Atmospheric Environment 74 (2013) 291–300

Contents lists available at SciVerse ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Temporal variation in airborne microbial populations and microbially-derived allergens in a tropical urban landscape

Anthony C. Woo [a], Manreetpal S. Brar [a], Yuki Chan [a], Maggie C.Y. Lau [b], Frederick C.C. Leung [a,c], James A. Scott [d], Lilian L.P. Vrijmoed [e], Peyman Zawar-Reza [f], Stephen B. Pointing [a,g,*]

[a] School of Biological Sciences, The University of Hong Kong, Pokfulam Road, Hong Kong, China
[b] Department of Geosciences, Guyot Hall, Princeton University, Princeton, NJ 08544, USA
[c] Bioinformatics Centre, Nanjing Agricultural University, Nanjing, China
[d] Department of Medicine and Dalla Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada
[e] Department of Biology and Chemistry, City University of Hong Kong, Tat Chee Avenue, Hong Kong
[f] Department of Geography, University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[g] School of Applied Sciences, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand

HIGHLIGHTS

- Tropical urban bacterial and fungal bio-aerosols displayed pronounced seasonality.
- Bio-aerosol composition was determined by climate rather than airborne pollutants.
- A range of bacterial and fungal pathogens was detected in bio-aerosols.
- Microbially-derived allergen levels exceeded safe levels during summer months.

ARTICLE INFO

Article history:
Received 15 October 2012
Received in revised form
21 March 2013
Accepted 23 March 2013

Keywords:
Airborne microorganisms
Bio-aerosols
Bacteria
Fungi
Urban aerosol

ABSTRACT

The microbial component of outdoor aerosols was assessed along a gradient of urban development from inner-city to rural in the seasonal-tropical metropolis of Hong Kong. Sampling over a continuous one-year period was conducted, with molecular analyses to characterize bacterial and eukaryal microbial populations, immuno-assays to detect microbially-derived allergens and extensive environmental and meteorological observations. The data revealed bio-aerosol populations were not significantly impacted by the level of urban development as measured by anthropogenic pollutants and human population levels, but instead exhibited a strong seasonal trend related to general climatic variables. We applied back-trajectory analysis to establish sources of air masses and this allowed further explanation of urban bio-aerosols largely in terms of summer-marine and winter-continental origins. We also evaluated bio-aerosols for the potential to detect human health threats. Many samples supported bacterial and fungal phylotypes indicative of known pathogenic taxa, together with common indicators of human presence. The occurrence of allergenic endotoxins and beta-glucans generally tracked trends in microbial populations, with levels known to induce symptoms detected during summer months when microbial loading was higher. This strengthens calls for bio-aerosols to be considered in future risk assessments and surveillance of air quality, along with existing chemical and particulate indices.

1. Introduction

Microorganisms disperse widely in the aerosphere (Griffin, 2007; Burrows et al., 2009). Numerous studies have demonstrated the recovery of cultivable taxa and/or DNA signatures, most notably from bacteria and fungi and including pathogenic species (Griffin, 2007). Biological materials accounts for up to 25% of aerosolized matter (Jaenicke, 2005), which includes allergens (e.g. endotoxins, glucans) of microbial origin (Horner et al., 1995). The emerging field of aerobiology, the study of these bio-aerosols, has been identified as a high priority and immediacy issue in relation to not only public health, but also climate, environment, ecology, epidemiology and environmental engineering (Peccia et al., 2008).

* Corresponding author. School of Applied Sciences, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand. Tel.: +64 9921 9999; fax: +64 9921 9627.
E-mail address: steve.pointing@aut.ac.nz (S.B. Pointing).

1352-2310/$ – see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.atmosenv.2013.03.047

The occurrence of bio-aerosols within indoor built environments has received much attention (Li et al., 2007; Rintala et al., 2008; Tringe et al., 2008) due to public health concern. In contrast, outdoor environments are often well-characterized in terms of particulate or chemical pollution as part of air-quality monitoring efforts (Pöschl, 2003), and yet relatively little is known about biodiversity or the spatio-temporal dynamics of outdoor bio-aerosols in urban areas. In particular the impact of urbanization on natural microbial populations, and potential to detect possible threats of outdoor bio-aerosols to human health, are not well understood. Early studies employed cultivation approaches to characterize outdoor urban bio-aerosols, including presence of pathogenic taxa (Mancinelli and Shulls, 1978). More recently, molecular surveys have provided critical insight as to the biodiversity of these aerosols (Maron et al., 2006; Brodie et al., 2007; Frohlich-Nowoisky, 2009; Bowers et al., 2011b), but also highlight new paradigms in their ecology that have yet to be resolved.

In an urban setting, an additional factor is the contribution of microbial allergens to public health issues. There is a lack of data on the occurrence of microbial allergens (e.g. bacterial endotoxins, fungal glucans and mycotoxins) in outdoor urban environments, although their contribution to allergies from indoor air studies is well documented (Horner et al., 1995). Evidence is also emerging that urban human populations may experience elevated incidence of allergy due to lack of exposure to 'natural' environmental microbiota (Hanski et al., 2012), and so understanding the reservoir and dynamics of microbial bio-aerosols in urban areas potentially has major public health implications.

The coastal metropolis of Hong Kong provides an ideal model system against which to test hypotheses related to the impact of urbanization on bio-aerosols since a large population (approx. 7 million people) inhabit a steep and relatively well-defined gradient of urbanization, with a comprehensive network of meteorological stations that measure a broad suite of climatic and pollutant variables. A high level of regional relevance exists since Hong Kong exhibits one of the world's highest rates of respiratory allergy as a result of poor air quality (Leung, 1993; Leung et al., 1997), although the relative contribution of abiotic and biotic components to allergy symptoms is unknown. Global relevance arises as Hong Kong represents an 'extreme' case that mimics the situation likely to be faced in other cities worldwide as human population (and hence urbanization) increases. Here, we present the first multi-domain assessment of urban microbial bio-aerosols over a continuous annual cycle. We use high-throughput molecular analyses and immunoassays to detect changes in microbial populations and microbially-derived allergens. Multivariate analyses indicate that local climate and regional air masses best explained changes within urban and rural areas, and urbanization *per-se* does not affect microbial dispersal in Hong Kong. We also highlight the potential for identification of pathogenic and allergenic signals from such data sets.

2. Material and methods

2.1. Sampling locations, environmental data and bio-aerosol collection

Near-ground (2 m average elevation) air samples were collected at 4 outdoor meteorological station locations along a gradient of urbanization in Hong Kong, a seasonal-tropical metropolis with a population of approximately 7 million (Fig. 1, Supplementary Table S1). The gradient was defined according to criteria based upon long-term environmental and population monitoring, by the Environmental Protection Department of the HKSAR government. Aerosols were sampled on a weekly basis over one year during November 2008 until October 2009 with randomization within any given week to account for differences in urban inputs (i.e. weekends, public holidays). A comprehensive suite of environmental and climatic variables was measured during this period (Supplementary Table S1) for each sampling interval ($N = 192$, corresponding to 48 samples per location over a one year period). Aerosol samples were collected using a two-stage bioaerosol cyclone sampler (National Institute for Occupational Safety and Health, CDC, US) attached to universal sampling pumps (SKC Inc, 224-PCXR8, US). The sampling time was 12 h, and mean air volume collected was 2500 L for DNA and 1800 L for allergens. Filters inserted into cassette apparatus but not exposed to air-flow were used as controls. All filters and cassettes were UV-sterilized and rinsed with 70% alcohol before use.

Air fractions for endotoxin and β-glucan analysis were sampled in parallel to DNA sampling using endotoxin-free cassettes (Zefon International, 4 μm PC 37 mm 3 PC ENDOTEC). Two endotoxin/β-glucan samples were collected per month for each sampling site during an 8-month sampling period from March–October 2009 and analyzed in duplicate ($N = 64$), although a technical error (equipment malfunction during automated analysis) resulted in 25 glucan analyses being voided (final $N = 39$). All samples were stored at −80°Cuntil processed. Calibrations of the flow rate for the pumps were made every month to ensure sampling volume consistency. Meteorological data were obtained from the Hong Kong Observatory website: (http://www.hko.gov.hk/cis/data/awsext_e.htm). Pollution variables including Air Pollution Index, sulfur dioxide, respirable suspended particulates, nitrogen oxides, nitrogen dioxide, carbon monoxide and ozone were obtained from the Environmental Protection Department, Hong Kong SAR website: (http://epic.epd.gov.hk/ca/uid/airdata/p/1).

2.2. DNA recovery, PCR amplification, terminal RFLP and pyrosequencing

Total DNA was extracted directly from the filters and tubes using the DNeasy Plant Mini Kit (Qiagen, CA, US), after first washing with kit lysis buffer for 10 min. The remaining steps of the extraction were carried out according to the manufacturers instructions. Recovered DNA was quantified using Nanodrop (Thermo-scientific). The PCR reaction comprised a 25 μl PCR mixture containing 0.1–2 μl of DNA template, 0.5 μM of each primer, 2.5 units of high fidelity *Taq* polymerase (Takara), 1× PCR buffer provided by the manufacturer, 200 μM of each dNTP, and H_2O. Amplification of target genes was achieved using primer pair A341F and 1058R for Archaea (Baker et al., 2003; Nercessian et al., 2003), 341F and 907R for bacteria (16S rRNA) (Muyzer et al., 1993), NS1/NS2 (18S rRNA) and ITS1F/ITS4 (ITS) for eukarya (White et al., 1990) (Supplementary Tables S3, S4). For pyrosequencing, barcodes were assigned to forward primers to enable discrimination of samples after sequencing. The PCR reaction involved an initial denaturation time of 5 min (16S rRNA and 18S rRNA) or 3 min (ITS); 30 cycles at 95 °C for 1 min, 55 °C for 1 min (16S rRNA and 18S rRNA) or 51 °C for 1 min (ITS), 72 °C for 1 min, and a final extension at 72 °C for 10 min. Positive and negative controls were run for every PCR.

For t-RFLP analysis, PCR was performed with FAM-labeled forward 341F primer (Muyzer et al., 1993) and NS1 primer (White et al., 1990) targeting 16S rRNA and 18S rRNA respectively ($N = 192$). Restriction digests (*Msp*I and *Cfo*I for 16S/18S rRNA) of FAM-labeled PCR amplicons were analyzed by capillary electrophoresis (3730 Genetic Analyzer Biosystems). Samples were further interrogated via pyrosequencing using the Roche GS Junior System (454 Life Sciences Corp., Branford, CT, USA). For each amplicon library, purification was carried out with Agencourt AMPure XP Bead (Beckman Coulter, Inc., CA, USA) according to manufacturers instructions. The library was quantified with Quant-iT PicoGreen

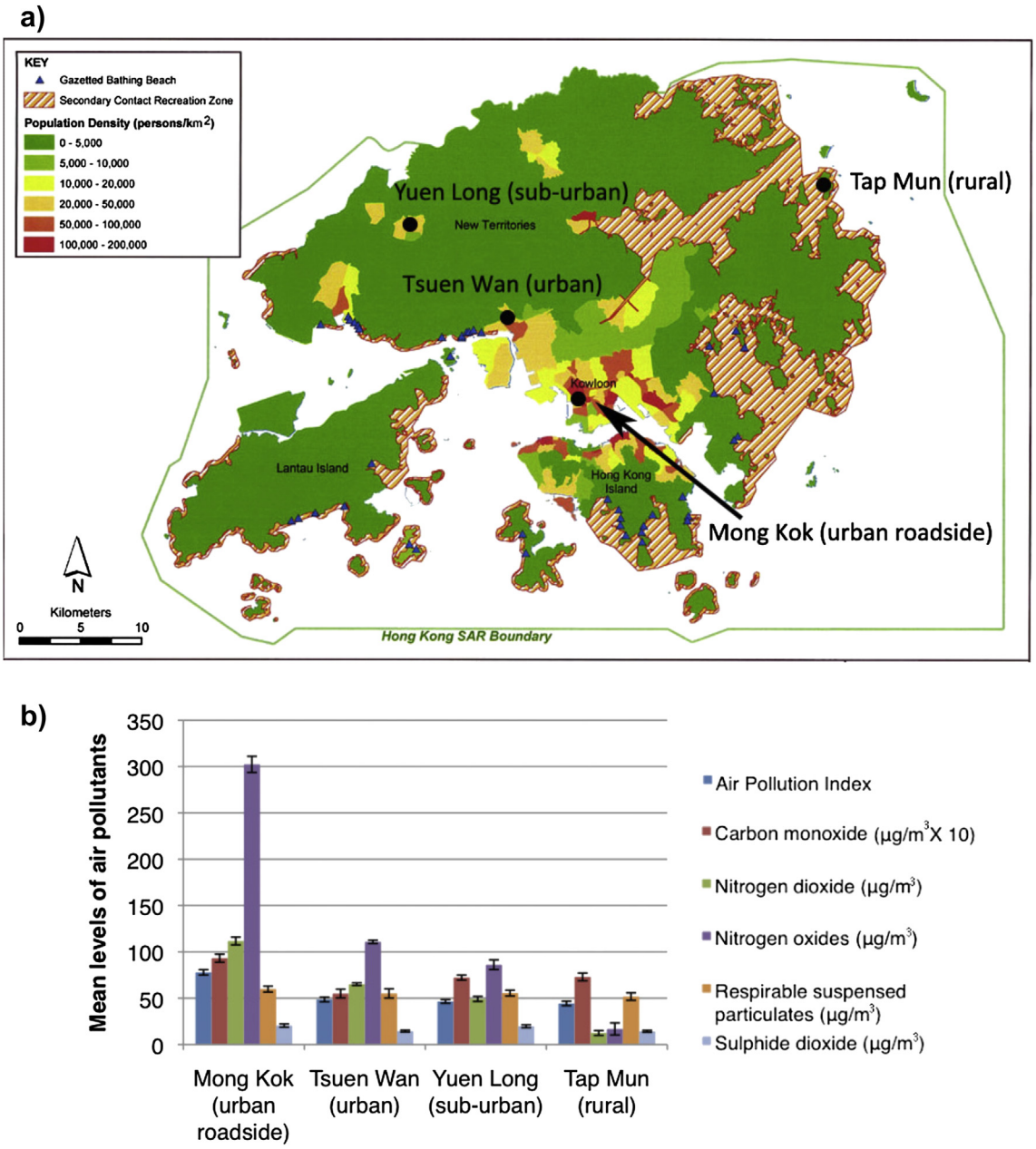

Fig. 1. (a) Map of the seasonal-tropical metropolis of Hong Kong indicating sampling locations along the gradient of population density (map courtesy HKSAR government, http://www.epd.gov.hk/epd/eindex.html). (b) Variation in key pollutants at sampling locations (values show mean annual levels; individual daily measurements were used in data analysis of temporal trends). Difference in overall pollutant loading (API) in aerosols were significant between locations (ANOVA, $F = 40.82$, $P < 0.001$).

dsDNA Assay Kit (Invitrogen Life Technologies, NY, USA) using FLUOstar OPTIMA F fluorometer (BMG Labtech GmbH, Offenburg, Germany) and library quality was assessed with FlashGel System (Lonza Group Ltd., Basel, Switzerland). Emulsion-PCR was carried out with GS Junior Titanium emPCR Kit (Lib-L, 454 Life Sciences Corp.) according to the emPCR Amplification Method Manual – Lib-L, Single-Prep. The sequencing reaction was carried out with the GS Junior Titanium Sequencing Kit and GS Junior Titanium PicoTiterPlate Kit (454 Life Sciences Corp.) according to the manufacturers instructions. The sequencing run was conducted in 200 cycles.

2.3. Sequence processing and analyses

For t-RFLP fragments, true peaks and bin fragments of similar size were identified using scripts in Perl and R (Abdo et al., 2006). ANOSIM analyses af biotic and abiotic variables were performed using PRIMER 6.1.5 (Primer-E Ltd, Plymouth,UK) and confidence intervals for ANOSIM were run with 999 random permutations.

Pyrosequencing reads were sorted according to barcoding prior to analysis and processing using the software package MOTHUR (Schloss and Handelsman, 2005). De-noising was carried out, with sequences removed from analysis if they met any of the following criteria: the length was shorter than 350 bp; with an average quality score less than 25; contained ambiguous characters or more than 6 homopolyers; did not contain the primer sequence or barcode. In order to remove sequences that were probably due to pyrosequencing errors, sequences were pre-clusted using a pseudo-single linkage algorithm as implemented in MOTHUR. After filtering and chimera check, the resulting data set contained 129,865 16S rRNA gene sequences ranging from 350 to 530 bp

(mean length 492 bp) and 81,411 ITS gene sequences ranging from 350 to 625 bp (mean length 496 bp). Chimera check was performed using UCHIME with the *de novo* mechanism (Edgar et al., 2011). Hierarchical clustering was performed with the remaining sequences to form clumps that were small enough to align using USEARCH (Edgar, 2010). A master set was created using the longest sequence from each clump. Sequences in the clumps and master set were aligned using MUSCLE (Edgar, 2004). The aligned sequences were merged into a final alignment with the master set as a guide. Alignment columns containing more than 90% (16S rRNA), 95% (ITS) gaps were trimmed using trimAL (Capella-Gutierrez et al., 2009). To correct the differences in sequencing depth among individual samples, the data sets were rarefied to 2600 sequences and 5000 sequences per sample for 16S rRNA gene sequences and ITS gene sequences respectively. Rarefaction was carried out using MOTHUR (Schloss and Handelsman, 2005) and phylogenetic trees constructed with FastTree (Price et al., 2009) to compare phylogenetic similarity between samples as calculated by the weighted UniFrac metrics (Lozupone et al., 2007). The distances in UniFrac matrix were calculated based on the fraction of branch length shared between two communities within a phylogenetic tree. Alpha diversity was assessed by constructing the rarefaction curves defined at 97% sequence similarity cutoff for O.T.Us. Taxonomic classification of 16S rRNA gene sequences was made using the ribosomal database project Classifier (Wang et al., 2007) and for 18S rRNA and ITS sequences using the SILVA database (Pruesse et al., 2007). ANOSIM, Mantel and BEST analyses of biotic and abiotic variables were performed using PRIMER 6.1.5 (Primer-E Ltd, Plymoyth,UK) and confidence intervals for ANOSIM and Mantel analyses were run with 999 random permutations. SPSS v19.0 (IBM corporation, New York, US) was used to carry out ANOVA and establish univariate models. Phylogenetic analysis of chlorophyte sequences was performed with the on-line version of MAFFT (Katoh et al., 2009) using the neighbor joining algorithm. Sequence data have been deposited in NCBI's sequence read archive under accession number SRA052054.1.

2.4. Back trajectory analysis

Back trajectory analysis was achieved using the internet-based HYSPLIT (Draxler and Rolph, 2012; Rolph, 2012). Back trajectories were calculated before the sampling period in January 2009 (winter) and July 2009 (summer). Each back trajectory was run for 13 days and was initiated on 1st, 8th, 15th, 22nd, and 29th day of the month. The atmospheric vertical level of release for each back trajectory was at 500 mAGL. GDS meteorological data were adopted for the analysis.

3. Results

3.1. Spatio-temporal sampling of urban aerosol environments

We assembled an extensive spatio-temporal data set of seasonal-tropical urban bio-aerosols from the city of Hong Kong that involved collecting bio-aerosols weekly over a continuous 12-month period (Nov 2008–Oct 2009) from four locations. Site selection was based upon analysis of key air pollution indicators and the level of urban development in order to cover a statistically supported gradient of urban influence on aerosols from inner-city to a rural location (Fig. 1, Supplementary Tables S1 and S2).

3.1.1. Spatio-temporal variation in aerosolized microbial populations

We used DNA recovered from filters as a general proxy for biomass, with unexposed filters acting as controls. Concentrations of DNA recovered ranged from 0.1 to 6.3 ng l^{-1} air sampled (Supplementary Fig. S1). There was a statistically significant relationship between the amount of DNA collected and sampling months (ANOVA, $F = 7.031$, $P < 0.001$). Generally, aerosol DNA recovered during summer was higher than in winter months. In contrast, there was no significant difference for concentration of DNA between sampling sites along the gradient from urban to rural locations (ANOVA, $F = 0.837$, $P > 0.05$).

Variation in bacterial and eukaryal assemblages among all air samples was examined using terminal restriction fragment length polymorphism (t-RFLP) of 16S rRNA and 18S rRNA genes in a multi-domain assessment (Fig. 2). Initial studies indicated that archaea formed a negligible part of the bio-aerosol as they were below detection limits of amplicon and real-time quantitative PCR (Supplementary Tables S3 and S4), and so they were not included in further study. Major differences between months were apparent for both bacteria (ANOSIM, Global R = 0.586, $P < 0.001$) and eukarya (ANOSIM, Global R = 0.507, $P < 0.001$). The major differences in phylogenetically-defined diversity for both bacteria and eukarya were seasonal, with distinct groupings of relatively closely related assemblages in cooler and drier months (Nov–Apr), giving way to two more variable groups of assemblage types in hotter and wetter months (May–Aug and Sep–Oct) (Fig. 2), Conversely we could establish no significant differences among samples from different sites along the gradient of urbanization, such that assemblages were not significantly different between roadside-urban, urban, sub-urban or rural locations for either bacteria (ANOSIM, Global R = −0.042, $P = 0.954$) or eukarya (ANOSIM, Global R = 0.019, $P = 0.227$).

3.2. Spatio-temporal variation in aerosolized allergens

Bacterial endotoxin and fungal glucan levels varied markedly during the sampling period (endotoxins: 0.35 to 29.21 EU m^{-3}, glucans 0.18–22.3 ng m^{-3}), with mean values broadly tracking the variations in biomass (Supplementary Fig. S2). Neither endotoxin nor glucan levels showed a statistically significant difference between sampling sites (ANOVA, $F < 1.5$, $P > 0.05$ in both cases). Endotoxin levels were generally higher during summer months and correlated positively with recovered microbial biomass levels, but not with any specific taxon.

3.3. Abiotic influences on bio-aerosols

Multiple rank correlations (BEST analysis) of abiotic variables and microbial assemblages (t-RFLP profiles) showed that temperature was the single most important variable in shaping the bacterial assemblage (ρw 0.272) (Supplementary Table S5) while temperature and rainfall were the most parsimonious combination explaining variation in eukaryal assemblages (ρw 0.224). Similarly, temperature and rainfall were also the variables best able to explain biomass (ρw 0.129) and endotoxin levels (ρw 0.273). Other variables associated with the level of urbanization such as respirable suspended particulates and chemical pollutants showed relatively weak correlation (Supplementary Table S5). Species Richness (Shannon's Index) calculated from the t-RFLP data revealed a strong seasonality for both bacteria (ANOVA, $F = 9.1$, $P < 0.05$) and eukarya (ANOVA, $F = 4.325$, $P < 0.05$) (Table 1 and Supplementary Table S6). In contrast, there were no statistically significant differences between sampling sites for either domain along the gradient of urbanization (ANOVA, $F < 1.5$, $P > 0.05$ in both cases).

In an attempt to further understand the relationship between bacterial and eukaryal bio-aerosols as a function of climatic and pollutant variables, we established univariate models using both linear and quadratic fit (Table 1 and Supplementary Table S7).

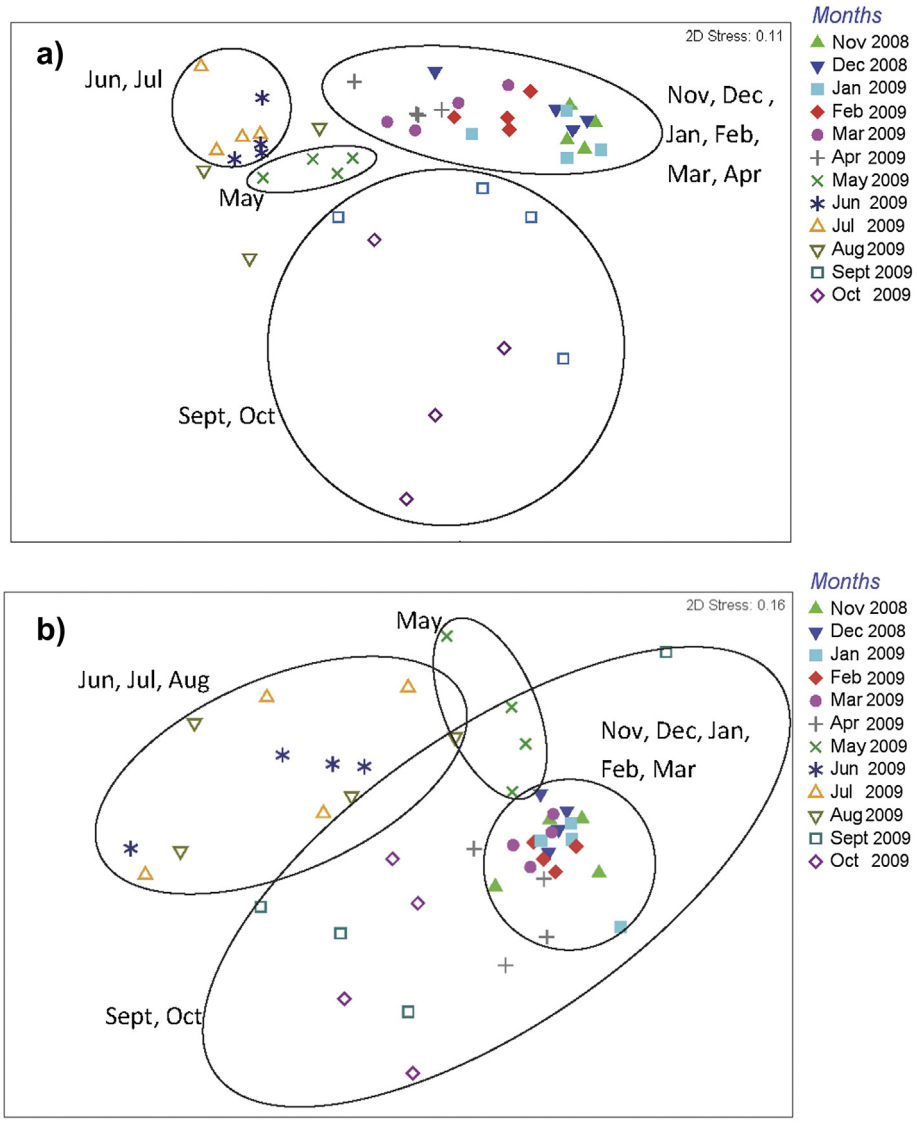

Fig. 2. Nonmetric Multi-Dimensional Scaling plots of Bray Curtis similarity for bacterial and eukaryal assemblages in Hong Kong's urban aerosphere as determined by t-RFLP analysis of rRNA genes. Assemblage shifts were apparent on a temporal scale: (a) Bacteria (ANOSIM, Global R = 0.586, $P < 0.001$) and (b) eukarya (ANOSIM, Global R = 0.507, $P < 0.001$). Each data point represents the mean of four samples from the same site within a month ($N = 48$), for ease of viewing. Ordinations of all samples separately ($N = 192$) produced similar results.

Table 1
Summary of significance levels for effects of climate-related and pollutant abiotic variables on microbial composition of aerosols. Comprehensive data for each test are shown in Supplementary Tables S3, S4, S5 and S7. Abbreviations: P indicates probability for significant correlations; NS, not significant; NT, not testable.

	Univariate models					Mantel test	
	Biomass	Endotoxin	β-Glucan	t-RFLP data		Pyrosequencing data	
				Shannon index H′ (Bacteria)	Shannon index H′ (Eukarya)	UniFrac (Bacteria)	UniFrac (Eukarya)
Climatic variables							
Mean pressure at M.S.L. (hPa)	$P < 0.001$	NT	NS	$P < 0.001$	NS	NS	NS
Temperature (°C)	$P < 0.001$	NS	NS	$P < 0.001$	$P < 0.001$	NS	$P < 0.05$
Relative humidity (%)	NS	$P < 0.01$	$P < 0.05$	$P = 0.001$	NS	NS	NS
Rainfall (mm)	$P < 0.05$	$P < 0.001$	$P < 0.05$	$P < 0.05$	$P < 0.01$	NS	NS
Wind speed	$P < 0.05$	NS	NS	$P < 0.001$	NS	NS	NS
Pollutant variables							
Air Pollution Index	NS	NS	NS	NS	NS	NS	NS
Sulphide dioxide (ug m^{-3})	NS	NS	NS	NS	$P < 0.05$	NS	NS
Respirable suspended particulates (ug m^{-3})	NS	NS	NS	NS	NS	NS	NS
Nitrogen oxides (ug m^{-3})	NS	NS	NS	NS	$P < 0.05$	NS	NS
Nitrogen dioxide (ug m^{-3})	NS	NS	NS	NS	NS	NS	NS
Carbon monoxide (ug m^{-3})	NS	NS	$P < 0.01$	$P < 0.01$	$P = 0.001$	NS	NS
Ozone (ug m^{-3})	NS	NS	NT	NS	NS	NT	NT

Various climatic factors were consistently shown to have a strong statistically significant relationship with biotic variables. In contrast, the relationships with pollutant variables were mostly insignificant and confidence levels were weak where statistical significance was demonstrated.

Backward air trajectory analysis clearly demonstrated different origins of air masses during the sampling period and these were readily differentiated between winter and summer (Fig. 3). The air masses arriving over Hong Kong during winter months were mostly from continental sources under the influence of the continental scale Siberian high-pressure system, and the trajectories suggest that a considerable amount of air mass was transported to Hong Kong from the far northwestern continent. Contrastingly, the air masses coming to Hong Kong in summer months followed a largely marine trajectory under the influence of the Asian Monsoon system.

3.4. High-throughput sequencing of bio-aerosols

To gain greater insight into the phylogenetic diversity of bio-aerosol assemblages, samples were further characterized using high throughput sequencing of rRNA genes. Based on the four statistically supported groupings for the 192 assemblages that associated along temporal differences (see Fig. 2), we selected an urban and rural sample from each statistically supported cluster ($N = 8$) for barcoded pyrosequencing of bacteria and eukarya (Figs. 4 and 5, Supplementary Figs. 5 and 6).

We identified 6394 unique phylotypes from the 129,865 quality trimmed 16S rRNA sequences with an OTU similarity cutoff of 97% (Supplementary Table S8 and Supplementary Fig. S3). These data allowed us to identify that the temporal shift in 16S rRNA-defined assemblages was mainly due to seasonal changes in the abundance of Cyanobacteria and Chlorophyta. The Proteobacteria dominated all of the samples and they accounted for at least 34% of the total bacterial assemblage for any given sample (Fig. 4). Among the Proteobacteria, the class Beta proteobacteria was the most abundant among samples collected in January and May (cooler, drier months) while Gammaproteobacteria were dominant in July and October (warmer, wetter months). Bacteria belonging to the order Burkholderiales accounted for 99% of all Beta proteobacteria sequences. Within the Gammaproteobacteria, these mainly comprised the Pseudomonadales (50%), Chromatiales (34%) and Xanthomonadales (12%). The Chromatiales accounted for at least 28% of total Proteobacteria among samples collected in July and October but only accounted for <0.6% samples in January and May. Alphaproteobacteria were also prevalent in all samples and more than 50% of the bacteria in this class were identified as Rhizobiales.

Cyanobacteria were more abundant for samples collected in January (winter) and they comprised 5.8% (rural) and 8% (urban) of all 16S rRNA gene sequences. The summer samples comprised less than 2% cyanobacteria regardless of location. We also recovered a number of sequences that indicated plastids and phylogenetic analysis (Supplementary Fig. S3) was used to further confirm that these affiliated with Bacillariophyta (diatoms), Chlorarachniophyceae (cercozoans) and Chlorophyta (green algae). Variation among chlorophyte sequences was pronounced, they accounted for 43% of summer phylotypes vs. <1% of winter/spring phylotypes. A relatively low number of phylotypes (<2%) were indicative of propagules from higher plants. These included regionally occurring wild species such as *Castanea mollissima* (Chinese chestnut) and a number of common agricultural plants (castor bean, maize, tobacco).

In assessments of eukaryal rRNA genes, initial pyrosequencing efforts revealed the majority of 18S rRNA gene sequences were fungal (mean 54.2%) with lower frequencies of algae (mean 2%), plants (mean 6.7%), and arthropod (mean 19%) phylotypes (13,044 filtered reads, data not shown). The 18S rRNA PCR primers gave less satisfactory results than ITS primers in pyrosequencing data aquisition, and so we restricted our future efforts to using the eukaryal ITS primers that provide greater resolution and fidelity for fungi and are proven to work well with pyrosequencing platforms (Monchy et al., 2011). We obtained 5031 unique phylotypes from the 81,411 quality trimmed ITS sequences with OTUs defined at 97% sequence similarity (Supplementary Table S8, Supplementary Fig. S3). The Ascomycota and Basidiomycota were prevalent among all samples (Fig. 5). All but one sample comprised >50% Ascomycota phylotypes and among these more than 85% of them were class Dothideomycetes dominated by *Cladosporium* spp. Among the Basidiomycota, over 90% of phylotypes indicated Agaricomycetes. A single rural summer sample was dominated by Basdiomycota phylotypes (>99% of phylotypes) and these were mainly affiliated with the species *Tremiscus helvelloides*, a saprobe of the locally occurring Chinese oil pine that was also common in other samples.

The high-throughput sequencing data supported our temporally extended t-RFLP analysis in that no significant relationship between either bacterial (ANOSIM $R = -0.125$, $P > 0.05$) or eukaryal assemblages (ANOSIM $R = -0.031$, $P > 0.05$) with sampling location occurred (Figs. 4 and 5, Supplementary Table S9). In contrast, there were statistically significant differences between samples collected in different months for both bacterial ($R = 0.542$, $P < 0.01$) and eukaryal ($R = 0.313$, $P < 0.05$) assemblages. We applied the Mantel test to identify any correlation between phylogenetic biodiversity derived from high-throughput sequencing (determined using UniFrac (Lozupone et al., 2007)) and abiotic variables (Supplementary Table S9). This test supported temperature (i.e. a proxy for seasonal climate) as a significant factor affecting diversity, but no support for the effects of pollutant levels. These data therefore support our broader analysis based upon t-RFLP data.

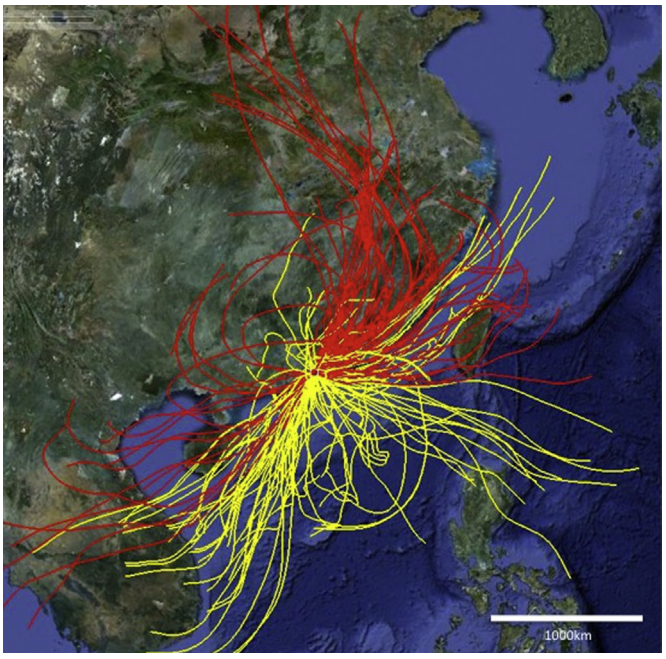

Fig. 3. Back air trajectories arriving over Hong Kong during December 2008 (winter, shown in red) and June 2009 (summer, shown in yellow). Each back trajectory was run for 13 days and was initiated on 1st, 8th, 15th, 22nd, and 29th day of the month to coincide with bio-aerosol sampling. The atmospheric vertical level of release for each back trajectory was at 500 mAGL. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

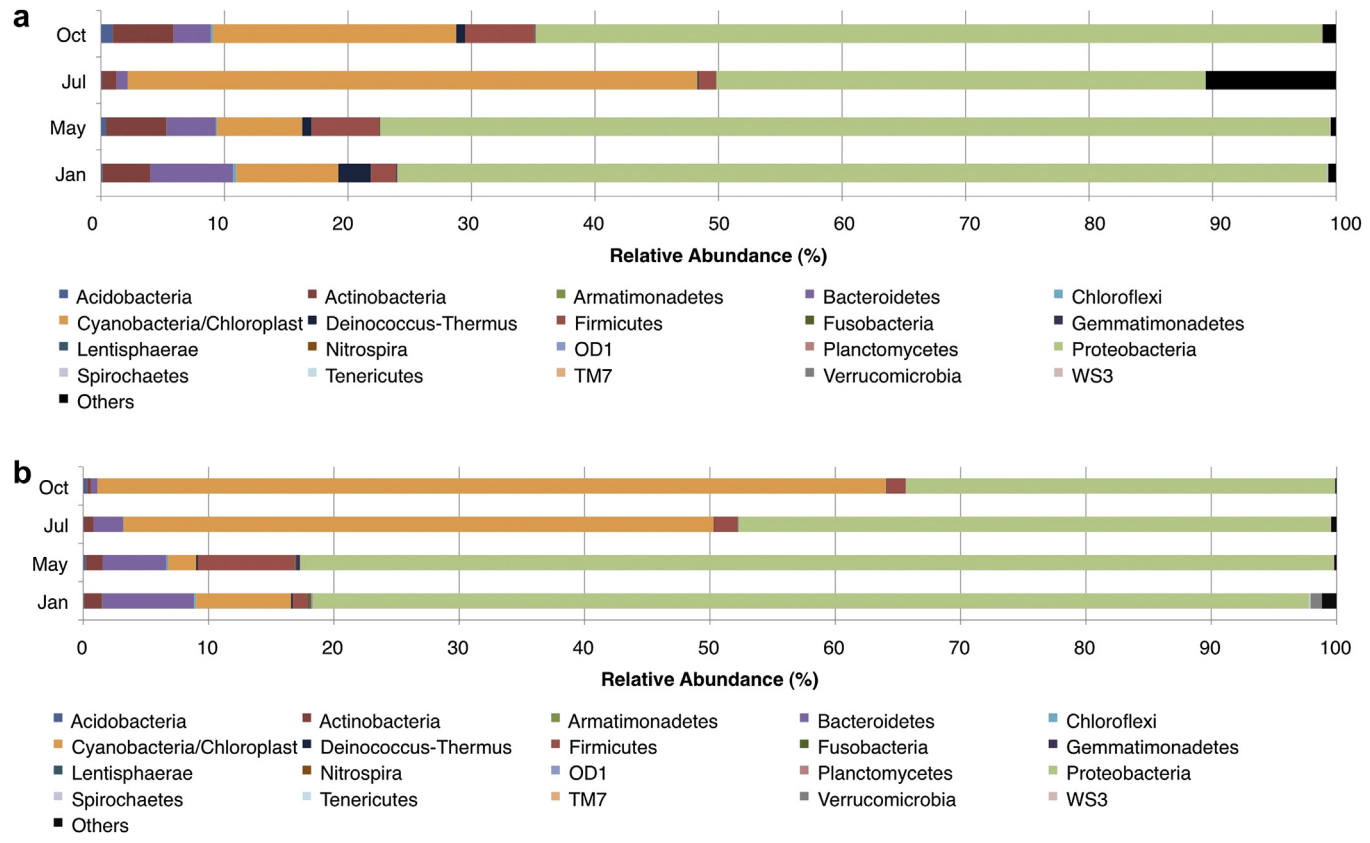

Fig. 4. Relative abundance of bacteria at phylum level in urban (a) and rural (b) bio-aerosols based on high throughput sequencing of 16S rRNA genes ($N = 129{,}865$). Additional analysis at class level is provided in the Supplementary online material.

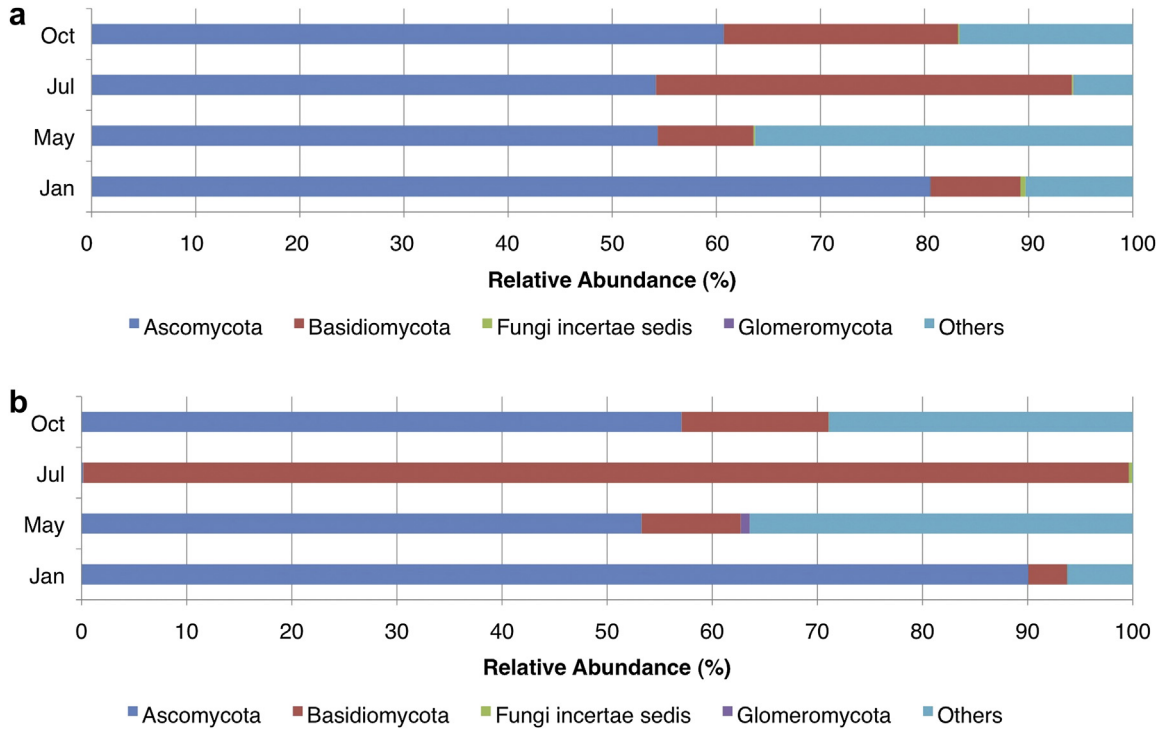

Fig. 5. Relative abundance of fungi at phylum level in urban (a) and rural (b) bio-aerosols based on high throughput sequencing of ITS region ($N = 81{,}411$). Additional analysis at class level is provided in the Supplementary online material.

3.5. Detection of putative pathogens in bio-aerosols

We screened our pyrosequencing data for a range of common bacterial and fungal pathogens, indicators of human presence and also species listed on the CDC bioterror watch list (http://www.bt.cdc.gov/agent/agentlist.asp) (Supplementary Table S10). We employed screening methods for our sequence-based data set that have been developed and applied in several other studies to indicate phylogenetic markers likely due to pathogens (Brodie et al., 2007; Luna et al., 2007; Tringe et al., 2008; Kembel et al., 2012). We detected various putative pathogenic bacterial signatures in all samples at low frequencies, including representatives from the following genera: *Legionella*, *Pseudomonas*, *Shigella*, *Salmonella*, *Staphylococcus* and *Streptococcus* (Supplementary Table S10). Several species listed on the CDC bioterror watch list (http://www.bt.cdc.gov/agent/agentlist.asp) were detected including *Escherichia coli* O157:H7, *Clostridium perfringens* and *Ricinus communis*. Phylotypes indicating bacterial pathogens accounted for 0.48% of all 16S rRNA bacterial gene sequences. Other phylotypes that are common indicators of human presence were also recovered, including *Staphylococcus epidermidis*, *Staphylococcus haemolyticus*, *Staphylococcus saprophyticus*, *Streptococcus salivarius*, *Propionibacterium acnes* and *Pseudomonas aeruginosa*. Among the eukaryal phylotypes, the opportunistic pathogen *Aspergillus terreus* and pathogenic species of the yeast *Candida* were encountered in most samples, but no other possible pathogenic fungi were identified. Putative fungal pathogens accounted for 1.2% of the fungal deep-sequencing data sets.

4. Discussion

This study clearly demonstrates that measurable variation in bacterial and fungal populations plus microbially-derived allergens, occurred seasonally over a continuous annual temporal scale in the urban landscape of Hong Kong. The level of urbanization and human occupation as measured by common indicators of urban pollution did not emerge as a key factor in shaping microbial communities. Rather, the shift in bio-aerosol composition was best explained by seasonal climatic variables. The marked seasonal difference in source of air masses arriving in the territory also emerged as a possible influence on bio-aerosols, and in particular as an explanation for sample variation in deep-sequencing data.

We present strong evidence for greater abundance of marine-derived phylotypes (notably Bacillariophyta and Chlorophyta) in summer when air mass is predominantly oceanic in origin. This is supported by other cultivation/microscopy studies of algal biogeography (Genitsaris et al., 2011). The winter air mass arrived from terrestrial sources and the abundant betaproteobacterial phylotypes that were characteristic of summer samplings indicated soil origins (Bowers et al., 2011b). Whilst we did not have local soil/urban surface samples for reference, it appears that dynamics of Hong Kong bio-aerosols may be most parsimoniously explained by sources of regional air masses. This is in contrast to strong local signals detected from studies of mid-western urban US locations that included specifically identified sources such as dog feces (Bowers et al., 2011a, 2011b). This may reflect differences related to climate and land use between our coastal tropical location and the temperate continental US location, and highlights that a more globally-focused effort to understand bio-aerosols and their biogeography is required in future since a single location is unlikely to provide unifying explanations and theories.

The composition of Hong Kong bacterial bio-aerosols revealed clear trends, but also marked differences compared to studies of other urban aerosols using similar approaches. Indeed taken together this and other recent studies indicate a considerable biogeographic variation probably exists for urban bacterial bio-aerosols. For example our aerosols were dominated by alpha and beta proteobacteria but variation in our samples was due to increases in cyanobacteria in winter and algae in summer. In Italy winter bio-aerosols were dominated by Actinobacteria and in summer by plastids (Franzetti et al., 2011), and in the US mid-west winter bio-aerosols were dominated by Bacteriodetes, Clostridiales and Fusobacteria and in summer by alpha and beta proteobacteria (Bowers et al., 2011b). In contrast fungal diversity at the phylum level in our study was remarkably similar to that from bio-aerosols in Germany (Frohlich-Nowoisky et al., 2009), although our deep sequencing effort revealed significantly greater O.T.U richness at lower taxonomic ranks, with winter assemblages dominated by ascomycetes related to *Cladosporium*, and in summer by basidiomycete Agaricales. This likely reflects fungal basidiospore release during wetter summer months (Frohlich-Nowoisky et al., 2009) and the identification of the basidiomycetes in our rural data set as saprobes of a locally occurring conifer strengthen this. This indicates that in addition to climatic and seasonal factors, our data set also hints at some local influence on fungal bio-aerosols for the rural location, and is supported by findings that agricultural land use in the US mid-west influences bacterial bio-aerosols (Bowers et al., 2011a). It may yet emerge that urban influence leads to greater evenness in bio-aerosols on temporal scales compared to rural locations but this requires further investigation. Interestingly, our time series of aerosol sampling included stochastic weather events (typhoons) and yet these did not significantly impact aerosol biodiversity. It is well documented that dust storms from continental desert sources can led to marked changes in bio-aerosols (Boreson et al., 2004; Jeon et al., 2011; Pointing and Belnap, 2012). Although dust storms can be significant regionally, the time-frame of our study did not include any major dust events in the region.

The detection of a relatively broad range of phylotypes indicative of pathogenic bacteria and fungi was at abundances that suggest these are legitimate signatures of actual populations (rather than singletons in libraries for example). Whilst these phylotypes formed over 1% of recoverable phylotypes, it is not possible from these data to directly correlate this with cell numbers and hence risk of infection. A recent study identified high-throughput sequencing libraries for indoor bio-aerosols supported greater abundance of pathogenic phylotypes than an outdoor sample (Kembel et al., 2012). This might suggest that built environments and human occupancy could be a reservoir for pathogenic taxa and influence recovery of these taxa in urban outdoor air, although our study did not reveal any trend along the urbanization gradient in Hong Kong to support this. Long distance transport of pathogenic microbial taxa in aerosols has also been demonstrated (Griffin, 2007) and so an additional non-local input may also occur. Establishing local versus non-local pathogen sources may be an interesting avenue for future research.

Microbially derived allergen levels in our study were higher than several other studies on urban outdoor environments in the USA (Mueller-Anneling et al., 2004; Solomon et al., 2006). Thresholds for endotoxins are not standardized internationally although a variety of recommendations do exist. A US study indicates exposure to a threshold ambient concentration of 17 EU m^{-3} for 24 h may result in acute airway obstruction (Mueller-Anneling et al., 2004). In our study, there were occasions mainly during summer when endotoxin levels exceeded the threshold concentration by almost 2-fold. The relationship between microbially derived allergens, symptoms, urban bio-aerosols and the urban environment require further study to identify causality.

5. Conclusions

We conclude by postulating that the level of urbanization and population growth are unlikely to impact regional airborne microbial populations, but rather climate as determined by regional air mass trajectories determine Hong Kong's urban bio-aerosols in a predictive manner. This implies limited negative impact of urbanization on natural microbial dispersal in this region, and no evidence for major increases in human associated microflora in urban aerosols along the urbanization gradient. Whether this holds for urban locations in temperate versus tropical, and coastal versus inland cities requires further investigation. We also acknowledge that a confounding factor may arise in our study since the Pearl River Delta region where Hong Kong is located is widely and heavily polluted and this could potentially have masked more subtle localized urban effects. In addition to air-related variables and regional land use, the nature of urban planning and even cultural habits (e.g. pet ownership (Bowers et al., 2011b)) may also emerge as important factors when making comparisons to other cities and regions. Our findings also suggest that broader issues in microbial ecology may benefit from incorporating bio-aerosol studies, for example the popularized notion of distance–decay relationships for biogeography of microbial taxa (Hanson et al., 2012) may be affected in part by regional air trajectories. Finally we identify that greater focus on risk assessment to human health of outdoor urban bio-aerosols may be warranted given the findings of our and other recent studies.

Acknowledgments

The authors acknowledge the co-operation of the Hong Kong SAR Meteorological Office and Environmental Protection Department.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.atmosenv.2013.03.047.

References

Abdo, Z., Schuette, U.M., Bent, S.J., Williams, C.J., Forney, L.J., Joyce, P., 2006. Statistical methods for characterizing diversity of microbial communities by analysis of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environmental Microbiology 8 (5), 929–938.

Baker, G., Smith, J.J., Cowan, D.A., 2003. Review and re-analysis of domain-specific 16S primers. Journal of Microbiological Methods 55, 541–555.

Boreson, J., Dillner, A.M., Peccia, J., 2004. Correlating bioaerosol load with PM2.5 and PM10cf concentrations: a comparison between natural desert and urban-fringe aerosols. Atmospheric Environment 38 (35), 6029–6041.

Bowers, R.M., McLetchie, S., Knight, R., Fierer, N., 2011a. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. ISME Journal 5 (4), 601–612.

Bowers, R.M., Sullivan, A.P., Costello, E.K., Collett Jr., J.L., Knight, R., Fierer, N., 2011b. Sources of bacteria in outdoor air across cities in the midwestern United States. Applied and Environmental Microbiology 77 (18), 6350–6356.

Brodie, E.L., DeSantis, T.Z., Parker, J.P., Zubietta, I.X., Piceno, Y.M., Andersen, G.L., 2007. Urban aerosols harbor diverse and dynamic bacterial populations. Proceedings of National Academy of Sciences of the United States of America 104 (1), 299–304.

Burrows, S.M., Elbert, W., Lawrence, M.G., Pöschl, U., 2009. Bacteria in the global atmosphere – part 1: review and synthesis of literature data for different ecosystems. Atmospheric Chemistry and Physics Discussion 9 (3), 9263–9280.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25 (15), 1972–1973.

Draxler, R.R., Rolph, G.D., 2012. HYSPLIT (HYbrid Single-particle Lagrangian Integrated Trajectory) Model Access via NOAA ARL READY Website. NOAA Air Resources Laboratory, Silver Spring, MD. http://ready.arl.noaa.gov/HYSPLIT.php (accessed 30.06.12.).

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32 (5), 1792–1797.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26 (19), 2460–2461.

Edgar, R.C., Haas, B.J., Clemente, J.C., Quince, C., Knight, R., 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27 (16), 2194–2200.

Franzetti, A., Gandolfi, I., Gaspari, E., Ambrosini, R., Bestetti, G., 2011. Seasonal variability of bacteria in fine and coarse urban air particulate matter. Applied Microbiology & Biotechnology 90 (2), 745–753.

Frohlich-Nowoisky, J., Pickersgill, D.A., Despres, V.R., Poschl, U., 2009. High diversity of fungi in air particulate matter. Proceedings of National Academy of Sciences of the United States of America 106 (31), 12814–12819.

Genitsaris, S., Kormas, K.A., Moustaka-Gouni, M., 2011. Airborne algae and cyanobacteria: occurrence and related health effects. Frontier in Bioscience (Elite Ed) 3, 772–787.

Griffin, D.W., 2007. Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. Clinical Microbiology Reviews 20 (3), 459–477.

Hanski, I., von Hertzen, L., Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., et al., 2012. Environmental biodiversity, human microbiota, and allergy are interrelated. Proceedings of National Academy of Sciences of the United States of America 109 (21), 8334–8339.

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B., 2012. Beyond biogeographic patterns: processes shaping the microbial landscape. Nature Reviews Microbiology 10, 497–506.

Horner, W.E., Helbling, A., Salvaggio, J.E., Lehrer, S.B., 1995. Fungal allergens. Clinical Microbiology Reviews 8 (2), 161–179.

Jaenicke, R., 2005. Abundance of cellular material and proteins in the atmosphere. Science 308 (5718), 73.

Jeon, E.M., Kim, H.J., Jung, K., Kim, J.H., Kim, M.Y., Kim, Y.P., Ka, J.-O., 2011. Impact of Asian dust events on airborne bacterial community assessed by molecular analyses. Atmospheric Environment 45 (25), 4313–4321.

Katoh, K., Asimenos, G., Toh, H., 2009. Multiple alignment of DNA sequences with MAFFT. Methods in Molecular Biology 537, 39–64.

Kembel, S.W., Jones, E., Kline, J., Northcutt, D., Stenson, J., Womack, A.M., et al., 2012. Architectural design influences the diversity and structure of the built environment microbiome. ISME Journal 6 (8), 1469–1479.

Leung, R., 1993. Prevalence of allergy and atopy in Hong Kong. Journal of Hong Kong Medical Association 45, 232–238.

Leung, R., Wong, G., Lau, J., Ho, A., Chan, J.K., Choy, D., et al., 1997. Prevalence of asthma and allergy in Hong Kong schoolchildren: an ISAAC study. European Respiratory Journal 10 (2), 354–360.

Li, Y., Leung, G.M., Tang, J.W., Yang, X., Chao, C.Y., Lin, J.Z., et al., 2007. Role of ventilation in airborne transmission of infectious agents in the built environment – a multidisciplinary systematic review. Indoor Air 17 (1), 2–18.

Lozupone, C.A., Hamady, M., Kelley, S.T., Knight, R., 2007. Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Applied and Environmental Microbiology 73 (5), 1576–1585.

Luna, R.A., Fasciano, L.R., Jones, S.C., Boyanton Jr., B.L., Ton, T.T., Versalovic, J., 2007. DNA pyrosequencing-based bacterial pathogen identification in a pediatric hospital setting. Journal of Clinical Microbiology 45 (9), 2985–2992.

Mancinelli, R.L., Shulls, W.A., 1978. Airborne bacteria in an urban environment. Applied and Environmental Microbiology 35, 1095–1101.

Maron, P.A., Mougel, C., Lejon, D.P.H., Carvalho, E., Bizet, K., Marck, G., et al., 2006. Temporal variability of airborne bacterial community structure in an urban area. Atmospheric Environment 40 (40), 8074–8080.

Monchy, S., Sanciu, G., Jobard, M., Rasconi, S., Gerphagnon, M., Chabe, M., et al., 2011. Exploring and quantifying fungal diversity in freshwater lake ecosystems using rDNA cloning/sequencing and SSU tag pyrosequencing. Environmental Microbiology 13 (6), 1433–1453.

Mueller-Anneling, L., Avol, E., Peters, J.M., Thorne, P.S., 2004. Ambient endotoxin concentrations in PM10 from Southern California. Environmental Health Perspectives 112 (5), 583–588.

Muyzer, G., de Waal, E.C., Uitterlinden, A.G., 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Applied Environment Microbiology 59 (3), 695–700.

Nercessian, O., Reysenbach, A.-L., Prieur, D., Jeanthon, C., 2003. Archaeal diversity associated with in situ samplers deployed on hydrothermal vents on the East Pacific Rise (13°N). Environmental Microbiology 5, 492–502.

Peccia, J., Milton, D.K., Reponen, T., Hill, J., 2008. A role for environmental engineering and science in preventing bioaerosol-related disease. Environmental Science and Technology 42 (13), 4631–4637.

Pointing, S.B., Belnap, J., 2012. Microbial colonization and controls in dryland systems. Nature Reviews Microbiology 10 (8), 551–562.

Pöschl, U., 2003. Aerosol particle analysis: challenges and progress. Analytical Bioanalytical Chemistry 375, 30–32.

Price, M.N., Dehal, P.S., Arkin, A.P., 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Molecular Biology and Evolution 26 (7), 1641–1650.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B.M., Ludwig, W., Peplies, J., Glockner, F.O., 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Research 35 (21), 7188–7196.

Rintala, H., Pitkaranta, M., Toivola, M., Paulin, L., Nevalainen, A., 2008. Diversity and seasonal dynamics of bacterial community in indoor environment. BMC Microbiology 8, 56.

Rolph, G.D., 2012. Real-time Environmental Applications and Display sYstem (READY) Website. NOAA Air Resources Laboratory, Silver Spring, MD. http://ready.arl.noaa.gov (accessed 30.06.12.).

Schloss, P.D., Handelsman, J., 2005. Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. Applied Environmental Microbiology 71 (3), 1501–1506.

Solomon, G.M., Hjelmroos-Koski, M., Rotkin-Ellman, M., Hammond, S.K., 2006. Airborne mold and endotoxin concentrations in New Orleans, Louisiana, after flooding, October through November 2005. Environmental Health Perspectives 114 (9), 1381–1386.

Tringe, S.G., Zhang, T., Liu, X.G., Yu, Y.T., Lee, W.H., Yap, J., et al., 2008. The airborne metagenome in an indoor urban environment. PLoS One 3 (4).

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73 (16), 5261–5267.

White, T.J., Burns, T., Lee, S., Taylor, J.W., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR Protocols: a Guide to Methods and Applications. Academic, New York, pp. 315–322.