, is located north of the city. It harbours international shipping and industrial activities (e.g. petrochemical industries) by more than 900 companies. In the south, near the precious-metals refinery plant of Umicore, elevated Pb, As and Cd concentrations are measured in the atmosphere (Vlaamse Milieumaatschappij, 2016). The city experiences a maritime temperate climate, with lowest and highest mean temperatures of 3 °C in January and 19 °C in July, respectively, and a mean annual rainfall of 848 mm (1981-2010; Royal Meteorological Institute of Belgium (RMI), http://www.kmi.be). In 2015, spring and summer were relatively dry (140 and 160 mm of rain, respectively; RMI, http://www.kmi.be). Mean atmospheric concentrations of the most important

pollutants at the nearest air quality monitoring stations (operated by Flanders Environment Agency; Fig. 1) in 2015 are provided in Table S1.

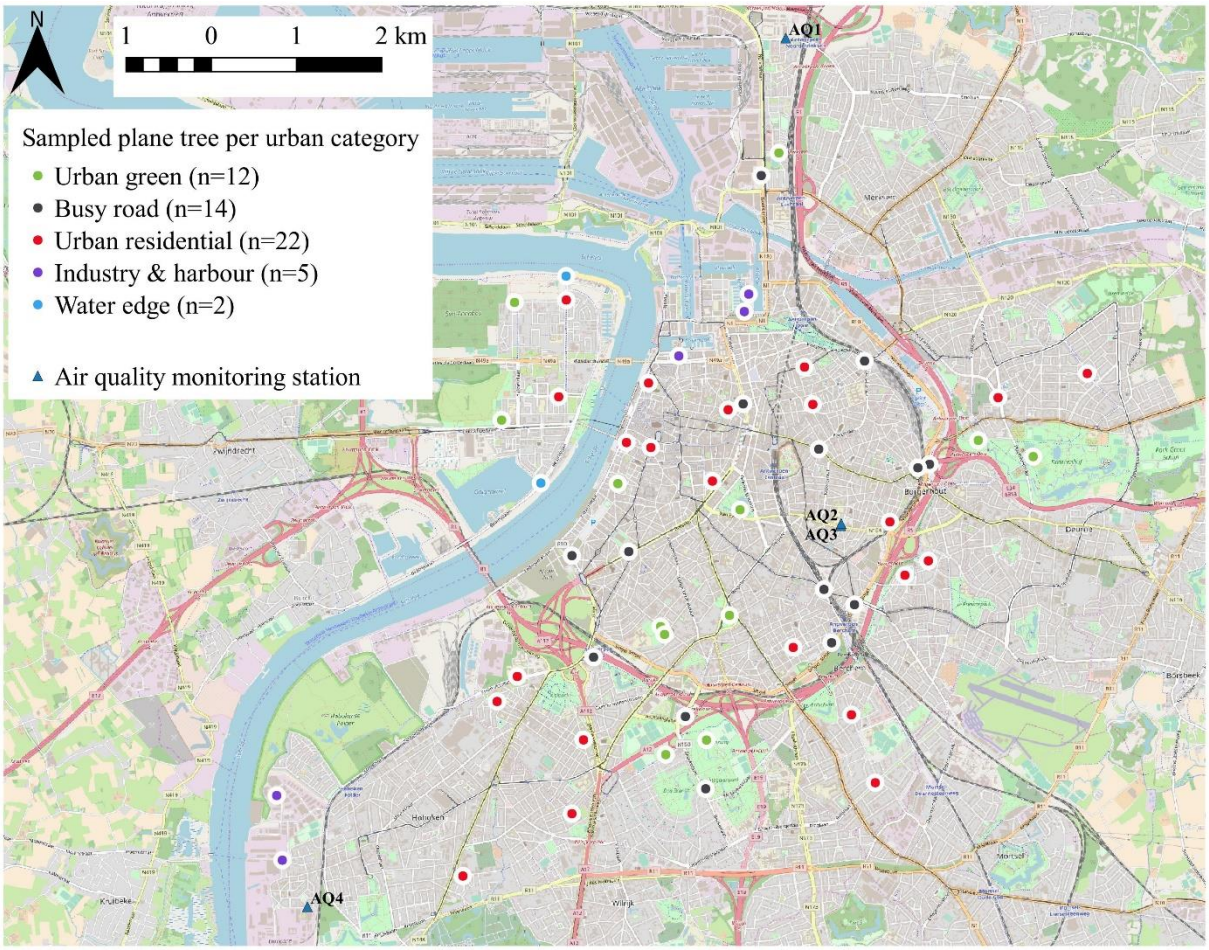

Fig. 1. London plane trees were sampled at 55 locations throughout the city of Antwerp within five urban categories (source of background map: OpenStreetMap contributors, https://www.openstreetmap.org/)

The London plane tree was selected as host species of PBC because it is a common tree throughout the city of Antwerp, and in Europe in general, it is easy to recognize and its canopy is usually out of reach for people to touch, which decreases chances of contamination. Fifty-five trees were selected based on leaf magnetic data (similar analyses as described in §Biomagnetic analyses) from 150 London plane trees in Antwerp (geolocated and characterised for hyperspectral biomonitoring of air quality, HYPERCITY project; Samson et al. 2016), so that a full and balanced range of air

pollution exposure representative for the city was attained.From each tree, two branches were cut using a telescopic pruner, between 2 and 5 m height above street level. Branches were intercepted before touching the ground. From each branch, immediately, two undamaged leaves were cut from the branch with scissors using gloves sterilized on site with 70% ethanol. The leaves of each plant were put in sterile 50 mL falcons (VWR) and transported on ice to the lab. From the same branches, four healthy, undamaged leaves per branch were sampled for magnetic analysis and stored in paper envelopes. For measurements of drop contact angle (DCA; a measure for leaf wettability), one leaf per branch was collected and transported to the lab in a cooled box. Leaf relative chlorophyll content (RCC) was measured on site on three attached leaves per branch (two measurements per leaf) using a SPAD-502Plus chlorophyll meter (Konica Minolta) and an average RCC was calculated for every tree. In the field, the sites of the sampled trees were allotted into five urban categories (Fig. 1): urban green (trees in public parks, green patches and semi-natural areas with large trees; n=12), busy roads (street trees < 10 m from one or more roads with high-intensity motorized traffic; n=14), urban residential areas (trees between buildings, from three-story high up to housing blocks, > 50 m from nearest busy road; n=22), industry & harbour (n=5) and water edge (trees < 50 m from river Scheldt, > 50 m from nearest busy road; n=2).

All samples and in situ measurements were performed on seven days between the 12th and 21st of August 2015. A period of eight days prior to the first sampling day was free of rain. During the sampling campaign, there were rain events in the evening of the 13th and 14th of August, of 17 and 10 mm, respectively (Flanders Environment Agency, http://www.waterinfo.be). Mean daytime temperature during the campaign was 21°C.

Sequencing & data processing

Phyllosphere bacteria were extracted from the leaves upon arrival in the lab, 1 to 6 h after samples were taken. Hereto, 5 mL sterile leaf wash solution (Redford and Fierer, 2009) was added to each

falcon tube. To suspend the phyllosphere bacteria, the tubes were vortexed for 5 min at maximum speed with the Vortex Genie® 2 (MoBio). The falcons were centrifuged at 1000 g for several seconds to spin down most of the buffer still sticking to the leaves. Leaves were removed from the falcons and the remaining wash solution was centrifuged in aliquots of 2 mL at 12300g for 2 min and the supernatant was discarded. The final pellet was resuspended in 750 µL of Bead Solution (PowerFecal DNA Isolation Kit, MoBio). This suspension was stored at -20°C until further processing within 30 days.

Further processing was focused on allowing data comparison with other relevant studies and was therefore based on the methodology of the Earth Microbiome Project (www.earthmicrobiome.org). . DNA was extracted using the PowerFecal DNA Isolation Kit, according to the manufacturer's instructions. This kit uses the same buffers and procedure as the PowerSoil DNA Isolation Kit. The extracted DNA was used as template for PCR amplification of the V4 region of the 16S rRNA gene with the Phusion High-Fidelity DNA polymerase (ThermoFisher Scientific) and barcoded primers (IDT) as described by Kozich et al. (2013) but with only 25 cycles. A PCR blank was included. Amplicons were purified and normalized using the SequalPrep™ Normalization Plate Kit (ThermoFisher Scientific) according to the manufacturer's instructions. All samples were pooled into one library and diluted to 2nM, based on the DNA concentration determined by a Qubit 3.0 Fluorometer (Thermo Fisher Scientific). The library was sequenced at the Centre for Medical Gentics (Edegem, Belgium) with Illumina MiSeq, using the 500-cycles MiSeq Reagent Kit v2 (Illumina).

The UPARSE pipeline (Edgar, 2013) was used to assemble the paired reads, conduct quality filtering, and cluster sequences into operational taxonomic units (OTUs). Assembly of the paired end reads was set to a maximum of 34 differences in the overlap zone and the minimum length of the overlap was set to 150 base pairs. A maximum per sequence expected error frequency value of 0.5 was used to quality-filter sequences and global singletons were removed. The remaining sequences were clustered into OTUs at 97% similarity using a *de novo* clustering approach.

Taxonomy was assigned using the RDP classifier (Wang et al., 2007) trained on the Greengenes database (version 13_08). The mitochondrial and chloroplast reads were removed. The sequences obtained in this study are available in the NCBI database with BioProject number PRJNA359642.

Biomagnetic analysis

In the lab, each leaf sample was split in two subsamples each containing two leaves. The one-sided leaf surface area of the subsamples was determined using a leaf area meter (LI-3100C, LI-COR), after which the subsamples were ovendried (at 60 °C for three days). The saturation isothermal remanent magnetisation (SIRM) was determined of each leaf subsample, as a proxy for exposure to traffic- and industry-derived PM (Hansard et al., 2011; Kardel et al., 2012; Mitchell and Maher, 2009; Muxworthy et al., 2003). Following the protocol described by Kardel et al. (2011), the dry subsamples were tightly packed in small sample pot using cling film and magnetized in a pulse DC magnetic field of 1 Tesla using a pulse magnetizer (Molspin Ltd, UK). Immediately after, the remanent magnetisation was measured twice using a Minispin magnetometer (Molspin), with a sensitivity of 10^{-9} A m^2. Method blanks were included as empty containers. The remanent magnetisation of a subsample (mean of duplo measurement) was multiplied by the assumed sample volume of the Minispin (10^{-5} m^3) to obtain the magnetic moment (A m^2). The leaf area-normalized SIRM (in A) of a subsample was obtained by dividing its magnetic moment by its one-sided leaf area. The data were averaged to tree level.

Leaf wettability analysis

Leaf wettability was determined immediately upon arrival in the lab, within 1 to 6 h after leaf collection. From each leaf taken for leaf wettability assessment, two pieces of about 2 cm^2 were cut out of the leaf blade, i.e. left and right from the mid vein, halfway the petiole offset and the leaf

tip. Both were attached to a flat, horizontal surface with double-sided tape, so that one piece had the adaxial leaf side up and the other the abaxial. Three drops of 7.5 µl of deionised water at room temperature were carefully positioned on each leaf piece using a micropipette. Immediately after, the two visually most symmetric drops per piece were photographed using a Canon EOS 550D digital camera with Canon MP-E 65mm micro lens (at 3:1 magnification). The digital pictures were processed with the image-processing software ImageJ (version 1.51g). The drop contact angle (the inside angle between the outside curve of the drop with the leaf surface) was determined for each drop from the red band image using the Java plug-in for sessile and symmetric drops 'Low-Bond Axisymmetric Drop Shape Analysis' (Stalder et al., 2010). Adaxial and abaxial DCA data (DCA_{ad} and DCA_{ab}, respectively) were each averaged to tree level.

Vehicle density and modelled atmospheric pollutant concentrations

Vehicle density was determined as the daytime average number of motorized vehicles passing per hour within a radius of 50 m of a tree using modelled traffic data (SGS, 2010). Modelled yearly NO_2, PM_{10} and $PM_{2.5}$ concentrations (in µg m^{-3}) were extracted for each location from air quality maps modelled following a IFDM-OSP model chain (Vranckx and Lefebvre, 2013), available at the city of Antwerp's open data platform (http://opendata.antwerpen.be/). Spatial data analyses were performed using QGIS 2.16 Nødebo.

Landscape metrics

The Urban Atlas 2012, developed by the EU's Copernicus Land Monitoring Service based on satellite image interpretation at 2.5 m resolution for the reference year 2012 (freely accessible at http://land.copernicus.eu/), was used to calculate land-cover statistics around each of the sampling sites. The land cover vector file was converted to a raster file with a spatial resolution of 5 m and

reclassed into twelve land-cover classes (Table S2). In seven concentric buffers around each sampling site, with a radius of 20, 50, 100, 200, 300, 400 and 500 m, landscape and class metrics were calculated using the LecoS plug-in developed by Jung (2016) for QGIS. Land-cover richness (LCR; i.e. the total number of different land-cover classes), land-cover diversity (Shannon H and Simpson λ indices) and the proportion of each land-cover class in the buffer zones were extracted from the land-cover map. In addition, class metrics were calculated also for green infrastructure (GI), which includes all vegetated land-cover classes, i.e. urban green, semi-natural & agricultural area, forest and sport & leisure. Moreover, the number of patches (measure of class fragmentation), Euclidean nearest-neighbour distance (measure of patch isolation; mean shortest edge-to-edge distance between patches of the same class) and Patch Cohesion index (measure of class connectivity) were calculated for GI in the buffer zones. Distances to the nearest GI and water infrastructure were extracted for each sampling point. Pair-wise geographic distances between the sampling sites were calculated from their geographic coordinates using the open-source application Geographic Distance Matrix Generator (http://biodiversityinformatics.amnh.org/open_source/gdmg).

Statistical analyses

All statistical analyses were performed in R (R core team 2016). From the bacterial communities, we calculated the OTU richness (number of different OTUs) and the alpha diversity using the Shannon diversity index of each sampling site using the R package *vegan* (Oksanen et al. 2016). The pairwise beta diversity was calculated as (i) Bray-Curtis dissimilarity (β_{bc}) using *vegan* and (ii) Simpson dissimilarity β_{sim}, nestedness-resultant dissimilarity β_{nes} and Sørensen dissimilarity β_{sor} using the *betapart* package (Baselga et al. 2017). This was done on the incidence data rather than abundance data, as suggested by Baselga (2013) for non-visible populations that are difficult

to measure without bias. Likewise, the multi-site beta diversity indices β_{SIM}, β_{NES} and β_{SOR} were calculated following R scripts provided by Ensing & Pither (2015).

The relationships of the phyllosphere bacterial alpha diversity (OTU richness and Shannon diversity) with the environmental site variables (model vehicle density, model NO_2, $PM_{2.5}$ and PM_{10} concentrations), leaf variables (leaf SIRM, RCC, DCA_{ad} and DCA_{ab}), the distances of the sampling sites to the nearest GI and water surface and the land cover metrics in the seven buffer zones (land cover richness, Shannon diversity index and land cover proportions) were investigated with Kendall's τ correlation. Differences in alpha diversity between the urban categories in which the sites were located (busy roads, urban residential, urban green, industrial & harbour areas and water edge) were tested with a Kruskal-Wallis rank sum test. Relationships between different land cover proportions, environmental site and leaf variables were investigated with PCA and Kendall's correlation analyses.

To explore the relationship between pairwise dissimilarities in community structure and pairwise differences in space and environment (Anderson et al. 2010), the relationships between pairwise β diversity (Bray-Curtis β_{bc}, Simpson β_{sim}, nestedness-resultant β_{nes} and Sørensen β_{sor} dissimilarities) on the one hand and spatial distance and differences in environmental site variables, leaf variables and land cover on the other hand were tested with Mantel tests from the package *vegan*. The beta diversity indices β_{bc} and β_{sor} were tested for differences between the urban categories with a Permanova and for dispersion effects with permdisp using *vegan* (adonis and betadispr).

Following the workflow by Gloor et al. (2017), the centered log-ratio of the unrarified reads (further on called clr-abundances) were calculated with the count zero multiplicative approach, using the R package *zCompositions* (Palarea-Albaladejo & Martin-Fernandez 2015). A principal component analysis (PCA) was performed on the clr-abundances of the full dataset using *vegan* (Oksanen et al. 2016). Linear models were built to explain the variation in community composition using the site scores on the first two PCA axes as dependent variables and the environmental, leaf and

landscape variables as explanatory variables. Differences in PCA scores between urban categories observed on site was tested with a Kruskal-Wallis test.

The clr-abundances of the 50 most dominant taxa (genera and families) were calculated as a subset. The core community was identified as the group of OTUs that occurred in the phyllosphere of all 55 trees, and the clr-abundances were calculated for the sum of reads of the core OTUs and the sum of the non-core OTUs. The clr-abundances of the most dominant taxa and of the full core community were tested for significant relationships with the variables that related significantly with the site scores along one of the two PCA axes, using Pearson correlation. Because of multiple testing, p values were adjusted for false discovery rate (p_{adj}) following the correction by Benjamini & Hochberg (2000) using the FDR-based software routines supplied by Pike (2011). With the package *ALDEx2* (Fernandes et al. 2013, 2014; Gloor et al. 2016), the effect of urban category observed at the site of the trees on the clr-abundance of the 50 most abundant taxa and the full core community was tested and pair-wise comparisons were performed between busy road, urban residential, urban green, and industry & harbour.

Results

Alpha diversity lower near more green infrastructure

The OTU richness and Shannon diversity of the PBC of London plane (Table S3) were significantly affected by the urban category observed at the sampling sites (Kruskal-Wallis p=0.0155 and 0.0313, respectively). The OTU richness was twice as low in the phyllosphere sampled in urban green and at the water edge than in the busy road, urban residential areas and industry & harbour sites (Fig. 2a). The urban categories differed significantly in leaf SIRM (p=0.0012; Fig. 2c) and the proportion of green infrastructure (GI) in all buffer zones around the sampled trees (with 20 to 500 m radius, all $p<0.01$; Fig. 2d).

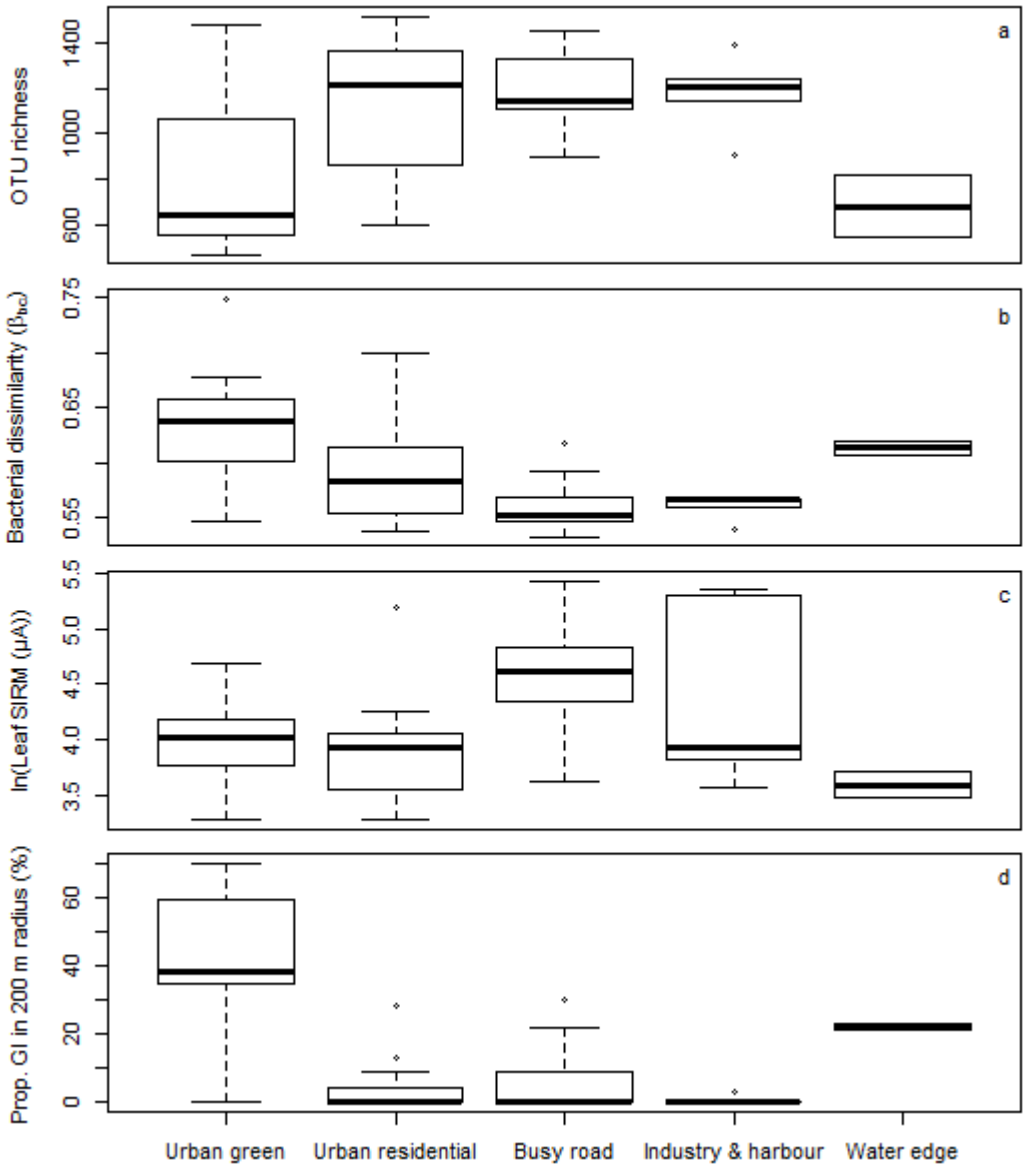

Fig. 2. OTU richness (a) and Bray-Curtis β diversity (b) of the phyllosphere bacterial communities, the air pollution proxy of ln-transformed leaf SIRM (c) and the proportion of green infrastructure in a buffer zone with 200 m radius around the sampling site (d) among the five urban categories (median values with hinges indicating 25th and 75th percentiles and whiskers 1.5 x the inter-quartile range)

Furthermore, when considering the individual urban factors (listed per site in Table S3), bacterial OTU richness and Shannon diversity significantly increased with the sites' distance to the nearest GI (Kendall's τ=0.20, p=0.033; τ=0.20, p=0.040, respectively), which did not significantly correlate with any of the other environmental site variables. The OTU richness related positively

with model vehicle density within a 50 m radius (τ=0.22, p=0.0187), which in turn correlated positively with our air pollution proxy (leaf SIRM; τ=0.27, p= 0.0042). The Shannon diversity related significantly with the abaxial drop contact angle (DCA), a measure of leaf wettability (τ=0.26, p=0.0059): diversity was higher when leaves were more hydrophilic. Relative chlorophyll content (RCC), air pollution (leaf SIRM, modelled $PM_{2.5}$, PM_{10} and NO_2 concentrations) nor proximity to the nearest open water significantly contributed to the explanation of bacterial alpha diversity.

With regard to the land-cover metrics, the OTU richness and diversity of the PBC did not significantly relate to the richness and Shannon diversity of Urban Atlas' land-cover in buffer zones with varying radii (20-500 m) around the London plane trees. The OTU richness and Shannon diversity of the PBC decreased significantly with the proportion of Urban Atlas' urban green within the buffer zones around the sampling sites up to a radius of 200 m. The best fits (with most negative Kendall correlation coefficient τ and lowest p value) were attained in the 20 m radius buffer zones (p=0.0210, τ=-0.26 for OTU richness and p=0.0099, τ=-0.28 for Shannon diversity). The proportion of all GI together (including urban green, agricultural and semi-natural lands, forests, sports & leisure) explained diversity even better than urban green alone, with the best fits in the 20 m radius buffer zones (p=0.0078, τ=-0.29 for OTU richness and p=0.0117, τ=-0.28 for Shannon diversity). The land cover proportions of urban green and GI in almost all buffer zones related significantly and negatively with the model vehicle density, but not with leaf SIRM in any of the buffer zones. The OTU richness related significantly and positively to the proportion of industry & harbour in buffer zones of 400 m and 500 m radius (τ=0.25, p=0.0074) and negatively to the proportion of open water in the 500 m radius buffers (τ=-0.23, p=0.0302). Proportions of the other land covers (urban fabric, roads, construction & dump sites, harbour and railway) nor the richness and Shannon diversity of land-cover in buffer zones with varying radii (20-500 m) significantly related to OTU richness & diversity. For all GI, the number of patches but not the Euclidean Nearest-Neighbour distance related significantly with the OTU richness, in buffer zones up to 100

m, and with Shannon diversity, in buffer zones of 50 m radius. The most relevant landscape variables are listed per site in Table S3.

Beta diversity affected by green infrastructure and air pollution

The pairwise Bray-Curtis (β_{bc}) and Sørensen (β_{sor}) beta diversities of the PBCs differed significantly between the urban categories (permanova, both p=0.001), but this seemed to be caused by both a dispersion effect (difference in multivariate spread; Permdisp p=0.0008 and <0.0001, respectively) and a location effect. Following the partitioning of beta diversity by Baselga et al. (2010), the multi-site incidence-based Sørensen beta diversity (β_{SOR}) of the PBC was 0.942, of which 0.928 could be attributed to dissimilarity due to taxa turnover (β_{SIM}) and 0.024 to nestedness-resultant dissimilarity (β_{nes}). None of the pairwise beta diversity indices related significantly with the spatial distance between the sites (Fig. 3). The pairwise incidence-based Sørensen dissimilarity β_{sor} and abundance-based Bray-Curtis dissimilarity β_{bc} both increased with increasing difference in the proportion of urban green within a radius up to 300 to 400 m, with the highest Mantel correlation coefficient in the 200 m radius. As part of the β_{sor}, the nestedness-resultant dissimilarity β_{nes} related also most significantly with the difference in the proportion of GI in the buffer zone with a 200 m radius. On the contrary, pairwise β_{sim}, the turnover component of β_{sor}, did not relate with the dissimilarity in proportion of urban green or GI, but related significantly with the difference in leaf SIRM and the difference in modelled atmospheric NO_2 concentrations (Fig. 3); relationships with the modelled $PM_{2.5}$ and PM_{10} concentrations were marginally insignificant. β_{nes} and β_{sim} did not relate with dissimilarity in any other site or leaf variable including vehicle density.

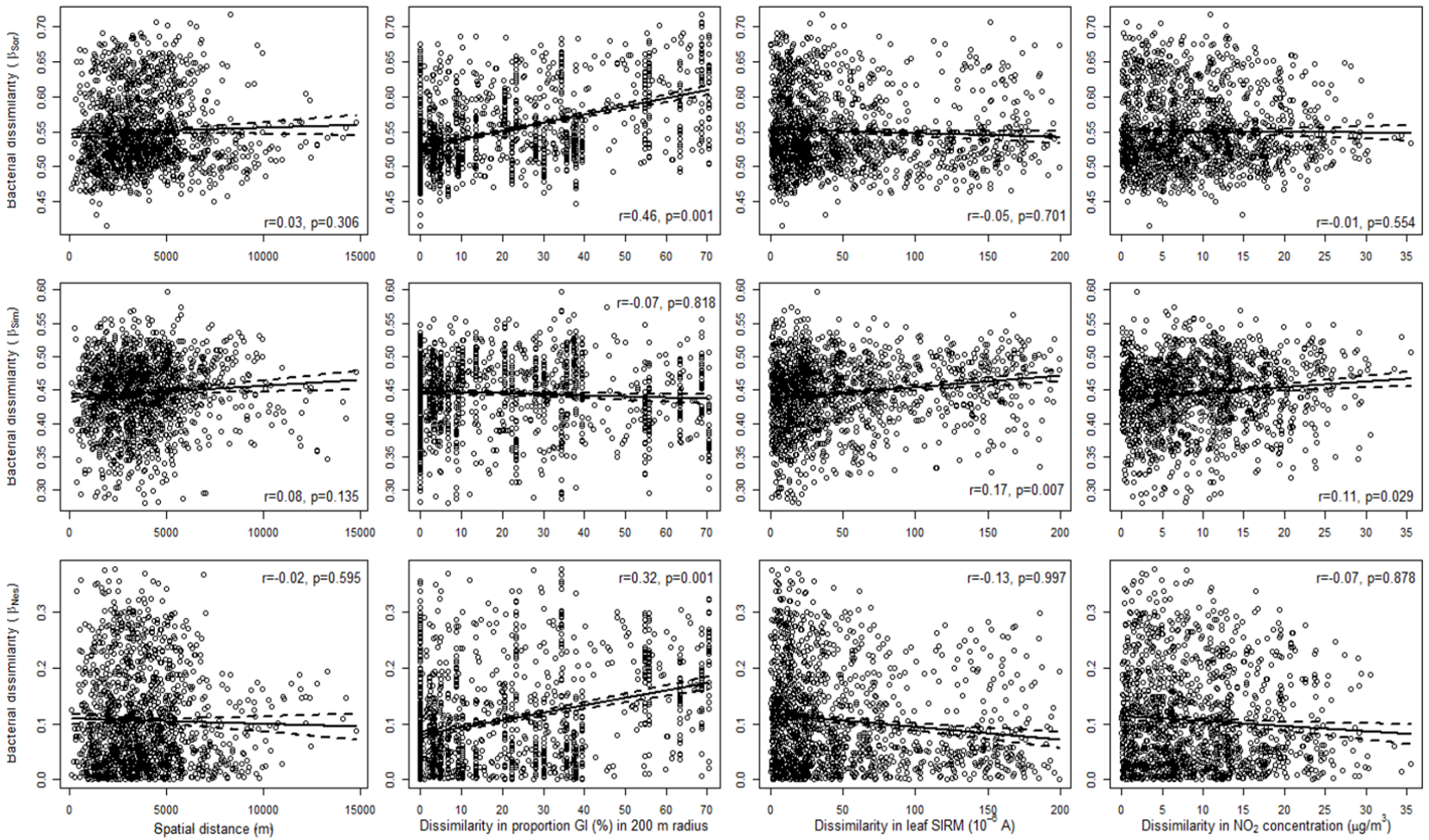

Fig. 3. Relationships of bacterial incidence-based pairwise Sørensen dissimilarity (β_{sor}; upper row), split up in dissimilarity caused by turnover (β_{Sim}; middle row) and by nestedness (β_{Nes}; lower row), with spatial distance between the sampling sites and pairwise dissimilarity in leaf SIRM, modelled NO_2 concentration and proportion of green infrastructure (GI) within a 200 m-radius buffer. Each point represents the dissimilarity between two sampling sites; lines indicate the fitted linear regression and its 95% confidence interval. Pearson correlation (r) and significance (p computed using Mantel tests) are shown for each relationship.

Following principal component analysis (PCA) on the clr-abundance (Fig. 4), the first and second principal component, explaining 16.1 and 6.4% of the whole-community variation, respectively, differed significantly between the urban categories observed at the sites (p=0.0418 and 0.0055, respectively). From all continuous leaf and environmental traits, the RCC, the distance to the nearest GI, the proportion of GI (in a radius up to 200 m; best fit for 100 m) and the proportion of open water (in a 300 m radius and more; best fit for 500 m) significantly contributed to the explanation of variation in the first principal component (Table S4, S5). The abaxial leaf DCA, the air pollution proxy leaf SIRM, the proportion of GI (in all buffer zones tested) and the proportion of industry & harbour land use (in buffer zones with radius between 50 and 300 m) were significantly related with the second principal component (Table S4, S5).

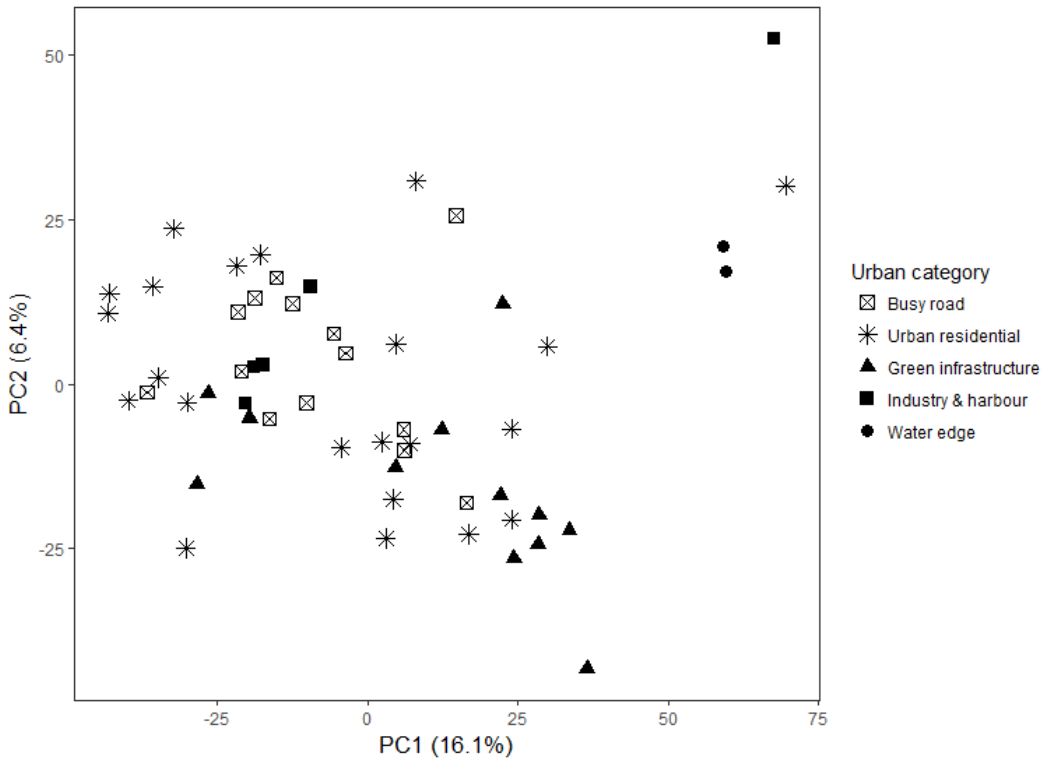

Fig. 4. Biplot of the PCA on the composition of the whole phyllosphere bacterial communities of each tree (n = 55), marked according to the urban category in which the sampled tree was located.

Abundant taxa related to landscape and leaf traits

Significant differences in clr-abundances between urban categories occurred for several of the abundant genera and families, i.e. with relative abundance >1% in at least three out of the 55 sites (Table 1). Pairwise comparison revealed that these differences are a reflection of significantly different clr-abundances between urban green sites on the one hand and busy roads, urban residential sites and/or industrial & harbour sites on the other hand. Trees in urban green sites differed most from busy road side trees, the first showing significantly higher clr-abundances of *Hymenobacter, Acidobacteraceae* and *Cystobacterineae* and significantly lower clr-abundances of *Acinetobacter, Flavisolibacter, Kaistobacter, Ralstonia, Rubellimicrobium, Skermanella* and *Geodermatophilaceae* than the latter (Fig. 5). In addition, indicator species analysis was applied to the clr-abundances of the most abundant families and genera to identify the taxa associated with the urban categories at the sites and indicated that *Hymenobacter* and *Sphingomonas* were associated with urban green sites (IndVal 0.345 and 0.288, respectively), while *Skermanella* was associated with industry & harbour sites (IndVal 0.396).

Table 1. P values (corrected for false-discovery rates) of the urban category effect on centered log-ratio abundances using ALDEx2 and p values (corrected for false-discovery rates) of Pearson correlations between centered log-ratio abundances of the most abundant genera and families (with relative abundance >1% in at least three out of the 55 sites) with continuous variables found to be significantly related to one of the first two PCA axes (RCC=relative chlorophyll content; DCA_{ab}=drop contact angle of abaxial leaf side). The proportions of green infrastructure (GI) and industry & harbour (IH) were calculated in buffer zones with a radius of 100 m, while those of open water (OW) were calculated in 500 m buffer zones, following the most significant relationships with the PCA axes. Blue indicates a significant negative correlation, red a positive one.

	Urban category	Leaf traits			Landscape traits		
		RCC	DCA_{ab}	Leaf SIRM	Prop. GI	Prop. IC	Prop. OW
Most abundant genera							
Acinetobacter	**0.0331**	0.4157	0.0952	0.0868	0.0225	0.1236	0.7361
Agrobacterium	0.7905	0.4373	0.7385	0.6846	0.2449	0.8160	0.7961
Arthrobacter	0.1564	0.2440	0.0344	0.7540	0.1166	0.7767	0.2726
Beijerinckia	0.0644	0.2440	0.0439	0.1484	0.0344	0.0842	0.4282
Chroococcidiopsis	**0.0060**	0.0269	0.3907	0.8213	0.0012	0.4705	0.0779
Chryseobacterium	0.4507	0.5967	0.7361	0.1476	0.6724	0.4214	0.3806
Clostridium	0.0859	0.6272	0.6703	0.3540	0.4705	0.4081	0.5355
Deinococcus	0.4357	0.0892	0.2842	0.2842	0.5967	0.4430	0.5741
Escherichia	0.9607	0.4422	0.3646	0.3829	0.6157	0.4206	0.6157
Flavisolibacter	**0.0018**	0.2127	0.0220	0.2088	0.0012	0.2467	0.5004
Hymenobacter	**0.0088**	0.5277	0.6371	0.5536	0.0012	0.6214	0.3646
Janthinobacterium	0.0840	0.3907	0.3716	0.6846	0.1944	0.8213	0.7518
Kaistobacter	**0.0135**	0.3646	0.0868	0.0779	0.0071	0.2301	0.4214
Methylobacterium	0.4554	0.4863	0.5105	0.7878	0.6157	0.5710	0.3646
Pantoea	0.4650	0.5758	0.4214	0.5536	0.7840	0.4157	0.5105
Pedobacter	0.7166	0.4214	0.5859	0.6846	0.5105	0.4980	0.5277
Polaromonas	0.3366	0.6598	0.5201	0.4705	0.3716	0.5967	0.6724
Pseudomonas	0.1043	0.3806	0.3242	0.5105	0.3540	0.5710	0.7200
Pseudonocardia	0.9781	0.6499	0.2301	0.6499	0.6547	0.6332	0.5583
Ralstonia	**0.0059**	0.4575	0.0554	0.0337	0.0437	0.1755	0.6547
Rickettsiella	**0.0433**	0.5175	0.7340	0.2114	0.7010	0.3242	0.0337
Rubellimicrobium	**0.0279**	0.0496	0.0365	0.5583	0.0344	0.1786	0.3896
Serratia	0.2770	0.2467	0.4289	0.6859	0.2012	0.5386	0.2503
Skermanella	**0.0004**	0.0536	0.0496	0.2302	0.0439	0.2043	0.1476
SMB53	0.1401	0.5105	0.5105	0.5241	0.5972	0.2302	0.6846
Sphingomonas	**0.0453**	0.4475	0.1813	0.7010	0.0021	0.5060	0.7909
Spirosoma	0.1118	0.1951	0.5993	0.8213	0.6846	0.4567	0.3716
Most abundant families							
Acetobacteraceae	0.9516	0.6157	0.3907	0.6805	0.3973	0.7305	0.5105
Acidobacteriaceae	0.0799	0.4025	0.0186	0.2440	0.0397	0.1628	0.2948
Caulobacteraceae	0.3412	0.7406	0.2012	0.2440	0.5082	0.4373	0.7998
Chitinophagaceae	0.1298	0.2503	0.3530	0.6724	0.2440	0.6214	0.0725
Clostridiaceae	0.0889	0.2938	0.6846	0.8102	0.3806	0.6157	0.3646
Comamonadaceae	0.0652	0.7200	0.3646	0.5940	0.6157	0.6846	0.2012
Cystobacterineae	**<0.0001**	0.2760	0.6846	0.3060	0.0002	0.2755	0.0894
Ellin6075	0.0660	0.2948	0.1166	0.6846	0.0439	0.3540	0.3788
Enterococcaceae	0.1971	0.6875	0.3806	0.5967	0.6703	0.6157	0.3052
Geodermatophilaceae	**0.0002**	0.3716	0.0344	0.0521	0.0012	0.1836	0.6157
Intrasporangiaceae	0.2266	0.4567	0.0706	0.4529	0.0541	0.4214	0.0677
Isosphaeraceae	0.3707	0.6157	0.1065	0.5818	0.1166	0.1166	0.7010
Kineosporiaceae	0.0808	0.6214	0.6157	0.5511	0.0223	0.6157	0.1376
Lactobacillaceae	**0.0464**	0.5758	0.6157	0.0439	0.3540	0.3646	0.7385
Methylocystaceae	0.2284	0.6157	0.0868	0.3646	0.1166	0.0698	0.4438
Microbacteriaceae	0.3062	0.4430	0.8213	0.5105	0.6846	0.7010	0.8160
Micromonosporaceae	0.5875	0.7865	0.2129	0.1944	0.6908	0.6730	0.7870
Nocardioidaceae	0.1989	0.6157	0.0668	0.4214	0.1341	0.4914	0.0962
Oxalobacteraceae	0.2047	0.1687	0.5758	0.6875	0.0541	0.8239	0.1813
Planococcaceae	0.1091	0.5967	0.0668	0.7573	0.0439	0.1786	0.6724
Simkaniaceae	0.1749	0.5737	0.1166	0.5752	0.2012	0.5489	0.4579
Sphingobacteriaceae	0.3817	0.8213	0.5940	0.2427	0.3672	0.1686	0.7850
Xanthomonadaceae	0.2330	0.7870	0.3806	0.3907	0.0441	0.3716	0.7865

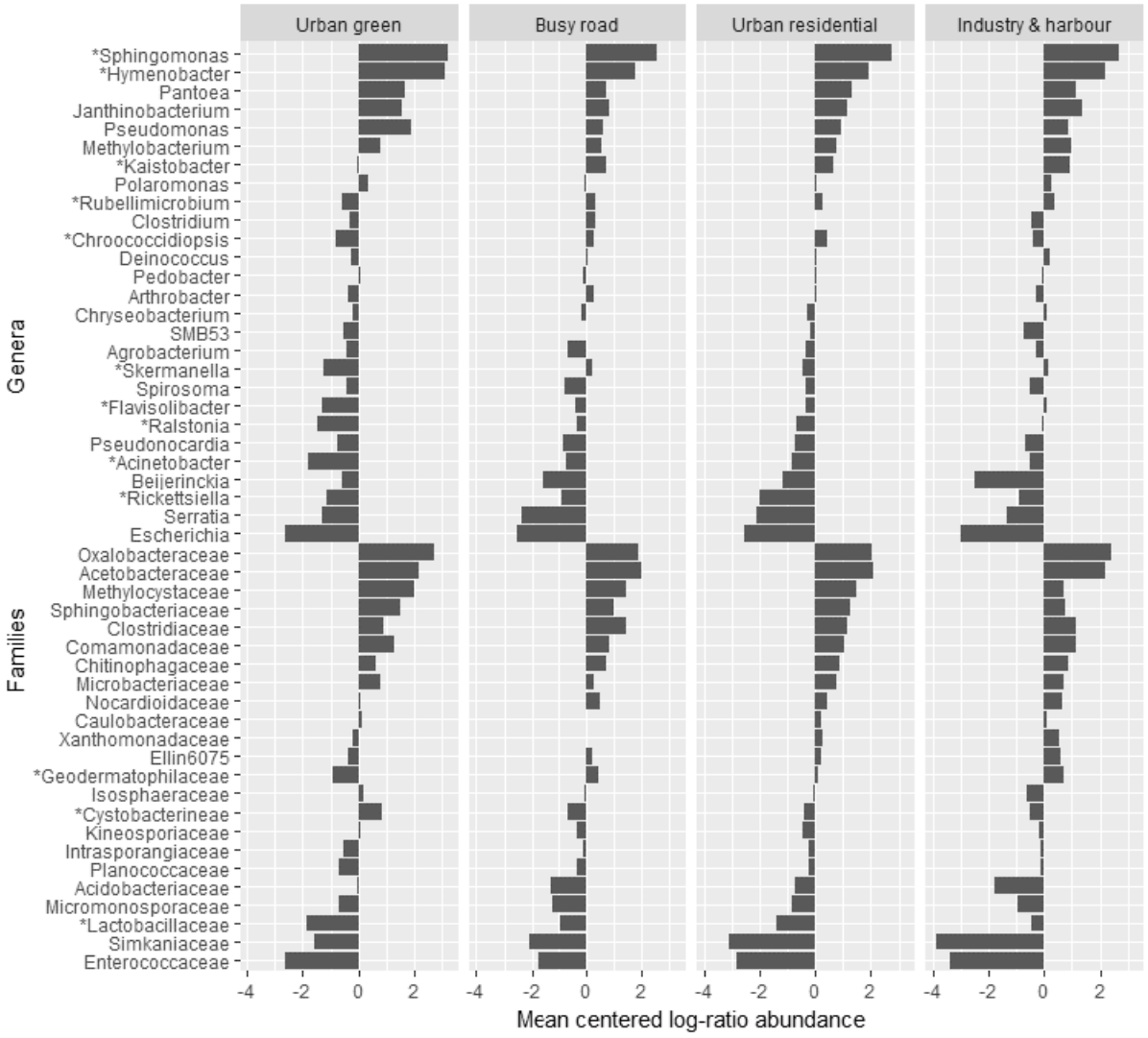

Fig. 5. Mean centred log-ratio abundance of the most abundant genera and families (in decreasing order of overall mean clr-abundance in all samples) in four urban categories, i.e. busy road (n=14), green area (n=12), industry & harbour (n=5), and urban residential (n=22) (water edge sites were excluded because of too low number of sites). Negative (respectively positive) values indicate that, on average, the number of reads of that taxon was lower (respectively higher) than the mean number of reads of all 50 most abundant taxa in the samples from sites in the given urban category. *: p value corrected for false discovery rates ≤ 0.05 indicating a significant effect of urban category at the site.

Plenty of the more abundant genera and families related with the leaf or landscape traits that were found to significantly relate with one of the two PCA axes (Table 1). Three genera - *Hymenobacter*,

Sphingomonas and *Beijerinckia* – and three families showed a positive relation with the proportion of GI within a 100 m radius, of which *Sphingomonas* and *Hymenobacter* are very abundant core community members (Fig. S1). Negative relationships with GI were observed in seven genera and four families (Table 1). The most negative relationships of clr-abundance with GI were observed for the genera *Chroococcidiopsis*, *Flavisolibacter*, *Acinetobacter* and *Skermanella,* and for the family *Geodermatophilaceae*. The relative abundances of these taxa were already low in sites with less than 20% GI (median 0.9%), and decreased even further to negligible numbers in sites with 60-80% GI (median 0.1%; Fig S1). The clr-abundance of none of the most abundant taxa related with the proportion of industrial and harbour area within a 500 m radius, while one genus *Rickettsiella* related with the proportion of open water in a radius of 500 m (Table 1).

The leaf SIRM related significantly and positively with the clr-abundances of one abundant genus, i.e. *Ralstonia* (Table 1), while the leaf RCC related with the clr-abundance of two genera and the DCA at the abaxial side with the clr-abundance of five genera and two families. The clr-abundance of all but one (*Arthrobacter*) of the genera and families that relate to the abaxial DCA show opposite relationships with the proportion of GI (Table 1). There was indeed a tendency of lower abaxial DCA at higher proportions of GI, but this relationship was not significant (τ=-0.11, p=0.2982).

Sixty-two OTUs occurred in all samples and were considered to be the core microbiome of the phyllosphere of Antwerp's urban plane trees (Table S6). Of these 62 OTUs, the highest relative abundances were observed for the genera *Sphingomonas*, *Hymenobacter*, *Pantoea* (*P. agglomerans*), *Janthinobacterium* and *Pseudomonas* (*P. fragi*), in order of decreasing abundance. The clr-abundance of the core community (all core OTUs together) related positively with the proportion of GI within a 100 m radius (p=0.0246) and differed significantly between urban categories (p=0.0098). According to pairwise comparisons, the core community was significantly

higher in clr-abundance in the green sites than at busy roads (p=0.0083) and the sites in urban residential areas (p=0.0400).

Discussion

Differences between urban categories

The bacterial phyllosphere diversity was strongly affected by the urban category the tree occurred in, but not as we expected. In contrast to our hypotheses, the bacterial OTU richness in the phyllosphere of trees at the urban green sites was only half as high as trees in urban residential areas, at busy roads and in industry & harbour sites. Trees in different urban categories also harboured distinct PBC compositions. Differences in taxon abundances occurred mainly between urban green sites on the one hand and urban residential sites, busy roads and/or industry & harbour sites on the other hand, with higher abundances of *Hymenobacter*, *Acidobacteraceae* and *Cystobacterineae*, and lower abundances of *Acinetobacter*, *Chroococcidiopsis*, *Flavisolibacter*, *Kaistobacter*, *Ralstonia*, *Rubellimicrobium*, and *Skermanella* at the green sites. The PBC in the urban green sites mainly consisted of the core community (68% relative abundance), i.e. genera that occur in the phyllosphere of all trees and dominated by *Sphingomonas, Hymenobacter, Janthinobacterium*, *Pantoea* (*P. agglomerans*) and *Pseudomonas* (*P. fragi*). *Hymenobacter* are often reported to occur in phyllosphere communities, even in urban settings (e.g. Gandolfi et al. 2017; Smets et al. 2016b; Laforest-Lapointe et al. 2017a), but their role is currently unknown. *Sphingomonas* is usually a dominant genus in the phyllosphere, because its members are able to use a wide variety of substrates characterizing this habitat (Delmotte et al., 2009; Kim et al., 1998; Vorholt, 2012). *Sphingomonas spp.* have previously been found to protect the plant against plant pathogens, probably by substrate competition (Innerebner et al., 2011). Hence, our observations imply that trees in non-green urban categories could be more vulnerable to certain plant pathogens.

The urban categories differed in the proportion of GI in the surrounding landscape as well as in air pollution exposure (as indicated by the leaf SIRM), which both show to relate with changes in diversity and composition of the bacterial communities. Therefore, we presume that the differences between urban categories are mainly related to effects of the proportion of GI and air pollution.

Green infrastructure in the surrounding landscape

The higher the proportion of (urban) GI in the surroundings of the sampling sites and the smaller the distance to the nearest GI, the lower the alpha diversity of the community of phyllosphere bacteria sampled on urban plane trees, contradicting our first hypothesis of lower richness in phyllosphere bacteria in areas with less (and further from) green infrastructure due to reduced dispersal. In fact, the significant positive relationship between the nestedness-resultant β_{nes} with the dissimilarity in the proportion of urban green indicates that the bacterial community on trees in areas with more GI is a subset of the communities that occur at sites with less green. The presence of more and closer GI thus reduces the diversity of the urban PBC. This conclusion endorses the findings of a study conducted on urban PBCs in Montreal, Canada (Laforest-Lapointe et al., 2017a), but is opposite to what is observed in urban areas for taxa such as insects and birds, of which the richness and diversity increase with increasing proportion of GI (Beninde et al. 2015). When looking into community composition, we found that as the proportion of GI around the tree decreased, in favour of other, more anthropogenically-disturbed land-cover classes, plenty of genera were introduced in the community while the relative abundances of the core community, and several of its members, decreased. More specifically, the relative abundances of *Hymenobacter* and *Sphingomonas* (both important members of the phyllosphere core community), the genus *Beijerinckia* and three families including the *Acidobacteriaceae* decreased with decreasing proportion of GI in a radius of 100 m around the tree, but *Hymenobacter* and *Sphingomonas* were still highly abundant in areas with no or low (<20%) GI proportion (Fig. S1). In contrast, many of

the abundant genera became more abundant in trees with less GI and more anthropogenically-disturbed land cover while they were virtually absent in areas with high GI proportion (>80%). The most pronounced increases in relative abundance with decreasing GI were for the genera *Chroococcidiopsis*, *Flavisolibacter* and *Acinetobacter*. *Chroococcidiopsis* are the most common members of the photosynthesising phylum of the cyanobacteria. This genus is known for their high desiccation tolerance, occurring in soils and on rock surfaces worldwide, but their presence was also observed in the phyllosphere of Atlantic forests (Brewer and Fierer, 2017; Büdel, 2011; Rigonato et al., 2016). *Flavisolibacter* are generally associated with soil (Yoon and Im, 2007), whereas *Acinetobacter* are mainly known as human skin microbiota and human pathogens, but are also found in soil, water and sewage (Towner, 2006). Hence, the taxa that are increased in urban, non-green-infrastructure land cover types seem to have a more generalist lifestyle than typical phyllosphere bacteria and potentially originate from typical urban sources (e.g. "rock" surfaces of urban infrastructure, soil, people). Due to the limitations of *16S rRNA* amplicon sequencing, it is however difficult to determine if they are active members of the urban phyllosphere community.

Although the distance to the nearest GI also related with alpha diversity (i.e., increasing diversity with increasing distance), it was negatively correlated with proportion of GI and provided a less good fit with alpha diversity indices than GI proportion. Also, the number of GI patches related with the alpha diversity indices, but not its Euclidean Nearest Neighbour distance. Moreover, the relationships of the differences in distance to the nearest green with the calculated β diversity indices were not significant. These three separate observations all suggest that the amount of GI rather than distance to the nearest GI is a factor influencing phyllosphere bacterial diversity and composition, indicating the importance of habitat change rather than reduced connectivity. This complies with observations by Turrini & Knop (2015) on other city-dwelling organisms, i.e. arthropods.

Moreover, the lack of a significant relationship between geographic distance between the trees and their difference in leaf bacterial community confirms that connectivity does not play an important role in the dispersal of their bacteria at city scale. The absence of such a relationship seems in contrast with the study by Finkel et al. (2012) on Tamarix. In addition, Rastogi et al. (2012) and Knief et al. (2010) found that plants grown in the same geographical location supported highly similar bacterial communities compared to those at other sites.

From the four processes that drive community dynamics (Vellend, 2010; Vacher et al., 2016), dispersal and speciation regulate the addition of species to communities. Speciation as a process of community dynamics can be excluded when a sufficient degree of dispersal is present (Nemergut et al., 2013), which, as described above, seems to be the case in this study (at city scale). Dispersal is generally accepted as an important driver of lower richness and diversity of animal taxa in an urban environment when the proportion, size and connectivity of GI is low (Beninde et al., 2015; Öckinger et al., 2009). The introduction of bacterial species in trees in areas with a low GI thus seems contradictory to the dispersal-limiting effect of low proportions of GI. However, sources of phyllosphere bacteria are not limited to the phyllosphere of other plants, but also encompass air, soil, water or other tree parts through airborne, waterborne or vectorborne deposition (Beattie & Lindow, 1999; Whipps et al., 2008). For phyllosphere, all land cover types could be potential bacterial emission sources, not only vegetation in GI but also any non-green land cover. These land cover types could be all sorts of anthropogenically developed or influenced land cover, as none of the non-green urban land-cover classes in specific (e.g., urban fabric, roads) related significantly with richness, diversity or composition of the PBC. The species with higher abundances in the phyllosphere of the trees in the green areas could be considered 'autochthonous' plane-phyllosphere-specific core species, adapted to persist in the phyllosphere, such as *Hymenobacter* and *Sphingomonas*. The bacteria coming from non-green land cover types enrich the urban phyllosphere and could be considered exotic to the tree-species-specific phyllosphere core

community. They seem to be adapted to highly disturbed sites and able to easily colonise new sites, like ruderals. However, they have a lower presence in the phyllosphere of trees in green buffer zones. This indicates that in the greener urban buffer zones, the selection pressure of the phyllosphere environment on the exotic species outweighs the rate of incoming exotic species by dispersal, either through less dispersal or through altered selection pressure. Dispersal in a city may be affected by local increases in wind-flow and turbulence within the built-up areas, due to channel effects, surface heating, and moving traffic (Belcher, 2005). As such, GI in cities could act as 'buffers' against the (atmospheric) dispersal of bacteria coming from other land cover/use types in the surroundings. Alternatively, the phyllosphere selection process may be affected by environmental factors. For example, in urban residential and commercial areas, humidity is lower and the urban heat-island effect is more pronounced in more built-up areas in comparison with more vegetated areas (Weng et al., 2004; Wang et al., 2015), making the environment in areas poor in GI harsher for typical phyllosphere bacteria to live in. Instead, complementary species that can cope better with drought or high temperatures (such as *Chroococcidiopsis*, its abundance relating negatively with the proportion of GI), are less affected and encounter less competition with sensitive species. Likewise, enhanced trophic interactions with other epiphytic micro-organisms such as fungi in the phyllosphere of trees in more humid conditions of urban green cannot be excluded. The phyllosphere selection process may also be affected by the traits of the plant host. The DCA describes the leaf wettability – the leaf surface's affinity with water: at high wettability, water droplets tend to spread (displaying low DCA), ultimately forming water films, while at low wettability, isolated drops (with high DCA) are formed (Shephard & Griffiths, 2006). For six out of the 17 abundant genera and families of which the clr-abundance related to the proportion of GI, a significant relation with the abaxial DCA was observed, most of them positive and all of opposite sign than the relation with GI. Hence it seems that water availability for phyllosphere bacteria in urban non-green environments is lower, due to both lower air humidity and decreased leaf

wettability. However, there are direct and indirect ways to link the leaf wettability with the abundances of several phyllosphere bacteria. Firstly, the growth and population density of some leaf bacteria are known to interact with the DCA, acting either as driver or response (Whipps et al., 2008; Knoll & Schreiber, 2000). Secondly, an indirect positive effect of the proportion of GI on the leaf wettability is possible. Leaf wettability is influenced by the leaf cuticular waxes, of which production, morphology and composition depend on shading, temperature, humidity, mechanical and water stress and air pollution (Shephard & Griffiths, 2006), i.e. environmental factors that change in relation with the proportion of GI in the urban surroundings. Although more research is needed to allow prediction of the complex interactions between different phyllosphere bacteria, landscape traits, and leaf properties such as leaf wettability, it is likely that different selection pressures occur in the urban phyllosphere of non-green and green environments. As a result, the incoming species exotic to the tree-species-specific phyllosphere core community die off or only settle for a short period of time and leave the phyllosphere again by resuspension processes such as evapotranspiration. Presumably, a combination of both dispersal limitation and altered selection pressure are driving the observed diversity patterns in urban phyllosphere communities.

The richness and diversity of our phyllosphere communities did, however, not relate with land-cover richness and diversity within the different buffer zones, so only the proportion of non-green land cover seems important, not its diversity. This rejects our second hypothesis, which associated an increase in PBC diversity in areas with increasing land use/cover diversity.

Air pollution

On top of the effect of the proportion of GI leading to community variation due to nestedness, shifts in community composition occurred in the bacterial communities in form of taxa turnover in which species disappear and are replaced by others that appear. This taxa turnover related with leaf SIRM

and modelled NO_2 concentrations. This is confirmed by the significant relation of leaf SIRM with the second PCA axis of clr-abundances, indicating shifts in community composition, and the lack of a relation with OTU richness and Shannon diversity, indicating no net loss of taxa. Leaf SIRM is a proxy metric for air pollution (Hofman et al., 2017). It is shown to relate with leaf-deposited PM and trace element content in particles (Hofman et al., 2014, Castanheiro et al., 2016) and to be an indicator time-integrated exposure to PM_{10}, NO_x concentrations, particle-bound trace elements and polycyclic aromatic hydrocarbons co-emitted with, and/or adsorbed onto, particles. So, it mainly relates with exposure to combustion exhaust emissions by motorized traffic but also by industrial activity (Hansard et al., 2011). Our data thus confirm the third hypothesis that urban air pollution favours some taxa and sets back others. Our data suggest that exposure to air pollution relates with changes in PBC composition, but not in richness, and thus relates with turnover of bacterial taxa, in a balanced equilibrium of taxa gain and taxa loss. Unexpectedly, only two of all most abundant taxa, the genus *Ralstonia* and the *Lactobacillaceae*, showed their clr-abundance to be significantly increased by leaf SIRM, and no abundant taxa decreased with leaf SIRM. So, most of the taxa turnover with leaf SIRM encompasses the less abundant taxa. Air pollution is hypothesized to affect phyllosphere bacteria either directly as a resource or stressor or indirectly via changes in leaf characteristics. Some natural phyllosphere bacteria can degrade mono-aromatic hydrocarbons, and can use them as a resource (Sangthong et al., 2016; Sandhu et al., 2007). Leaf SIRM also is a discriminatory tool for atmospheric metal pollution, as it relates well with the Fe, Pb, Zn, Mn and Cd content in leaves and in leaf-deposited particles (Castanheiro et al., 2016; Maher et al., 2008). Most of these heavy metals are toxic for some bacteria, while other bacteria show resistance to heavy metal toxicity (Nies 1999). For example, many *Ralstonia* strains have been reported to be resistant to metals (Goris et al. 2001), but the genus is highly diverse and we detected only one unknown OTU of the *Ralstonia* genus in the plane phyllosphere. Indirect effects of air pollution, through changes in leaf traits such as stomatal density, leaf wettability and specific leaf

area, mainly in response to ozone and sulphur dioxide (Cape, 1983; Elagöz et al., 2006, Shephard & Griffiths, 2006; Wuytack et al., 2011), could also affect phyllosphere bacteria. Vehicle density in a radius of 50 m did relate with OTU richness, showing higher richness as more motorized vehicles pass by each hour, but not with the PBC structure, seemingly contrary to the leaf SIRM with which it correlated. The reason probably is that vehicle density has a minor effect on the PBC through its positive relation with the proportion of GI, and although it relates with leaf SIRM, the latter indicates air pollution exposure from more than motorized traffic alone, including other PM sources such as industry and households. The relation between proportion of GI and leaf SIRM might explain the results of our previous study on *Hedera* sp. phyllosphere in and around the same city (Smets et al., 2016b), where the relative abundance of eight out of the ten most abundant taxa related with the leaf SIRM.

To conclude, we found that an increasing amount of nearby GI is related with a loss of bacterial taxa in the urban phyllosphere, whereas the degree of air pollution was significantly linked with the replacement of several taxa. These results imply that the urban phyllosphere in areas (with a radius of 200 m) with less vegetation are affected by high rates of incoming "exotic" bacterial taxa. Both dispersal and selection could explain this observation, as both nearby bacterial sources and variation in environmental factors depend on the surrounding land use. Besides urban land use, air pollution does seem to have its separate effect on the phyllosphere communities, possibly through direct selection pressure and/or indirect effects on host plant traits. Through their effects on PBC diversity and composition (Laforest-Lapointe et al., 2017b), it can be expected that air pollution and green infrastructure influence the ecosystem functioning of urban green and services provided by urban trees. The mechanisms involved could be, among others, atmospheric nitrogen fixation (Moyes et al. 2016), degradation of atmospheric pollutants (Sandhu et al. 2007), protection from

plant pathogen infection (Innerebner et al. 2011) and aerosolization of human pathogens (Mills et al. 2017), but this hypothesis requests further investigation.

Acknowledgements

The authors thank the Centre for Medical Genetics Antwerp for the use of their Illumina MiSeq sequencing system and Jolien Verhelst for access to the database of London plane trees in Antwerp, assembled in the HYPERCITY project (BELSPO Research Programme for Earth Observation – STEREO III, contract SR/00/303). This research was financially supported by the University of Antwerp, through a.o. the University Research Fund (KPBOF 2014, no. FFB 140090 'Tree leaf surface properties as dynamic drivers of particulate matter-leaf interaction and phyllosphere microbial communities'), and by the Agency for Innovation by Science and Technology in Flanders (ProCure project [IWT/50052]). We thank the anonymous reviewers for their insightful comments which helped us to improve the manuscript.

Conflict of interest

The authors declare no conflict of interest, financial or otherwise.

References

Alberti M, Correa C, Marzluff JM, Hendry AP, Palkovacs EP, Gotanda KM, et al. (2017). Global urban signatures of phenotypic change in animal and plant populations. Proceedings of the National Academy of Sciences of the United States of America 114, 8951–8956.

Anderson MJ, Crist TO, Chase JM, Vellend M, Inouye BD, Freestone AL, et al. (2010). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. Ecology Letters 14, 19–28.

Baselga A (2013). Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients. Methods in Ecology and Evolution 4, 552–557.

Baselga A, Orme D, Villeger S, De Bortoli J, Leprieur F (2017). betapart: Partitioning Beta Diversity into Turnover and Nestedness Components. R package version 1.4-1. https://CRAN.R-project.org/package=betapart

Belcher SE (2005). Mixing and transport in urban areas. Philosophical Transactions of the Royal Society A 363, 2947–2968.

Beninde J, Veith M, Hochkirch A (2015). Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. Ecology Letters 18, 581–592.

Benjamini Y, Hochberg Y (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. Journal of Educational and Behavioral Statistics 25, 60–83.

Bradley DJ, Gilbert GS, Martiny JB (2008). Pathogens promote plant diversity through a compensatory response. Ecology Letters 11, 461−469.

Brewer TE, Fierer N (2017). Tales from the tomb: the microbial ecology of exposed rock surfaces. Environmental Microbiology 20, 958−970.

Brighigna L, Gori A, Gonnelli S, Favilli F (2000). The influence of air pollution on the phyllosphere microflora composition of *Tillandsia* leaves (Bromeliaceae). International Journal of Tropical Biology and Conservation 48, 511−517.

Büdel B (2011). Cyanobacteria: habitats and species. Springer, 11–18 pp.

Cape N (1983). Contact angles of water droplets on needles of Scots pine (*Pinus sylvestris*) growing in polluted atmospheres. New Phytologist 93, 293–299.

Castanheiro A, Samson R, De Wael K (2016). Magnetic- and particle-based techniques to investigate metal deposition on urban green. Science of the Total Environment 571, 594–602.

Delmotte N, Knief C, Chaffron S, Innerebner G, Roschitzki B, Schlapbach R, et al. (2009). Community proteogenomics reveals insights into the physiology of phyllosphere bacteria. Proceedings of the National Academy of Sciences 106, 16428−16433.

Edgar RC (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods 10, 996−998.

Elagöz V, Han SS, Manning WJ (2006). Acquired changes in stomatal characteristics in response to ozone during plant growth and leaf development of bush beans (*Phaseolus vulgaris* L.) indicate phenotypic plasticity. Environmental Pollution 140, 395–405.

Ensing DJ, Pither J (2015). A novel multiple-site extension to pairwise partitioned taxonomic beta diversity. Ecological Complexity 21, 62–69

Fernandes AD, Macklaim JM, Linn T, Reid G, Gloor GB (2013). ANOVA-like differential expression (ALDEx) analysis for mixed population RNA-seq. PLoS ONE 8, e67019.

Fernandes AD, Reid JN, Macklaim JM, McMurrough TA, Edgell DR, Gloor GB (2014). Unifying the analysis of high-throughput sequencing datasets: Characterizing RNA-seq, 16S rRNA gene sequencing and selective growth experiments by compositional data analysis. Microbiome 2, 15.1–15.13. doi:10.1186/2049-2618-2-15.

Finkel OM, Burch AY, Elad T, Huse SM, Lindow SE, Post AF, et al. (2012). Distance-decay relationships partially determine diversity patterns of phyllosphere bacteria on Tamarix trees across the Sonoran Desert. Applied and Environmental Microbiology 78, 6187–6193.

Gandolfi I, Canedoli C, Imperato V, Tagliaferri I, Gkorezis P, Vangronsveld J, et al. (2017). Diversity and hydrocarbon-degrading potential of epiphytic microbial communities on *Platanus x acerifolia* leaves in an urban area. Environmental Pollution 220, 650−658.

Geographic Distance Matrix Generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/open_source/gdmg.

Gloor GB, Macklaim JM, Fernandes AD (2016). Displaying Variation in Large Datasets: a visual summary of effect sizes. Journal of Computational and Graphical Statistics 25, 971–979.

Gloor GB, Macklaim JM, Pawlowsky-Glahn V, Egozcue JJ (2017). Microbiome datasets are compositional: and this is not optional. Frontiers in Microbiology 8, 2224.

Goris J, De Vos P, Coenye T, Hoste B, Janssens D, Brim H, et al. (2001). Classification of metal-resistant bacteria from industrial biotopes as *Ralstonia campinensis* sp. nov., *Ralstonia metallidurans* sp. nov. and *Ralstonia basilensis* Steinle et al. 1998 emend. International Journal of Systematic and Evolutionary Microbiology 51, 1773–1782.

Grimm NB, Faeth SH, Golubiewski NE, Redman CL, Wu J, Bai X, et al. (2008). Global change and the ecology of cities. Science 319, 756-760.

Hansard R, Maher B, Kinnersley R (2011). Biomagnetic monitoring of industry-derived particulate pollution. Environmental Pollution 159, 1673-1681.

Hofman J, Maher BA, Muxworthy AR, Wuyts K, Castanheiro A, Samson R (2017). Biomagnetic monitoring of atmospheric pollution: a review of magnetic signatures from biological sensors. Environmental Science and Technology 51, 6648−6664.

Hofman J, Wuyts K, Van Wittenberghe S, Brackx M, Samson R (2014). On the link between biomagnetic monitoring and leaf-deposited dust load of urban trees: Relationships and spatial variability of different particle size fractions. Environmental Pollution 189, 63−72.

Innerebner G, Knief C, Vorholt JA (2011). Protection of Arabidopsis thaliana against leaf-pathogenic Pseudomonas syringae by Sphingomonas strains in a controlled model system. Applied and Environmental Microbiology 77, 3202−3210.

Jung M (2016). LecoS - A python plugin for automated landscape ecology analysis, Ecological Informatics 31, 18-21. http://dx.doi.org/10.1016/j.ecoinf.2015.11.006

Kardel F, Wuyts K, Maher B, Hansard R, Samson R (2011). Leaf saturation isothermal remanent magnetization (SIRM) as a proxy for particulate matter monitoring: inter-species differences and in-season variation. Atmospheric Environment 45, 5164−5171.

Kardel F, Wuyts K, Maher B, Samson R (2012). Intra-urban spatial variation of magnetic particles: monitoring via leaf saturation isothermal remanent magnetisation (SIRM). Atmospheric Environment 55, 111−120.

Khanna K (1986). Phyllosphere microflora of certain plants in relation to air pollution. Environmental Pollution Series A, Ecological and Biological 42, 191−200.

Kim H, Nishiyama M, Kunito T, Senoo K, Kawahara K, Murakami K, et al. (1998). High population of Sphingomonas species on plant surface. Journal of Applied Microbiology 85, 731−736.

Knief C, Ramette A, Frances L, Alonso-Blanco C, Vorholt JA (2010). Site and plant species are important determinants of the Methylobacterium community composition in the plant phyllosphere. The ISME journal 4, 719−728.

Knoll D, Schreiber L (2000). Plant–Microbe interactions: wetting of Ivy (*Hedera helix* L.) leaf surfaces in relation to colonization by epiphytic microorganisms. Microbial Ecology 41, 33–42.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology 79, 5112−5120.

Laforest-Lapointe I, Messier C, Kembel SW (2017a). Tree leaf bacterial community structure and diversity differ along a gradient of urban intensity. mSystems 2, e00087−17.

Laforest-Lapointe I, Paquette A, Messier C, Kembel SW (2017b). Leaf bacterial diversity mediates plant diversity and ecosystem function relationships. Nature 546, 145−147.

Last F (1955). Seasonal incidence of Sporobolomyces on cereal leaves. Transactions of the British Mycological Society 38, 221−239.

Lindow SE, Leveau JH (2002). Phyllosphere microbiology. Current Opinion in Biotechnology 13, 238−243.

Maher BA, Moore C, Matzka J (2008). Spatial variation in vehicle-derived metal pollution identified by magnetic and elemental analysis of roadside tree leaves. Atmospheric Environment 42, 364–373.

McKinney ML (2008). Effects of urbanization on species richness: a review of plants and animals. Urban Ecosystems 11, 161−176.

Mhuireach G, Johnson BR, Altrichter AE, Ladau J, Meadow JF, Pollard KS, et al. (2016). Urban greenness influences airborne bacterial community composition. Science of the Total Environment 571, 680–687.

Mills JG, Weinstein P, Gellie NJC, Weyrich LS, Lowe AJ, Breed MF (2017). Urban habitat restoration provides a human health benefit through microbiome rewilding: the Microbiome Rewilding Hypothesis. Restoration Ecology 25, 866–872.

Mitchell R, Maher BA (2009). Evaluation and application of biomagnetic monitoring of traffic-derived particulate pollution. Atmospheric Environment 43, 2095–2103.

Moyes AB, Kueppers LM, Pett-Ridge J, Carper DL, Vandehey N, O'Neil J, Frank AC (2016). Evidence for foliar endophytic nitrogen fixation in a widely distributed subalpine conifer. New Phytologist 210, 657–668.

Muxworthy AR, Matzka J, Davila AF, Petersen N (2003). Magnetic signature of daily sampled urban atmospheric particles. Atmospheric Environment 37, 4163−4169.

Nemergut RD, Schmidt SK, Fukami T, O'Neill SP, Bilinski TM, Stanish LF, et al. (2013) Patterns and processes of microbial community assembly. Microbiology and Molecular Biology Reviews 77, 342-356.

Nies DH (1999). Microbial heavy-metal resistance. Applied Microbiology & Biotechnology 51, 730–750.

Öckinger E, Dannestam Å, Smith HG (2009). The importance of fragmentation and habitat quality of urban grasslands for butterfly diversity. Landscape & Urban Planning 93, 31–37.

Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, et al. (2016). vegan: Community Ecology Package. R package version 2.4-1. https://CRAN.R-project.org/package=vegan

Palarea-Albaladejo J, Martin-Fernandez JA (2015). zCompositions -- R package for multivariate imputation of left-censored data under a compositional approach. Chemometrics and Intelligent Laboratory Systems 143, 85–96.

Pike N (2011). Using false discovery rates for multiple comparisons in ecology and evolution. Methods in Ecology and Evolution 2, 278–282.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/

Rastogi G, Sbodio A, Tech JJ, Suslow TV, Coaker GL, Leveau JHJ (2012). Leaf microbiota in an agroecosystem: spatiotemporal variation in bacterial community composition on field-grown lettuce. The ISME Journal 6, 1812–1822.

Redford AJ, Fierer N (2009). Bacterial succession on the leaf surface: a novel system for studying successional dynamics. Microbial ecology 58, 189−198.

Rigonato J, Gonçalves N, Andreote APD, Lambais MR, Fiore MF (2016). Estimating genetic structure and diversity of cyanobacterial communities in Atlantic forest phyllosphere. Canadian Journal of Microbiology 62, 953−960.

Samson R, Veroustrate F, Verhelst J, Wuyts K, Moreno J, Alonso L, et al. (2016). HYPERCITY project – Annual report 2015. Antwerp, 33p.

Sandhu A, Halverson LJ, Beattie GA (2007). Bacterial degradation of airborne phenol in the phyllosphere. Environmental Microbiology 9, 383−392.

Sangthong S, Suksabye P, Thiravetyan P (2016). Air-borne xylene degradation by *Bougainvillea buttiana* and the role of epiphytic bacteria in the degradation. Ecotoxicology and Environmental Safety 126, 273−280.

Shephard T, Griffiths DW (2006). Tansley review: the effects of stress on plant cuticular waxes. New Phytologist 171, 469–499.

SGS, 2010. Strategische Geluidsbelastingskaarten Agglomeratie Antwerpen, 090357-2-v1. SGS Belgium NV.

Smets W, Moretti S, Denys S, Lebeer S (2016a). Airborne bacteria in the atmosphere: Presence, purpose, and potential. Atmospheric Environment 139, 214−221.

Smets W, Wuyts K, Oerlemans E, Wuyts S, Denys S, Samson R, et al. (2016b). Impact of urban land use on the bacterial phyllosphere of ivy (*Hedera* sp.). Atmospheric Environment 147, 376−383.

Stalder AF, Melchior T, Müller M, Sage D, Blu T, Unser M (2010). Low-bond axisymmetric drop shape analysis for surface tension and contact angle measurements of sessile drops. Colloids and Surfaces A: Physicochemical and Engineering Aspects 364, 72−81.

Towner K, 2006. The genus acinetobacter, The prokaryotes. Springer, pp. 746−758.

Turrini T, Knop E (2015). A landscape ecology approach identifies important drivers of urban biodiversity. Global Change Biology 21, 1652–1667.

Urban Atlas 2012, http://land.copernicus.eu/

Vacher C, Hampe A, Porté AJ, Sauer U, Compant S, Morris CE (2016). The phyllosphere: microbial jungle at the plant-climate interface. Annual Review of Ecology, Evolution, and Systematics 47, 1–24.

Vellend M (2010). Conceptual synthesis in community ecology. The Quarterly Review of Biology 85, 183–206.

Vlaamse Milieumaatschappij (2016), Luchtkwaliteit in het Vlaamse Gewest. Jaarverslag Immissiemeetnetten – 2015 [in Dutch]. Aalst, 270 p.

Vorholt JA (2012). Microbial life in the phyllosphere. Nature Reviews Microbiology 10, 828−840.

Vranckx, S., Lefebvre, W., 2013. Actualisering en verfijning luchtkwaliteitskaarten Stad Antwerpen [in Dutch]. Final Report. Study commissioned by the City of Antwerp, 2013/RMA/R/318, VITO, 97 pp.

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007). Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73, 5261−5267.

Wang Y, Bakker F, de Groot R, Wortche H, Leemans R (2015). Effects of urban trees on local outdoor microclimate: synthesizing field measurements by numerical modelling. Urban Ecosystems 18, 1305–1331.

Weng Q, Lub D, Schubring J (2004). Estimation of land surface temperature–vegetation abundance relationship for urban heat island studies. Remote Sensing of Environment 89, 467–483.

Whipps JM, Hand P, Pink D, Bending GD (2008). Phyllosphere microbiology with special reference to diversity and plant genotype. Journal of Applied Microbiology 105, 1744–1755

Wuytack T, Wuyts K, Van Dongen S, Baeten L, Kardel F, Verheyen K, Samson R (2011). The effect of air pollution and other environmental stressors on leaf fluctuating asymmetry and specific leaf area of *Salix alba* L. Environmental Pollution 159, 2405–2411.

Yoon M-H, Im W-T (2007). *Flavisolibacter ginsengiterrae* gen. nov., sp. nov. and *Flavisolibacter ginsengisoli* sp. nov., isolated from ginseng cultivating soil. International Journal of Systematic and Evolutionary Microbiology 57, 1834−1839.

Yutthammo C, Thongthammachat N, Pinphanichakarn P, Luepromchai E (2010). Diversity and activity of PAH-degrading bacteria in the phyllosphere of ornamental plants. Microbial Ecology 59, 357−368.