Sources and sinks of the air microbiome

Lau, Kenny Jia Xu

2021

Lau, K. J. X. (2021). Sources and sinks of the air microbiome. Doctoral thesis, Nanyang Technological University, Singapore. <https://hdl.handle.net/10356/153351>

<https://hdl.handle.net/10356/153351>

<https://doi.org/10.32657/10356/153351>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 22:39:37 SGT

**NANYANG
TECHNOLOGICAL
UNIVERSITY**

SINGAPORE

SOURCES AND SINKS OF THE AIR MICROBIOME

KENNY LAU JIA XU
Interdisciplinary Graduate School
Singapore Centre for Environmental Life Sciences Engineering (SCELSE)

SOURCES AND SINKS OF THE AIR MICROBIOME

KENNY LAU JIA XU

**Interdisciplinary Graduate School
Singapore Centre for Environmental Life Sciences Engineering
(SCELSE)**

A thesis submitted to the Nanyang Technological University in partial
fulfillment of the requirement for the degree of
Doctor of Philosophy

2021

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

24 December 2020

.....

Date

.....

Kenny Lau Jia Xu

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

24 December 2020

.....

Date

.....

Prof. Stephan C. Schuster

Authorship Attribution Statement

This thesis contains material from 2 paper(s) published in the following peer-reviewed journal(s) where I was the first and/or corresponding author.

Chapter 4 is published as Micheál Mac Aogáin*, Kenny J. X. Lau*, Cai Zhao, Jayanth Kumar Narayana, Rikky W. Purbojati, Daniela I. Drautz-Moses, Nicolas E. Gaultier, Tavleen K. Jaggi, Pei Yee Tiew, Thun How Ong, Mariko Siyue Koh, Albert Lim Yick Hou, John A. Abisheganaden, Krasimira Tsaneva-Atanasova, Stephan C. Schuster and Sanjay H. Chotirmall. Functional metagenomics reveals a core macrolide resistome related to microbiota in chronic respiratory disease.

The contributions of the co-authors are as follows:

- This is a joint first author paper between Dr. Micheal Mac Aogain and me where we designed the experiments, performed statistical analyses, and wrote the manuscript
- Dr. Cai Zhao and Jayanth performed antimicrobial resistance gene analysis by further constructing a co-occurrence network graph
- Dr. Tuang Yeow, Dr. Thun How and Dr. Albert were involved in the patient recruitment and procurement of clinical data and specimens
- Rikky performed bioinformatic analysis
- Dr. Daniela sequenced the sputum and inhaler swab samples
- Dr. John and Dr. Mariko provided intellectual contributions, patient recruitment
- Prof. Krasimira oversaw the mathematical methods and statistics
- Prof. Stephan and Prof. Sanjay conceived the design of the overall study and experiments from data analysis, interpretation, statistics, manuscript, and procurement of funding

Chapter 5 is published as Kenny J. X. Lau, Irvan Luhung, Balakrishnan N.V. Premkrishnan, Vineeth Kodengil Vettath, Anthony Wong, Akira Uchida, Tuang Yeow Poh, Daniela I. Drautz-Moses, Ana Carolina M. Junqueira, and Stephan C. Schuster. Structure vs. Chemistry: Alternate Mechanisms for Controlling Leaf Microbiome. (Under Review)

The contributions of the co-authors are as follows:

- I swabbed the leaves, extracted DNA, performed both reactive oxygen species assay and surface wettability tests, statistical analyses and wrote the manuscript
- Premkrishnan performed bioinformatic processing of the sequenced data
- Anthony performed sample preparation and coordinated efforts for electron microscopy imaging at A*Star
- Dr. Tuang Yeow guided and assisted me in the reactive oxygen species assay
- Dr. Akira helped me with the improvisation of the swab technique
- Dr. Daniela sequenced the swab samples with the Illumina HiSeq 2500
- Prof. Stephan and Prof. Ana supervised me throughout this work from experimental design, data analyses to writing

24 December 2020

.....

Date

.....

Prof. Stephan C. Schuster

Acknowledgement

My Ph.D. journey took many unexpected turns, often hovering between states of frustration and eureka moments. Firstly, I thank my main supervisor, Prof. Stephan Schuster, for giving me freedom to explore and seek creative solutions to answer hypotheses in my experiments. Secondly, I thank my co-supervisor, Prof. Sanjay Chotirmall, who also closely monitor my progress throughout my Ph.D. I am also actively participating in his group meetings and the discussions also help to cross-pollinate ideas. I am grateful to both my supervisors who have equipped me with essential skills of a researcher, giving me opportunities to present at conferences and contribute to various types of publications, from genome announcement, review to research articles.

This research work would not have been possible if I have not met Dr. Enzo Acerbi, who first introduced me to the SCELSE Air Microbiome team in 2016. He was my supervisor during my final year project as an undergraduate student. It was the field work and the steep learning curve in metagenomics that enticed me to pursue a Ph.D. study in this field. I also wish to express gratitude to the following research fellows who made my Ph.D. journey less daunting. Firstly, Dr. Irvan Luhung who led our team in the sampling of a moldy office, where the results became an excellent case study in Chapter 3 of my thesis. He guided me in the experimental design and methods of quantitative PCR on samples collected in the built environment. Secondly, Dr. Tuang Yeow who showed me how to reproduce and modify protocols from existing literature using the reagent and equipment available. Dr. Micheál Mac Aogáin who taught me the importance of independent learning and validation of data in bioinformatics.

Abstract

Background

Microbial source tracking is necessary to study, model and predict the movement of airborne microorganisms, especially in the context of respiratory diseases, agricultural pathogens, food spoilage and mould growth in buildings. The indoor built environment, human lung, and plant surfaces were investigated as potential sources of the air microbiomes. The findings in this thesis are expected to help identify the sources and map the fluxes of bioaerosols between the built environment, humans, and plants.

Results

Source tracking was conducted at two indoor offices with poor air quality. Metagenomics was used to characterise both the indoor air and surface microbiome. Bayesian inferential analysis was performed to quantify probabilities of transfer between the air and surfaces. Computational fluid dynamics modelling was also used to examine airflow patterns within the indoor space. Results from the metagenomic, probabilistic and airflow models reveal that cellulosic, gypsum and textile surfaces are likely reservoirs of indoor air microbes in moisture-damaged buildings.

Secondly, to study the microbial exchanges between the human lung and environment, source tracking was also performed on the sputum, inhalers of respiratory patients, including the air and surfaces of their residential homes. Results suggests that there was a significant amount of microbial transfer between the lung and inhaler. Inhaler surfaces harbour antibiotic resistance genes such as macrolide, beta-lactam, and multi-drug resistance genes via contact transfer with the human oral cavity. Given the similarity between the lung and inhaler microbiomes, a swab sample of the inhaler may be a good surrogate of the lung microbiome. Further findings suggest that the aircon, air purifiers and fans in homes are potential microbial sources exhibiting a

modest transfer of 5-10% to the lung microbiome, whereas the inhaler shows a transfer probability of more than 80% in microbial exchanges with the lung.

Lastly, plants have large surface areas of leaves that may be potential sources to the air microbiome. Source tracking investigation was performed between the leaves of outdoor plants and outdoor air. Results suggest that leaves contribute to less than 5% of the outdoor air in Singapore. Further experiments revealed that the adaxial and abaxial of the leaf surfaces harbour different microbiomes. Metagenomics profiling showed that the abaxial leaf surfaces have lesser bacteria than the adaxial counterpart. This could potentially be explained by the microbe-plant interactions where plants could deploy defence strategies to guard their stomata from phytopathogens. The leaves of two species of garden plants were sampled. *Rhapis excelsa* showed increased in reactive oxygen species production while *Cordyline fruticosa* has a hydrophobic surface nanostructure that may explain the reduction in bacteria observed on the abaxial surface.

Conclusions

In summary, the study of sources and sinks of the air microbiome requires a combinatorial approach of metagenomics, Bayesian probability, Markov Chain and computational fluid dynamics modelling. Results from the combined analyses can better inform indoor occupants of the sources of contamination so that relevant disinfection actions can be taken. Findings in this thesis also showed that inhalers are potential reservoirs for antibiotic-resistant microbes. It was also found that the fan blades, aircon panels and air purifiers in residential homes are potential environmental sources of microbes in indoor air. Lastly, the leaf surfaces of outdoor plants only contributed to less than 5% the outdoor air, which is likely due to the low wind speeds in Singapore unlike studies conducted in temperate and coastal sites. This work can provide a basis for modelling of bioaerosols transport and their impact on public health.

“Study hard what interests you most in the most undisciplined, irreverent and original manner possible” – Richard Feynman

Contents Page

Acknowledgements	v
Abstract	vi
List of Figures, Tables, Abbreviations and Publications	xii

Chapter 1. Literature Review

1.1 Introduction	1
1.2 Challenges of air sampling	1
1.3 Next-generation sequencing and microbiome approaches	3
1.4 Sources of bioaerosols	4
Sources in the built environment	4
Humans as sources of the air microbiome	7
Plants as sources of the air microbiome	8
1.5 Source identification methods	10
Algorithm of the source tracker method software	12
1.6 Challenges in modelling the diffusion and transport of airborne microbes	15
Algorithm behind computational fluid dynamics simulation	17
1.7 Research gaps and aims of study	21

Chapter 2. Source tracking of the air microbiome

2.1 Introduction	23
2.2 Materials and methods	
Swab and air sampling methods	27
Source tracker and modifications used	29

Computational fluid dynamics modelling -----	30
2.3 Results	
Metagenomic and source tracking results of office A -----	30
Metagenomic and source tracking results of office B -----	34
2.4 Discussion -----	37
Explanations on the modifications made to source tracker software -----	43
Comparing results between the modified model and original source tracker -----	46
2.5 Conclusions -----	48

Chapter 3. Humans as sources of the air microbiome

3.1 Introduction -----	50
3.2 Materials and methods	
Sample collection -----	51
Taxonomic and functional assignment of metagenomic sequence reads -----	52
Statistical analyses and data visualization -----	53
Source tracker analysis and modifications used -----	54
3.3 Results	
Functional metabolic dysbiosis of the airway metagenome -----	55
The core airway resistome -----	57
The microbial ecology of the core airway resistome -----	60
Inhaler device as environmental antibiotic resistance reservoirs -----	63
3.4 Discussion -----	67
3.5 Conclusions -----	69

Chapter 4. Plants as sources of the air microbiome

4.1 Introduction	71
4.2 Materials and methods	
Air sample collection	72
Leaf swab sample collection	72
DNA extraction, library preparation and sequencing	74
Taxonomical identification of plants	75
Adaxial and abaxial leaf swabs	76
Bioinformatics and taxonomic assignment	77
Source tracker and modifications used	77
Statistical analyses	78
Processing of leaf samples for scanning electron microscopy (SEM)	78
Reactive oxygen species (ROS) assay	78
Surface water contact angle measurement	79
4.3 Results	
Plants leaves contributes to < 5% of the tropical outdoor air microbiome	80
Significant reduction in bacterial species observed on abaxial leaf surfaces	83
Co-occurrence network reveals nine clusters of microbial species	86
Leaf surface chemical and physical properties	91
4.4 Discussion	94
4.5 Conclusions	96

Chapter 5. Conclusions and future perspectives

5.1 Conclusions and future perspectives	98
References	100

List of figures

Fig 1.1 Potential sources of microbes in the built environment.

Fig 1.2 Conditional probability of balls in the sink urn given two sources.

Fig 1.3 Conditional probability of balls given two sources, one unknown and one sink.

Fig 1.4 Random assignment of each sink sequence was tagged to a source and combined probability scores were calculated in an iterative process.

Fig 1.5 Schematic model of mass conservation in each cell of the CFD mesh.

Fig 1.6 Normal stresses at the boundary planes of each cell.

Fig 1.7 Tangential stresses at the boundary planes of each cell.

Fig 1.8 Solving for u and v vectors for particle movement from one cell to another.

Fig 2.1 A three-way combinatorial approach for source tracking.

Fig 2.2 NIOSH hierarchy of controls in the context of controlling microbial source hazards.

Fig 2.3 Sampling locations in office A.

Fig 2.4 Sampling locations in office B.

Fig 2.5 Metagenomic profiles of air and swab samples collected in office A.

Fig 2.6 Fungi cultivated from ceiling, chair, and floor in office A.

Fig 2.7 Markov chain transition graph of microbial transfers between sites in office A.

Fig 2.8 CFD simulation of airflow in office A.

Fig 2.9 Metagenomic profiles recovered from air and swab samples collected in office B.

Fig 2.10 Markov chain transition graph of microbial transfer between sites in office B.

Fig 2.11 CFD simulation of airflow in office B.

Fig 2.12 Reframing the Bayesian hypothesis underlying the sourcetracker2 algorithm.

Fig 2.13 Forward and inverse problem modelling using Bayes' Theorem.

Fig 2.14 Results from the original source tracker software without modifications.

Fig 3.1 Dysbiosis of functional airway metagenomes in COPD and BE patients.

Fig 3.2 Core airway resistome across respiratory disease states and non-diseased (healthy) individuals.

Fig 3.3 COPD and bronchiectasis patients have the greatest heterogeneity in their microbiomes.

Fig 3.4 α -diversity analysis of the microbiomes across cohorts.

Fig 3.5 Co-occurrence network graph of gene-microbe associations of the core macrolide resistome.

Fig 3.6 Overlap of microbes between airway and inhaler metagenomes.

Fig 3.7 Overlap of antibiotic resistance genes between airway and inhaler metagenomes.

Fig 3.8 Relative abundance of microbial species found on the indoor surfaces, in air and human sputum samples.

Fig 3.9 Source tracking and identification of potential sources of microbes to the lung microbiome.

Fig 4.1 Flowchart of processing pipeline.

Fig 4.2 Swab sampling locations at the N1.3 balcony.

Fig 4.3 Map of leaf swab sampling locations.

Fig 4.4 Relative contributions of nearby vegetation to the air microbiome

Fig 4.5 Biotic and abiotic surface microbiomes at the outdoor balcony versus air.

Fig 4.6 Source tracking investigation performed at the N1.3 balcony.

Fig 4.7 SEM images and bubble plot of microbes, plants, and arthropod-related reads.

Fig 4.8 Dot plot of species richness expressed as Chao-1 index on the leaf surface.

Fig 4.9 Co-occurrence network analysis of microbes on the adaxial and abaxial leaf surfaces.

Fig 4.10 Relative abundance of cluster A species.

Fig 4.11 Relative abundance of cluster B species.

Fig 4.12 Read counts of microbes from each co-occurrence cluster.

Fig 4.13 Reactive oxygen species detection on the leaf surface.

Fig 4.14 Levels of reactive oxygen species and structural differences of the leaf surface.

Fig 4.15 Surface wettability of leaf surfaces.

Fig 4.16 SEM image of the nanoimprinted abaxial leaf surface at 250x magnification.

Fig 4.17 SEM image of the nanoimprinted abaxial leaf surface at 1000x magnification.

List of tables

Table 1.1 Comparison of different types of air samplers.

Table 1.2 Comparison of the source identification methods.

Table 2.1 List of fungal species that are commonly found in moisture-damaged buildings and their associated growth on different types of building materials.

Table 2.2 The relationship between equilibrium relative humidity (ERH) and building materials.

Table 2.3 Recommended Air Changes per Hour (ACH) for different room types.

Table 3.1 Demographics of chronic respiratory patients in the study.

Table 4.1 Mann Whitney test of Chao-1 index between adaxial and abaxial leaf surface.

List of abbreviations

ACH: Air Changes per Hour

ANOVA: Analysis of Variance

ARGs: Antibiotic Resistance Genes

ASHRAE: American Society for Heating, Refrigerating and Air-Conditioning Engineers

ATP: Adenosine Triphosphate

BE: Bronchiectasis

CAM: Correlation Analysis Method

CARD: Comprehensive Antibiotic Resistance Database

CDC: Centre for Disease Control and Prevention

CFD: Computational Fluid Dynamics

COPD: Chronic Obstructive Pulmonary Disease

DAM: Database Analysis Method

DNA: Deoxyribonucleic Acid

EPA: European Protection Agency

ERH: Equilibrium Relative Humidity

ETI: Effector Triggered Immunity

GBLMs: Generalised Boosted Linear Models

GOLD: Global Initiative for Obstructive Lung Disease

HVAC: Heating Ventilation Air Conditioning

HYSPLIT: Hybrid Single Particle Lagrange Integrated Trajectory model

KEGG: Kyoto Encyclopaedia of Genes and Genomes

LCA: Lowest Common Ancestor

MAM: Model Analysis Method

MAP: Mitogen-Activated Protein

MCMC: Monte Carlo Markov Chain

MDB: Moisture Damaged Building

MEA-Strep: Malt-Extract Agar with Streptomycin

MEGAN: Metagenome Analyser

MERV: Minimum Efficiency Reporting Value

MST: Microbial Source Tracking

NCBI: National Center for Biotechnology Information

NIOSH: National Institute for Occupational Safety and Health

nMDS: non-Metric Dimensional Scaling

NTU: Nanyang Technological University

PAMPs: Pathogen Associated Molecular Patterns

PCA: Principal Component Analysis

PCoA: Principal Coordinate Analysis

PCR: Polymerase Chain Reaction

PM: Particulate Matter

PPE: Personal Protective Equipment

PTI: PAMPs Triggered Immunity

rRNA: Ribosomal Ribonucleic Acid

ROS: Reactive Oxygen Species

RPKM: Reads per kilobase mapped reads

SA: Severe Asthma

SARS-CoV2: Severe Acute Respiratory Syndrome Coronavirus 2

SASS: Smart Air Sampling System

SBS: Sick Building Syndrome

SCELSE: Singapore Centre for Environmental Life Sciences and Engineering

SEM: Scanning Electron Microscope

SGH: Singapore General Hospital

TTSH: Tan Tock Seng Hospital

UV: Ultraviolet

VOCs: Volatile Organic Compounds

List of publications

Elena S. Gusareva*, Enzo Acerbi*, Kenny J. X. Lau, Irvan Luhung, Balakrishnan N. V. Premkrishnan, Sandra Kolundžija, Rikky W. Purbojati, Anthony Wong, James N. I. Houghton, Dana Miller, Nicolas E. Gaultier, Cassie E. Heinle, Megan E. Clare, Vineeth Kodengil Vettath, Carmon Kee, Serene B. Y. Lim, Caroline Chénard, Wen Jia Phung, Kavita K. Kushwaha, Ang Poh Nee, Alexander Putra, Deepa Panicker, Koh Yanqing, Yap Zhei Hwee, Sachin R. Lohar, Mikinori Kuwata, Hie Lim Kim, Liang Yang, Akira Uchida, Daniela I. Drautz-Moses, Ana Carolina M. Junqueira, and Stephan C. Schuster (2019). Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proceedings Of The National Academy Of Sciences*, 116(46), 23299-23308. doi: 10.1073/pnas.1908493116

Kenny J. X. Lau*, Ana Carolina M. Junqueira*, Akira Uchida, Rikky W. Purbojati, James N. I. Houghton, Caroline Chénard, Anthony Wong, Sandra Kolundžija, Megan E. Clare, Kavita K. Kushwaha, Alexander Putra, Nicolas E. Gaultier, Cassie E. Heinle, Balakrishnan N. V. Premkrishnan, Vineeth Kodengil Vettah, Daniela I. Drautz-Moses, Stephan C. Schuster (2019). Complete Genome Sequence of *Agrococcus* sp. Strain SGAir0287, Isolated from Tropical Air Collected in Singapore. *Microbiology Resource Announcements*, 8(32). doi: 10.1128/mra.00616-19

Vineeth Kodengil Vettath*, Ana Carolina M. Junqueira*, Akira Uchida, Rikky W. Purbojati, James N. I. Houghton, Caroline Chénard, Daniela I. Drautz-Moses, Anthony Wong, Sandra Kolundžija, Megan E. Clare, Kenny J. X. Lau, Nicolas E. Gaultier, Cassie E. Heinle, Balakrishnan N. V. Premkrishnan, Elena S. Gusareva, Enzo Acerbi, Liang Yang, and Stephan C. Schuster (2017). Complete Genome Sequence of *Bacillus altitudinis* Type Strain SGAir0031 Isolated from Tropical Air Collected in Singapore. *Genome Announcements*, 5(45). doi: 10.1128/genomea.01260-17

Pei Yee Tiew, Fanny Wai San Ko, Sze Lei Pang, Sri Anusha Matta, Yang Yie Sio, Kenny J.X. Lau, Micheal Mac Aogain, Tavleen Kaur Jahhi, Fransiskus Xaverius Ivan, Nicolas E. Gaultier, Akira Uchida, Daniela I. Drautz-Moses, Huiying Xu, Mariko Siyue Koh, David Shu Cheong Hui, Augustine Tee, John Arputhan Abisheganaden, Stephan C. Schuster, Fook Tim Chew and Sanjay H. Chotirmall (2020). Environmental fungal sensitisation associates with poorer clinical outcomes in COPD. *Eur Respir J*, 56(2). doi:10.1183/13993003.00418-2020

Micheal Mac Aogain*, Kenny J.X. Lau*, Zhao Cai, Jayanth Kumar Narayana, Rikky W. Purbojati, Daniela I. Drautz-Moses, Nicolas E. Gaultier, Tavleen K. Jaggi, Pei Yee Tiew, Thun How Ong, Mariko Siyue Koh, Albert Lim Yick Hou, John A. Abisheganaden, Krasimira Tsavena-Atanasova, Stephan C. Schuster and Sanjay H. Chotirmall (2020). Metagenomics Reveals a Core Macrolide Resistome Related to Microbiota in Chronic Respiratory Disease. *American Journal Of Respiratory And Critical Care Medicine*, 202(3), 433-447. doi: 10.1164/rccm.201911-2202oc

* joint-first authorship.

Chapter 1. Literature Review

1.1 Introduction

The biological fraction of particulate matter (PM) in air, collectively known as bioaerosols, consist of bacteria, fungi, viruses, pollen and other biological materials¹. PM10 refers to inhalable coarse particles with a diameter of 2.5 to 10 μm in diameter while PM2.5 refers to fine particles that are smaller than 2.5 μm in diameter². Bioaerosols in PM2.5 consists of viruses in the nanometre size range, bacteria in the 1 μm size range and fungal spores or hyphal fragments that are typically in the 1-3 μm size range³. Bioaerosols can be dispersed either as a single unit or as larger aggregates in air. Smaller particles tend to drift with air currents and are transported across long distances by wind⁴. In contrast, larger particles are heavier and tend to settle more quickly. Larger aggregates may offer microbes more protection against direct exposure to environmental stresses such as ultraviolet radiation and desiccation³.

Long distance transport of pathogens can result in damage to agricultural crops⁵ and outbreak of diseases among livestock in farms⁶. Short-range transport can also give rise to localised transmission of infectious diseases⁷ and poor indoor air quality can affect work and living spaces within the built environment⁸. The field of bioaerosol research has been gaining traction due to interest in studying the spread of SARS-CoV2⁹. Therefore, it is important to track sources to understand the modes of transmission of airborne microorganisms. In this thesis, source tracking investigations were conducted using the Smart Air Sampling System (SASS) 3100 instrument and Copan 4N6 nylon Floq swabs. Samples were sequenced using the next-generation sequencing platform and microbes were profiled using bioinformatics tools.

1.2 Challenges of air sampling

Air has very limited nutrients for microbial growth and is a medium of transport for microbial propagule dispersal in dormant forms at a stage of their life cycles¹⁰. The concentration of

microbes in air is highly dependent on emissions from local sources. To accurately assess the microbial composition in air, a robust air sampling instrument and analytical method are both required. Microbial concentration in air is estimated to be 10^3 to 10^6 cells/m³ air^{11,12}, which is approximately 3 to 6 log fold fewer than the amount of biomass recovered from soil and marine environments¹³. Therefore, the major challenge in air sampling is to collect sufficient biomass in air over a short period of time. There is no consensus in the bioaerosol research community on a standardised air sampling method¹⁴. Hence, the choice of air sampler depends on the experimental design and application.

There are four main types of air samplers that have been reported in literature, the filtration-based air samplers, cyclonic liquid impinger, agar impactor and gelatin air samplers¹⁵⁻¹⁸. Table 1.1 summarizes the advantages and disadvantages of each type of air samplers.

Table 1.1 | Comparison of different types of air samplers. The advantages and disadvantages of four main types of air samplers for different experimental design and applications.

	Advantages	Disadvantages
Filtration-based air samplers	<ul style="list-style-type: none"> • More effective at capturing solid particles¹⁵. • Higher microbial diversity and OTUs observed with next-generation sequencing and bioinformatics method based on Shannon index¹⁵. • High volumetric flow rate of up to 310 liters/min. 	<ul style="list-style-type: none"> • Larger pressure drops (5%) than other air filters¹⁵. • Lower extraction efficiency of the filter with minimal 4 ml of buffer required to achieve the best efficiency¹⁵. • Electrostatic stress on filter makes it harder for cultivation of viable microbial cells¹⁵.
Cyclonic liquid impinger	<ul style="list-style-type: none"> • Preserve integrity and viability of microbial cells¹⁹. • Easier sample preparation for enumeration and microscopy¹⁸. • High volumetric flow rate of up to 310 liters/min. 	<ul style="list-style-type: none"> • Lower microbial diversity and lesser OTUs recovered with next-generation sequencing and bioinformatics method based on Shannon index¹⁵.
Agar impactor	<ul style="list-style-type: none"> • Direct cultivation of microbial cells on agar plates. 	<ul style="list-style-type: none"> • Only selects for culturable microorganisms on agar plates.
Gelatin air samplers	<ul style="list-style-type: none"> • Higher recovery of viable microbial cells and viruses that are culturable¹⁹. 	<ul style="list-style-type: none"> • Lower volumetric flow rate of up to 125 liters/min.

The SASS3100 is an electrostatic filter-based air sampler. It is chosen because of its high volumetric flow rate for the efficient capture of biomass. Furthermore, it is also easy to use and does not require the handling and transport of sterile liquid solution. Lastly, it is also compatible for use with next-generation sequencing and bioinformatics approaches.

1.3 Next-generation sequencing and microbiome approaches

The advent of next-generation sequencing tools has revolutionized microbial ecological studies of airborne microbes in the recent decade²⁰. Previous studies heavily relied on isolation and cultivation of environmental microbes and scientist could only study the culturable fraction of microbes collected. The Great Plate Anomaly states that 99% of the microorganisms seen under the microscope cannot be grown in the laboratory²¹. DNA sequencing technologies matured and disrupted the field, enabling researchers to now identify and study microbes that were once uncultivable based on their genetic sequences. In fact, as predicted by Moore's law, the cost of DNA sequencing has become cheaper and affordable over the years. It is now widely available and is now an essential toolkit for microbiologists. Two commonly used sequencing methods namely amplicon²² and metagenomic shotgun sequencing²³ have been performed in environmental surveys. Amplicon sequencing involves amplifying a targeted 600-800 bp DNA, usually a marker gene like the rRNA gene with polymerase chain reaction (PCR). This technique is based on sequence similarity matches of the database of marker gene with highly conserved and variable sequences that can distinguish microbes between genera. Whilst metagenomic shotgun sequencing is untargeted and is based on sequence similarity matches of whole genome or large contig sequences assembled from short reads of 250 bp fragments off the Illumina sequencer. Comparing the two methods, amplicon sequencing is cheaper as it sequences only a short fragment and tends to give accurate matches up to the genus level with extensive reference sequences available in public database. In contrast, metagenomic shotgun

sequencing provides better relative abundance estimates as microbes usually have a single whole genome unlike the rRNA genes that have varying copy numbers for different microbial species. In addition, metagenomic shotgun sequencing is untargeted and it does not require any universal PCR primers. However, metagenomic reference databases are less extensive as compared to the amplicon-based databases. As a result, metagenomic approach may sometimes give unprecise identity matches based on the top closest possible hit if the whole genome of the microbe has not yet been sequenced. Although both amplicon and metagenomic approaches are powerful tools, they are unable to distinguish live from dead cells. Viability assays are still needed to test if the microbes are metabolically active²⁴.

1.4 Sources of bioaerosols

Sources in the built environment

The sources of microbes in the built environment may originate from exchanges of outdoor air, water sources (water pipes and cooling towers), moist surfaces (toilet flushing), and movement of occupants, fomites on contact transfer with objects and also in concealed building spaces²⁵ (Fig 1.1).

Fig 1.1 | Potential sources of microbes in the built environment. Microbes are carried indoors by air, water sources, moist surfaces, movement of occupants, fomites and in concealed spaces.

Air is a transport medium for microbes within buildings as occupants inhale and interact with surfaces and dust. Indoor air microbes have bearings on human health and the transport process of microbes are key to understanding the source of the problem²⁵. Notably, it is the composite of these four factors that may influence diffusion and transport: air leakage via unintentional openings in the building, internal movement of air between spaces, mechanical ventilation, and natural ventilation. Movement of microbes in residential homes are largely driven by air leakage and natural ventilation, depending on the occupants' preferences. Home-based air-conditioning systems are also used to cool indoor spaces, but these systems do not draw air from outdoor sources. In contrast, office buildings are managed by engineers that operate the HVAC systems using ventilation strategies to initiate air exchanges to save costs and energy²⁶. Air leakage via unintentional openings in buildings can lead to the diffusion of outdoor airborne microbes indoors. This could be due to unintended engineering defects which may create gaps such as leaky window seals and cracked walls. Air particle infiltration is driven by indoor and outdoor air pressure difference and changes with external weather conditions such as wind, temperature, and humidity²⁷.

Second factor, the internal movement of air between spaces, is largely shaped by the interior building layout such as wall, partition panels, furniture and the locations of mechanical ventilation and exhaust panels. Computational Fluid Dynamics (CFD) modelling can aid in the investigation of airflow patterns to help track sources of microorganisms indoors. Airborne microbes may settle on these surfaces and surfaces with high moisture content can promote microbial growth as microbial sinks or environmental reservoirs.

The third factor, mechanical ventilation, refers to HVAC systems that draw air from outdoor source to incorporate it with indoor air. Most of the indoor air is being recirculated. Mechanical ventilation is the key driver of airflow in office buildings to disperse indoor airborne microbes. The outdoor air is filtered through a Minimum Efficiency Reporting Value (MERV) rated filter

before it enters the room. Higher MERV value means smaller pore size with better filtering capability. Most office buildings use low MERV rated filters due to cost-savings and are not very effective in preventing the infiltration of outdoor air microbes. In the interest of saving cost and energy, building engineers also turn off HVAC systems at night and over weekends. Switching off the systems intermittently, stop pressurized airflow and allow microbes that are already collected on the filters to grow²⁸. The cooling coil systems also tend to be sources of contamination due to the high moisture content²⁹.

Lastly, natural ventilation refers to opening windows letting outdoor air exchange and circulate within the buildings. Most residential homes rely on this mode of air exchange and outdoor microbes can drift with stochastic air streams into the buildings driven by wind.

There are also notable differences between HVAC systems in residential homes and office buildings. Residential homes have HVAC units that recirculate interior air for temperature control with low-efficiency particle filtration and no outdoor air intake. Outdoor air enters the homes through unintentional building leakage in closed homes or driven by outdoor weather conditions through opening of windows depending on the preference of the human occupant. Office buildings including schools and workspaces usually have central HVAC systems with particle filtration and liquid aerosolization associated with air conditioning cooling coils that can be potential sources of microbes. These centralized HVAC systems are designed to have outdoor air intakes mixed with recirculated indoor air. According to the American Society of Heating, Refrigerating and Air Conditioning Engineer (ASHRAE) 62.1 standard, office buildings should have a minimum outdoor intake of 0.3 L/s per m² for offices³⁰. Larger buildings require operations and maintenance through proper filter changing, systems inspections and repair, control sensor calibration.

Every building has its own unique microbiome and is constantly shaped by a variety of factors such as geographic location, age of building, occupancy rates, HVAC, ongoing operations, and

maintenance frequency³¹. Therefore, the indoor air microbiome is highly variable and reflects the activities of human occupants as well as the conditions of the mechanical and plumbing systems.

Humans as sources of airborne microbes

The human body has 10^{12} - 10^{14} microbes living on the skin and in the gastrointestinal tract³². Microbes can be dispersed by humans either through direct contact with surfaces or respiratory droplets. It has also been reported that humans can release up to 10^6 bioaerosols per hour and every individual has a unique microbial cloud³³. These human-associated microbes were detected on surfaces, dust and bioaerosols in the built environment³⁴⁻³⁶. Hence, humans are the biggest contributor of bioaerosols indoors and plays a significant role in influencing the indoor air and surface microbiomes.

Pathogens may be aerosolised when patients breathe, speak, cough or sneeze³⁷. These pathogen-containing bioaerosols can stay afloat in air for hours, facilitating the spread of infectious diseases^{38,39}. Therefore, it is critical to study the microbiomes of the air, surfaces, and human lungs to identify microbial sources and better model routes of transmission between the environment and the human host.

The lung was initially thought to be sterile probably because of the difficulty to obtain samples from the lower airway without performing an invasive bronchoscopy procedure⁴⁰. With advances in next-generation sequencing and bioinformatics approaches, the lung was shown to harbour microbes and a healthy lung microbiome typically consists of *Prevotella*, *Veillonella* and *Streptococcus* spp^{41,42}. Sputum has been clinically used as it is non-invasive which only requires the patient to perform deep chest cough and expel mucus from the lower airway unlike bronchoscopy samples. However, sputum samples tend to be contaminated by microbes from the upper airway and saliva. Nonetheless, the sputum sample is still a good surrogate of the lung microbiome⁴³.

Asthma is a disorder where the individual has a chronic airway inflammation with recurring symptoms of wheezing, breathlessness, chest tightness and coughing while patients with chronic obstructive pulmonary disease (COPD) have more severe airflow restrictions that are not fully reversible with damage in the airway and parenchyma tissues⁴⁴. Bronchiectasis (BE) patients exhibit even more severe reduction in lung function than COPD as the lungs in BE patients are not able to perform mucus clearance resulting in inflammation and damage of bronchial walls⁴⁵. An important distinction between asthma and COPD/bronchiectasis is that inflammation is driven by different group of immune cells. Eosinophils, mast cells and CD4+ T lymphocytes are responsible for the inflammation in asthma⁴⁶ while neutrophils, macrophages and CD8+ T lymphocytes are involved in COPD and bronchiectasis⁴⁷. Corticosteroid inhalers have been demonstrated to be effective against eosinophilic response in asthma but are less effective in COPD and bronchiectasis⁴⁸.

The lung microbiomes of asthmatic individuals are typically enriched in pathogens such as *Haemophilus* and *Neisseria* spp, with reductions in commensal *Prevotella* and *Veillonella* spp. The reduction of commensal species is more prominent in COPD and bronchiectasis patients. By studying the microbial exchanges between the lung, air, and surfaces, the bioaerosols fluxes between air, surfaces and human lung can be determined.

Plants as sources of airborne microbes

Plants shed and exchange microorganisms especially through the aerial parts of their system. Leaves constitute a large surface area with potential air-surface exchanges. The leaf surface is estimated to have 10^6 - 10^7 bacterial cells per square centimeter⁴⁹. The leaf surface is a stressful and nutrient-limiting environment. Epiphytes are leaf surface inhabiting microbes and they develop adaptive traits to enable them to be resistant against abiotic stresses such as ultraviolet radiation⁵⁰, oxidative stress and desiccation⁴⁹. Epiphytes can also work together to form biofilm and produce extracellular polymeric substances to help them survive the environmental

stress^{51,52}. To colonise the leaf surface, epiphytes will first encounter the cuticle of the plant. The cuticle layer, the outer most surface layer of the leaf, is composed of biopolyester, cutin, wax, phenolics, cutan and polysaccharides. The cuticle has water-permeable properties that can influence surface charges and pH and also functions as a physical barrier between the plant and epiphytes^{53,54}. The functional relationship between leaves and the microbiome is not well understood. Some studies suggest that epiphytic bacteria can secrete biosurfactants to increase diffusion of water across the leaf cuticle surface, thus increasing the moisture content to favour bacterial growth⁵⁵⁻⁵⁷. Some plants also produce methanol as a metabolic by-product during cell wall degradation for growth and development⁵⁸. *Methylobacterium* is able to metabolise methanol and have been found on leaf surfaces in several plant species⁵⁹. Hence, to survive, the leaf microbial communities need to develop adaptations against stresses such as ultraviolet radiation, reactive oxygen species (ROS) and desiccation.

Plants perform transpiration and gas exchanges through their stomata. The stomata can become vulnerable ports of entry for pathogenic microbes. Each stoma is guarded by two guard cells which can recognize an array of pathogen-associated molecular patterns (PAMPs) which includes flagellin, chitin, and chitosan⁶⁰. Upon detection of PAMPs, the plant will be alerted of possible pathogen invasion and trigger the closure of the stomatal pores⁶⁰⁻⁶². Studies have shown that pathogens like *Pseudomonas syringae* can evade recognition by secreting toxins to suppress stomatal closure and exploit the opportunity to enter the leaf tissues^{61,63}.

To combat against pathogens, plants have a two-tier innate immune system known as the pattern-triggered immunity (PTI) and the effector-triggered immunity (ETI). PTI is triggered upon recognition of PAMPs and is the first line of defense in plants through phytohormone signaling, secretion of antimicrobial compounds, production of reactive oxygen species (ROS), stomatal closure and mitogen-activated protein (MAP) kinase cascades^{64,65}. Plants need phytohormone ethylene to develop resistance against phytopathogens such as *Pseudomonas*

syringae and *Magnapotha oryzae* via the increased ROS production⁶⁶⁻⁶⁸. This was evident in a separate study that showed that *Arabidopsis* mutant ethylene insensitive2 (*ein2*) had a different leaf microbiome as compared to the wild-type plant⁶⁹. The ETI is a second-tier response following after PTI signalling. ETI is activated by intracellular resistance (R) proteins which sense pathogen effector proteins via the intracellular nucleotide-binding/leucine-rich repeat (NLR) receptors⁷⁰. ETI upregulates the PTI responses to increase ROS oxidative burst, nitric oxide, and salicylic acid phytohormone production, which then induces rapid programmed cell death to contain the pathogen⁷¹. ROS plays a key role in plant defence, and it shapes the leaf microbiome.

Different plants have its own unique microbial cloud and it is possible to track the microbial signature of plant species based on its microbiome. Lymperopoulou, et al. ⁷² measured and found that the abundance of bacteria and fungi were increased by 2 to 10-fold in the 50 m downwind air above plants as compared to upwind air samples with no plants. The study also showed that wind can carry up to 50% of the bacteria from the plants. A separate study has also shown that wall surfaces and surrounding floor areas can become contaminated with archaea, bacteria and fungi from the plants if plants were placed in a cleaned indoor controlled environment for 6 months⁷³. Therefore, plants like humans can also shed their microbiomes to their environmental surroundings.

1.5 Source identification methods

Identifying microbial sources can be challenging especially when microbes are not visible on-site during inspection. There are three methods to source identification, namely database analysis method (DAM), correlation analysis method (CAM) and model analysis method (MAM)¹⁰. Table 1.2 summarises the advantages and disadvantages of each method.

Table 1.2 | Comparison of the source identification methods. Advantages and disadvantages of source identification methods based on the search and source tracking algorithms¹⁰.

Source Identification Method	Software Tools	Advantages	Disadvantages
Database Analysis Method	BLAST search, MetaMetaDB, Global catalogue of microorganisms (GCM) databases	<ul style="list-style-type: none"> • Does not require samples to be collected from source locations. • Accurate match scores, requiring further analysis to calculate source contribution rates. 	<ul style="list-style-type: none"> • Requires a long time to compute. • Unknown source relationship to sample.
Correlation Analysis Method	Principal component analysis (PCA)	<ul style="list-style-type: none"> • Known source relationship to samples. • Easy to use. 	<ul style="list-style-type: none"> • Require samples to be collected from source locations. • Unable to quantify the source contribution rates
Model Analysis Method	Bayesian sourcetracker	<ul style="list-style-type: none"> • Calculates the source contribution rates of local sources based on samples collected. • Easy to use. 	<ul style="list-style-type: none"> • Require samples to be collected from source locations. • Calculations can only be made among samples collected.

DAM can be used to assess if a sample is a source based on sequence similarity to an isolate in the database. This can be performed using Basic Local Alignment Search Tool (BLAST) against sequences of isolates in the MetaMetaDB⁷⁴ and Global Catalogue of Microorganisms (GCM) databases⁷⁵. DAM is limited by the number of isolates and strain information in the database. Furthermore, some bacterial species may be generalists and are ubiquitous in various environments across the globe. The DAM method cannot calculate source contribution rates for paired samples collected within the same location.

Besides DAM, CAM has also been used for source identification. In this method, a principal component analysis (PCA) is performed and plotted in 2-dimensional axes to find correlation between the air and potential source samples. CAM was used by Bowers, et al. ⁷⁶ to show that

the airborne bacterial communities closely resembles the bacterial communities of the leaf. The major advantage of CAM is that the similarities between samples can be easily visualized. However, it is difficult to calculate the source contribution rates based on the arbitrary axes and eigenvector values derived from a PCA analysis.

Source Tracker Method (STM) is a new model analysis model (MAM) created by Knights, et al.⁷⁷. The STM software calculates Bayes' estimates based on air and potential source samples taking in account of unknown source environments. MAM has significant advantage over DAM and CAM methods as it gives quantitative probabilities as source contribution rates for samples collected within the same location. STM has been used in several bioaerosol-related studies such as source tracking of bacterial communities in air over Tokyo⁷⁴ and tracking the source of airborne fungi in mountainous and urban areas of China^{78,79}.

Among the three methods, the MAM method is the best option for addressing the source-sink problems in this thesis as all air and potential source samples were collected within the same localised sites. The STM software also gives quantitative probabilities of source contribution between samples.

Algorithm of the source tracker method software

The algorithm behind the original Bayesian source tracker is illustrated in Fig 1.2. Fig 1.2 shows a simplified source-tracking problem using three types of coloured balls in two source urns and a single sink urn. The number of balls and percent relative abundance in each source urn were tabulated. An inference on the conditional probability was made for red, green, and blue balls transferred into the sink urn from each source urn.

Fig 1.2 | Conditional probability of balls in the sink urn given two sources. The first table shows the actual counts of balls in each source urn. Then, in the second table, the percent relative abundance was tabulated against the sum of balls in each urn. The last table shows probability scores that were inferred based on the proportions of balls from each urn by dividing across the column sum of each coloured ball.

Fig 1.3 is a simplified conditional probability model where each coloured ball represents a microbial specie from each source environment. However, in the realistic world a fraction of the microbes may come from an unknown source, and this must be accounted for. To factor in this unknown source, the authors set an alpha parameter of 0.001 where it assumes that out of 100 sequences in the sink sample, $100 * 0.001 = 0.1$ of the read counts account for an unknown source. This alpha parameter is kept default at 0.001 unless specified. Fig 1.3 shows an updated version of the model with an unknown source. The absolute counts of each ball were updated with an addition of 0.1 and the probability scores were re-tabulated by dividing the new count with the column sum for each colour type.

Fig 1.3 | Conditional probability of balls given two sources, one unknown and one sink.

The first table shows the actual counts of balls in each source urn, with an addition of 0.1. The second table is the revised probability scores by taking counts from the first table divided by the column sum of each coloured ball.

To complete the probability model, the authors calculate the probability of a sequence in a sink assigned to a source randomly. The assignment is random where each sequence in the sink is given an index. A combined probability score is calculated for each indexed sink sequence by taking the probability of assignment multiplied by the conditional probability shown in Fig 1.4. Only when the sequence is assigned to the unknown, sourcetracker2 will update and add 1 to the alpha value for the unknown. Otherwise, no changes are made. The assignment is then re-shuffled in the next iteration using the Gibbs sampler. The process continues until a converge solution is achieved. Users can define a burn-in parameter of n to instruct sourcetracker2 to discard the first n sets of iterations before logging and averaging the scores together to get a final percent contribution of each source to the sink sample.

Fig 1.4 | Random assignment of each sink sequence was tagged to a source and combined probability scores were calculated in an iterative process. Each sink sequence was randomly assigned to a source and the probability of assignment was tabulated. For instance, the assigned green sink sequence was picked, and a combined probability score was calculated for the green ball.

1.6 Challenges in modelling the diffusion and transport of airborne microorganisms

The dispersal of microorganisms in the atmosphere can be thought of as a series of three-phase processes, take-off – transport – deposition⁸⁰. During take-off, a surface is regarded as a source and the atmosphere as a sink. Subsequently, the microbes become airborne and are transported in air across distances and are influenced by activities of insects, birds, and humans as well as passive stochastic events like wind and rain and classical physics with settling of bigger and heavier particles due to gravity. Once a microbe lands on a surface, the atmosphere now

becomes a source and the surface a sink. Hence, source tracking involves the tracking and modelling of the trajectories taken by airborne microorganisms from a source to their end destinations.

There were attempts to model the movement of spores mathematically. One of the earliest attempts was performed by Legg and Powell ⁸¹ in 1979 when they first created a two-dimensional model to predict the movement of barley mildew, *Erysiphe graminis*, in a barley crop. Their model yielded some initial success, but it did not fully account for some actual experimental observations where the spores are elliptical and that they tend to clump to each other falling faster than single spores. Improvements were made to the initial model to also account for gusts and random walk movements⁸². It remains challenging because spore movements do not follow a simple Gaussian or exponential decay profile due to the turbulent air flow. This has led to the development of more complex types of diffusion and transport models that give close approximation for the movement of bioaerosols.

The two most used models today are computational fluid dynamics (CFD) model and Hybrid Single Particle Lagrange Integrated Trajectory model (HYSPLIT). CFD is used for simulating movement of airflow in closed systems while HYSPLIT is used for long distance transport with metrological factors for open systems¹⁰. For example, CFD has been employed in agricultural farms to monitor gaseous pollutant emissions⁸³. CFD has also been used to analyze the spread of viruses or infectious agents in closed environments like the airplane⁸⁴ and indoor spaces⁸⁵. On the other hand, HYSPLIT has been used to explore the routes of transport of microbes across land and ocean environments^{86,87}. In this thesis, source-sink relationships are investigated in the built environment such as office spaces, residential homes, and balcony and hence CFD is the best modelling tool for the purpose of this work.

Algorithm behind computational fluid dynamics simulation

CFD is an *in silico* numerical method used to solve vectors in a 3-dimensional Cartesian coordinates system⁸⁸. First, the mesh of a geometry model is created. Imagine a barbeque wire mesh made up of small individual squares and each square represents a cell in the system. The simulation takes place within each cell and calculations must fulfill the laws of conservations in Physics when a vector passes through from one cell to its neighbor cell⁸⁸. This is represented by the vector operator ∇ in equation (1), when solving for vectors i , j and k relative to the change in x , y and z coordinates.

$$\nabla \equiv i \frac{\partial}{\partial x} + j \frac{\partial}{\partial y} + k \frac{\partial}{\partial z} \quad (1)$$

The velocity vectors are denoted as u , v and w , incorporating the velocity vectors and the rate of change of time with the ∇ equation in (1), we get equation (2).

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + u \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z} \quad (2)$$

The laws of conservations must be observed for mass, momentum, and energy in each cell. The rate of change in mass can be calculated by taking the rate of change in density, ρ , multiplied by the change of volume measured as changes in x , y and z in (3).

$$\frac{\partial m}{\partial t} = \frac{\partial}{\partial t} (\rho \Delta x \Delta y \Delta z) \quad (3)$$

Based on the continuity equation, the rate of change of mass equals to the net rate of change of mass entering the cell, taking account of velocity vectors that bring infinitesimal small particles into and out of the cell. In this model, the density is assumed to be constant (Fig 1.5).

Fig 1.5 | Schematic model of mass conservation in each cell of the CFD mesh. The net inflow of mass is equals to the rate of change of mass in each cell.

Besides conservation of mass, we need to also consider the momentum of the particles based on Newton's second law. The x-component can be expressed as follows:

$$F_x = m a_x \quad (4)$$

$$a_x = \frac{Du}{Dt} \quad (5)$$

There are two sources of forces experienced by the particles in motion namely, body forces and surface forces. Body forces is already described in (4) and (5). Surface forces are like friction opposing the motion of the object. For example, along the x-component, a particle will experience a force perpendicular or normal to the x-boundary planes when entering the cell. This force is termed normal stress stress, σ . Particle also experiences forces that are parallel to the boundary planes of the cell and this force is termed as tangential stress, τ . Both forces are illustrated in Fig 1.6 and 1.7.

Fig 1.6 | Normal stresses at the boundary planes of each cell. The normal stresses are resistance forces perpendicular to the boundary planes of the cell.

Fig 1.7 | Tangential stresses at the boundary planes of each cell. The tangential stresses are resistant forces parallel to the boundary planes of the cell.

Therefore, if both normal and tangential stresses are incorporated in our model. The x-, y- and z-component can be derived as follows:

$$\frac{Du}{Dt} = \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \tau_{yx}}{\partial y} + \frac{\partial \tau_{zx}}{\partial z} + \sum F_x^{body\ forces} \quad (6)$$

$$\frac{Dv}{Dt} = \frac{\partial \tau_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \tau_{zy}}{\partial z} + \sum F_y^{body\ forces} \quad (7)$$

$$\frac{Dw}{Dt} = \frac{\partial \tau_{xz}}{\partial x} + \frac{\partial \tau_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} + \sum F_z^{body\ forces} \quad (8)$$

The stresses can be derived numerically using Newton's law of viscosity as follows:

$$Viscosity, \tau = \mu \frac{du}{dy} \quad (9)$$

μ in the equation refers to the Dynamic viscosity and this equation is important when solving for Newtonian fluids in CFD. Since we are solving using the PIMPLE algorithm, the model assumes constant temperature and density. This simplifies to the calculations of pressure-linked velocity vectors.

The PIMPLE algorithm is a combination of PISO (Pressure-Implicit-of-Split-Operations) and SIMPLE (Semi-Implicit-Method-of-Pressure-Linked-Equations) solvers in CFD. The solver does the calculations by an iterative trial-and-error process to find a converged solution ensuring that the continuity equation is met. First, the computer will make a guess of the pressure, p^* . Then it will correct for pressure where $p' = p - p^*$ to obtain velocity gradients of u^* , v^* and w^* along x, y and z as shown in (10), (11) and (12). Lastly, both the left and right sides of equation (13) must be balanced.

$$u - u^* = -D^u \frac{\partial p'}{\partial x} \quad (10)$$

$$v - v^* = -D^v \frac{\partial p'}{\partial y} \quad (11)$$

$$w - w^* = -D^w \frac{\partial p'}{\partial z} \quad (12)$$

$$\frac{\partial}{\partial x} \left(D^u \frac{\partial p'}{\partial x} \right) + \frac{\partial}{\partial y} \left(D^v \frac{\partial p'}{\partial y} \right) + \frac{\partial}{\partial z} \left(D^w \frac{\partial p'}{\partial z} \right) = \frac{\partial u^*}{\partial x} + \frac{\partial v^*}{\partial y} + \frac{\partial w^*}{\partial z} \quad (13)$$

This iterative process is illustrated as a grid in x-y plane in Fig 1.8.

Fig 1.8 | Solving for u and v vectors for particle movement from one cell to another. The computer algorithm first guesses the pressure to find the velocity gradient and corrects the pressure and velocity vectors in an iterative manner. The movement of the particle in the x - y plane can be expressed as n , s , e , w directions.

Turbulence, advection, diffusion, and destruction of fluid particles must also be accounted for in transport equations of the model. Therefore, the k -omega shear stress transport model is selected as it is used for low Reynolds numbers fluid and is suitable for air ventilation simulation in buildings.

1.7 Research gaps and aims of study

Previous bioaerosol studies have uncovered the air and surface microbiomes of the built environment and identified potential sources using the Bayesian source tracker method. These studies, however, did not consider the airflow patterns, bioaerosol fluxes and source emission rates from surfaces, air, and respiratory droplets. Therefore, the aim of this thesis is to identify microbial sources by mapping source emission rates and fluxes of bioaerosols between the air, surface, humans, and plants in the built environment.

The primary objective of this work is to identify potential sources, quantify source emission rates and predict possible trajectory paths of bioaerosols. Source tracking of bioaerosols is a highly complex problem that requires an interdisciplinary solution. Metagenomics, Bayesian probability, Markov Chain, and computational fluid dynamics (CFD) models would be used to address the source tracking problem.

The second objective is to examine the impact of human's breath on surfaces and the air microbiome. Having access to chronic respiratory patients' sputum specimens will enable the study of microbial transfers between the human lung and the environment.

Finally, the last objective of this work is to examine potential source emissions from plant leaves as leaves have large surface areas that are exposed to wind and rain. Leaf microbes may be aerosolised by wind and dispersed by splashing of rain droplets.

Chapter 2. Source tracking of the air microbiome in buildings

2.1 Introduction

Sequencing-based studies have shown that the indoor built environment consists of mostly bacterial and fungal species originating from outdoor air through gaps in building openings^{89,90}. The microbial composition of outdoor and indoor air is influenced by temperature and relative humidity and hence varies at different geographical locations and seasons⁹¹. The microbiome of the tropical outdoor air in Singapore is largely dominated by fungi⁹². It was observed that the tropical airborne microbes follow a precise diel cycle in tandem with predictable diurnal weather patterns⁹². While most fungi originate from outdoor sources, a review study has shown that humans, pets, plants, plumbing systems, heating ventilation and air conditioning systems (HVAC), and dust resuspension are also potential indoor microbial source contributors⁹³.

Moisture-damaged buildings (MDB) are often caused by water intrusion due to plumbing, HVAC, roof leakages or flooding. Moulds are ubiquitous in air, and it is common for fungi such as *Aspergillus* spp., *Penicillium* spp., *Acremonium* spp. and *Wallemia* spp to proliferate in the presence of excessive moisture^{94,95}. Furthermore, building engineers also attempted to save energy by sealing leaky windows to reduce the electricity needed to cool the room and switching off the HVAC during non-office hours at night⁹⁶. While these approaches result in cost-savings, the reduced ventilation rates due to tighter window rubber seals trap moist air within the building especially when the HVAC is turned off for prolonged periods at night, causing even more damp and mould-related problems^{97,98}.

Numerous studies have hypothesized that long-term exposure to these fungal-dominated MDB environments can lead to sick building syndrome (SBS)⁹⁹⁻¹⁰². Nonetheless, the cause-and-effect relationship between MDB and SBS still remain controversial^{95,103}. Additionally, diagnosis of SBS is also largely descriptive with generic symptoms of headache, dizziness, nausea, eye, nose or throat irritation, itchy skin and cold¹⁰⁴. Therefore, source tracking investigations are

warranted to help identify sources and mitigate the risk of SBS. However, it is challenging and expensive to study the aerodynamic movement of particles due to the infeasibility of conducting experiments using fluorescent tracers in custom-built wind tunnels and sensors to track particle movements. An alternative attractive solution is to use *in silico* models instead to simulate airflows indoors. Though powerful, all computational simulation models are only predictions of the actual particle flow trajectories, which should be validated with some form of experimental evidence.

There are two major categories of *in silico* modelling. The first category involves the use of probabilistic models to predict mould growth in buildings. Temperature, moisture, ventilation rates and material properties can be used as input parameters to predict possible building locations with moisture risks that may be conducive for mould growth and contamination using Monte Carlo simulations^{105,106}. In probabilistic models, the output is dependent on the randomness and events that take place between the discrete source and sink states. It is computationally faster and requires less storage space¹⁰⁷.

The second category involves the use of deterministic models based on computational fluid dynamics (CFD) simulations to predict airflow and the movement of particles in a room with underlying sets of Navier-Stokes equations as shown in a study predicting mould growth in buildings¹⁰⁸. In deterministic models, the output is calculated based on the parameters and input used. It is computationally expensive as flow is calculated at every time step continuously in every cell. Data generated from the simulation results and the 3D geometry model require significantly more storage space¹⁰⁷.

A three-way combinatorial approach of metagenomics, probabilistic and deterministic models would be used to address the source and sink problems encountered in this thesis (Fig 2.1).

Fig 2.1 | A three-way combinatorial approach for source tracking. Air and swab samples were collected to gather experimental information of species present and their relative abundances. Combined probabilistic and CFD models can give additional insights into probabilities of microbial source emissions and transfers between sampled locations.

The National Institute for Occupational Safety and Health (NIOSH) in United States Centre for Disease Control and Prevention (CDC) established a set of guidelines for workplace safety using a model of hierarchy of controls starting from the most effective to the least effective method of risk management in descending order: source elimination, substitution, engineering, administrative and personal protective equipment (PPE) controls to remove workplace health hazards¹⁰⁹ (Fig 2.2). The NIOSH model of hierarchy of controls can be applied to contain microbial source exposure in the built environment. Disinfecting and replacing contaminated sources are the most effective approaches as sources of hazards are eliminated. Chemical fumigation and dehumidifying regimes are effective in reducing the microbial load in air but are not effective solutions for de-contaminating surfaces. Lastly, isolating the contaminated room and advocating the use of surgical masks can reduce direct exposure and contact times but it does not solve the root of the mould problem.

and collecting air samples in the same room^{35,112-114}. Subsequently, DNA were extracted from microbes collected in swab and air samples and microbial communities were identified by either amplicon or metagenomic sequencing followed by downstream bioinformatics analyses. To compute probabilities of relative source contributions of swab surfaces to the corresponding indoor air, the Bayesian source tracker software was typically used.

There are limitations associated with this existing source tracking method. Firstly, this method only accounts for unidirectional transfer of microbes from surfaces to air and attempts to track airborne microbes back to its possible source. Conversely, it can be argued that the air could also potentially be a source, seeding surfaces with microbes through settling of dust. Secondly, the source tracker analysis does not consider the airflow pattern in the room. To overcome these problems, the Bayesian hypothesis can be tweaked to predict future time steps instead of backwards time steps for source tracking. In this case, each sampled location is thought of as source emission locations rather than source contribution events. The modifications would be further discussed in detail in the discussion section. Probability of transfers can be plotted in a Markov Chain transition graph. This reframed probabilistic model can be used to map various sources and provide information on the strength of source emissions. The CFD simulation can further supplement the Markov Chain graphs to visualise source-sink exchanges in a room.

2.2 Materials and methods

Swab and air sampling methods

Both air and surface samples were collected from two moisture-damaged offices in NTU, Office A (1°20'40.3"N 103°40'42.6"E) and Office B (1°20'43.1"N 103°40'45.0"E) respectively. Swab were collected from surfaces that were deemed to be most representative in the room such as the chair, table, floor, partition walls, glass panel and ceiling. Air samples were collected at a flow rate of 150 L/min for 4 hours using the Smart Air Sampling System, SASS

3100 at 12 noon and 12 midnight. After sampling, the air filters and swabs were transported back to the laboratory. DNA were extracted using the DNeasy PowerWater kit (Qiagen, Germany) and sequenced on the Illumina HiSeq 2500. Sequence reads were then queried using KAIJU, a faster and sensitive taxonomic classifier¹¹⁵. All output files were visualised in MEGAN v6.12.0. The sampling locations of each office were outlined in Fig 2.3 and 2.4, respectively.

Fig 2.3 | Sampling locations in office A. Swab locations were arbitrarily chosen by taking samples of all representative surfaces in the room. Swabs were collected from the chair, ceiling, floor, and table. Air samples were collected as duplicate samples in office A, a neighbouring office space, hallway and outdoor balcony.

Fig 2.4 | Sampling locations in office B. Like in office A, swab locations were arbitrarily chosen by taking samples of all representative surfaces in the room. Swabs were collected from partition panel, laptop bag, back of chair, ceiling and aircon panel in office B. The air samples were collected as duplicates in office B and the adjacent director's office.

Source tracker and modifications used

Metagenomes of the air and swabs were exported in the biom format and imported into sourcetracker2 for Monte-Carlo simulations with a burn-in parameter of 100 iterations. Probabilities were first calculated by defining air as sink and surfaces as source and vice versa. The Bayesian hypothesis were modified, and results obtained were re-interpreted as follows:

$$\text{Original Source Tracker:} \quad \text{Source} \xleftarrow{P(\text{Source} | \text{Sink})} \text{Sink} \quad (1)$$

$$\text{Reframed Bayesian Hypothesis:} \quad \text{Location B} \xleftarrow{P(B | A)} \text{Location A} \quad (2)$$

$$\text{Location n} \xleftarrow{P(n | A)} \text{Location A} \quad (3)$$

In (2) and (3), the probability of a bioaerosol particle landing at locations B to n is calculated given the microbiomes observed at location A.

Output of transfer probabilities from source tracker were re-interpreted as transfers between locations and were plotted as Markov Chain transition graphs.

Computational fluid dynamics modelling

The geometries of the moisture-damaged mouldy rooms were reconstructed to scale as 3D models in Blender v2.7.1^{116,117}. Each geometric object such as the chair, table and aircon panels were exported as separate STL files. Meshing of the models was created using blockMesh and snappyHexMesh in OpenFOAM v1912¹¹⁸.

Boundary conditions for all objects were defined as wall functions except for air inlets which were defined with a fixed inlet air velocity of 0.1 m/s with a turbulent kinetic energy of 5% intensity and outlets defined as zero gradient functions with fixed flux atmospheric pressure of 0 Pa. The simulation was performed using the PIMPLE solver algorithm in OpenFOAM using the Reynolds averaged Navier-Stokes equations for Newtonian fluids and incompressible flow. The completed simulations were visualized in Paraview v5.0.1¹¹⁹. The geometric models were set to an opacity of 0.3 and were overlaid with glyph arrows that represent the velocity vectors and kinematic pressure scalars in the rooms. Temperature and relative humidity parameters were assumed to be constant in the CFD simulations performed.

2.3 Results

Metagenomic and source tracking results of office A

Metagenomic results show that the air in office A (meeting room air) is highly contaminated by fungi, with 20% of reads assigned to *Aspergillus versicolor*, 5% of reads to *Aspergillus sydowii* and 1% of reads to *Steccherinum ochraceum* (Fig 2.5). The bar chart shows the top 12 species ranked by relative abundance, other species ranked 13 and below, and higher taxonomic levels as well as no hits and unassigned reads.

Fig 2.5 | Metagenomic profiles of air and swab samples collected in office A. *Aspergillus* and *Penicillium* spp. were identified as the most dominant taxa in the room, with highest abundance found in the ceiling sample. *Wallemia mellicola* was detected with highest relative abundance on the floor. About 50% or more reads were either no hits or not assigned.

The dominant fungal species in office A were mainly of the genera, *Aspergillus*, *Penicillium* and *Wallemia* spp. Both *Penicillium* and *Aspergillus* species were also successfully cultivated from swabs collected from the ceiling, chair, and floor in the office on Malt-Extract agar supplemented with 100 µg/ml streptomycin (Fig 2.6). Based on the metagenomic data in Fig 2.5, results suggest that the ceiling, table, and floor were strong potential sources of the airborne contamination in office A.

Fig 2.6 | Fungi cultivated from ceiling, chair, and floor in office A. The fungal isolates were grown from swabs in malt extract agar with 100 µg/ml streptomycin. Polymerase chain reaction and Sanger sequencing was performed to identify the fungal species.

The relative abundances in percentage read counts were used for analysis in source tracker. The output probabilities were interpreted based on the modified Bayesian hypothesis and were plotted as Markov Chain transition graph as shown in Fig 2.7. The model shows that the airborne fungi in office A have about 7% chance of leaving the contaminated office into the neighbouring clean office and 8% chance of leaving the room into the hallway with outdoor air exchanges possibly due to the opening and closing of the door of the meeting room.

Fig 2.7 | Markov chain transition graph of microbial transfers between sites in office A.

The ceiling exhibited the strongest dual source-sink relationships in both directions, suggesting that the ceiling is the main source of the mould problem. Fungi were also detected on the floor, chair, and table. The probabilistic model provides information on the source emission rates between surfaces and air exchanges for each pairwise samples.

Airflow simulation of office A was conducted using CFD (Fig 2.8). Since the HVAC systems have MERV-8 rated filters installed, the air entering the room was filtered for coarse PM10 particles but not fine PM2.5 particles. The air entered the room through the diffusers and flowed over the fungal infested chairs, table, and floor surfaces. The air became contaminated as it picked up fungal spores and hyphal fragments which were circulated onto the sidewalls and upwards to the ceiling panels, before exiting through the return air grilles. Most of the contaminated air were trapped within the room and were recirculated back into the room as convective currents.

Fig 2.8 | CFD simulation of airflow in office A. Air flow trajectories was simulated using the PIMPLE algorithm to solve velocity vectors and pressure with the Reynolds averaged Navier-Stokes equations. The model assumes an inlet air flow rate of 0.1 m/s from the four ceiling HVAC diffusers and air as Newtonian fluid with incompressible flow in the system.

The air flow pattern in Fig 2.8 explains the Markov chain probabilistic model in Fig 2.7 showing bidirectional exchange between the ceiling and air, where the air is also a strong source to the ceiling due to the convective air flow.

Metagenomic and source tracking results of office B

Metagenomic results show that the air in office B (main office air) is contaminated by fungi, with less than 5% of reads assigned to mostly *Aspergillus* species (Fig 2.9). The bar chart shows the top 12 species ranked by relative abundance, other species ranked 13 and below, and higher taxonomic levels as well as no hits and unassigned reads. Surfaces in office B were mainly colonised by *Aspergillus* species and they bear high resemblance to the air samples from the main office and director's office. (Fig 2.9).

Fig 2.9 | Metagenomic profiles recovered from air and swab samples collected in office B.

Aspergillus spp. were identified as the most dominant taxa in the room, with the highest relative abundance detected among the textile material surfaces. The meeting room ceiling tile had higher proportions of *Actinomycespora* bacteria.

Actinomycespora spp. and *Cordyceps confragos* are soil-related microbes that were detected on the meeting room ceiling tile. Unlike office A, the ceiling tile in office B had brown patches of soil stain. It was probably contaminated during renovation works and was carried from outdoors by workers. These microbes were also not found in air and hence the ceiling tile was not considered as a strong source of the air in office B. *Aspergillus restrictus* was successfully isolated from the partition panels in MEA-streptomycin. Further analysis with the probabilistic source tracking model shows that textile materials such as the laptop bag, partition panels and recycling bag contributed largely as sources to the main office air (Fig 2.10).

Fig 2.10 | Markov chain transition graph of microbial transfer between sites in office B.

The partition panels, glass panel, chair backrest, recycling bag and panel behind shelf were potential sources to the indoor office air. The director's room and the main office are connected directly through a door.

CFD air flow analysis can help track the possible dispersal paths of fungal spores. In Fig 2.11, air entered the office through the five aircon panels near the ceiling. The blue arrows indicate the locations where there is slightly lower kinematic pressure. The laptop bag and chair backrest had visible signs of mould and were in the proximity of the areas with blue arrows (Fig 2.11). The model suggests that there were higher concentrations of fungal spores entrapped in areas of the room between the confined hidden spaces within the furniture.

Fig 2.11 | CFD simulation of airflow in office B. Air flow trajectories was simulated using the PIMPLE algorithm to solve velocity vectors and pressure with the Reynolds averaged Navier-Stokes equations. The model assumes an inlet air flow rate of 0.1 m/s from the five panels near the ceiling and air as Newtonian fluid with incompressible flow in the system.

2.4 Discussion

Aspergillus spp. were mainly identified in both offices A and B. Office A had a temperature profile of 23.4 °C and relative humidity of 87.1% while office B had a temperature profile of 22.9 °C and relative humidity of 71.7%. Office A suffered from a more severe mould problem than office B. It is likely due to the higher relative humidity and poor air ventilation in the enclosed room. Temperature was maintained at 23 °C by the HVAC but the relative humidity was not regulated and exceeded the range of normal healthy levels. According to the U.S. Environment Protection Agency (EPA), the healthy indoor relative humidity should be maintained below 60%, and ideally between 30% and 50%¹²⁰. In fact, both temperature and relative humidity are interdependent parameters as temperature also influences the saturation vapour pressure⁹².

The indoor air temperature can have an influence on the indoor air microbiome as different microbes grow preferentially at different temperature range. However, there are conflicting

arguments in the literature where some studies reported positive correlations between higher temperatures and higher relative abundance of indoor air microbiomes^{36,120,121} while two other studies reported inverse correlations^{91,122}. The HVAC is usually operated in the range of 18 to 25 °C to maximise the comfort of the human occupants and many microbes also thrive in the same temperature range¹²³. Temperature in both offices A and B were kept constant at 23 °C and is likely not the driving factor for the mould problem observed.

Relative humidity could be another factor that may be associated with the mould problem seen in both offices A and B. It has been shown by other studies that high relative humidity may be associated with microbial growth on surfaces¹²³⁻¹²⁵ and may mediate higher contact transfer between fomites and human occupants¹²⁶. Carpet when exposed to high relative humidity can promote microbial growth due to its thick adsorbent woven fabric. *Wallemia* sp. was detected in the carpet of office A and was also detected in a carpet tile incubated in a controlled chamber at 90% relative humidity in a separate study¹²⁷. In contrast, low relative humidity may be associated with increased aerosolization rates from surfaces¹²⁸⁻¹³⁰. Bioaerosols may drift longer in air at lower relative humidity as the more humid the air is, the more water the aerosols absorb, and the more likely the particles become heavier and fall to the ground faster¹³¹. Despite numerous associations reported, a study demonstrated that relative humidity had no direct control over fungal growth as *Penicillium* and *Aspergillus* spp. have been shown experimentally to grow even at low relative humidity of 11% if water is available on the surface¹³². Therefore, surface moisture level plays a more crucial role than relative humidity. A more reliable measure of moisture on surfaces is to use Equilibrium Relative Humidity (ERH) instead of the ambient relative humidity. ERH can be measured by using a resistive electrolytic¹³³, capacitance¹³⁴ or dew point¹³⁵ hygrometer. The higher the moisture level on a surface, the higher the likelihood that a surface is contaminated with microbes. ERH differs from the relative humidity of air as the moisture level on a surface is affected by the surface

temperature of a material that may differ from the air. It has been shown that materials with a higher ERH tends to be more susceptible to microbial contamination¹³⁵. However, it is challenging to measure ERH on-site without damaging the room as a sample of the material must be torn or broken off and placed in the instrument chamber.

Table 2.1 | The relationship between equilibrium relative humidity (ERH) and building materials. Materials with a higher ERH are more susceptible to microbial growth³¹.

Building materials	Equilibrium relative humidity (ERH)
Wood	70%
Gypsum	85%
Concrete	90-95%

The gypsum ceiling tile in office A is the strongest source and has the highest bidirectional exchange with air, followed by source emissions from the table and carpet floor. In office B, textile materials such as the laptop bag, partition walls, chair backrest are strong sources of *Aspergillus* fungi. Results suggest that different types of materials have different propensity for microbial colonisation. It has been shown that bio-based construction materials are more susceptible to microbial contamination¹³⁶. *Aspergillus* and *Penicillium* were the most reported fungal species found in moisture-damaged buildings^{137,138}. Table 2.2 summarises the types of building materials that the top four fungal species seen in office A are often associated with.

Table 2.2 | List of fungal species that are commonly found in moisture-damaged buildings and their associated growth on different types of building materials. Bio-based materials are more prone to fungal colonization and *Aspergillus* and *Penicillium* spp. appear to be generalists and they grow on most building substrates.

Species	Building Materials
<i>Aspergillus</i> spp.	Ceramic ^{138,139} , gypsum ¹⁴⁰ , paint ^{138,139} , glue ^{138,139} , glass fiber ¹⁴¹ , silicon organic plastics ¹⁴² , fluorine plastics ¹⁴² , cellulose ¹⁴² , polyimide ¹⁴² , rubber ¹⁴²
<i>Penicillium</i> spp.	Painted wallpaper ¹⁴³ , glass fiber ¹⁴³ , concrete, gypsum ¹⁴⁴ , polyethylene ¹⁴² , polyimide ¹⁴² , rubber ¹⁴² , cellulose ¹⁴² , silicon organic plastics ¹⁴² , polyvinyl chloride ¹⁴²
<i>Acremonium</i> spp.	Painted wallpaper ¹³⁸ , glass fiber ¹⁴¹
<i>Wallemia</i> spp.	Dust ¹²⁷ , carpet ¹²⁷

Of the four species shown in Table 2.2, *Aspergillus* and *Penicillium* spp. appear to be generalists and are ubiquitously found on most building surfaces, whereas *Acremonium* and *Wallemia* spp. seem to be specialists that grow only on certain preferred substrates under specific conditions. Notably, cellulosic materials and gypsum boards are one of the most vulnerable substrates. Cellulosic materials are at the highest risk as cellulose can be easily metabolised as food source by many naturally occurring microbes. Gypsum board contains traces of inorganic calcium and sulphur minerals that can support fungal growth under moist conditions. Besides cellulose and gypsum, the next group of susceptible materials includes paint, textile, glue, wallpaper, and wood materials³¹. Paint, for example, contains carboxymethyl cellulose, which can act as a carbon source for microbes¹⁴⁵. Materials with rough and porous surfaces are also prone to contamination as dust, organic particles and moisture are easily trapped within, providing a microenvironment that is favourable for microbial growth¹⁴⁶.

Source tracking investigation remains useful to identify contaminated building surfaces so that disinfection can be more targeted to eliminate the source of the problem. CFD simulation result reveals that office A is more poorly ventilated than office B. The contaminated air in office A was continuously being recirculated, aggravating the mould problem. Office B had more furniture with hidden spaces that may obstruct and congregate airflow. The hidden spaces can become microenvironments suitable for mould growth. Both offices were mechanically ventilated with MERV-8 filters to keep outdoor microbes out of the buildings. MERV-8 filters are cheap and are usually replaced once after every 3 to 6 months by the building facility management. However, MERV-8 filters are not very effective for filtering *Aspergillus* and *Penicillium* conidia spores that are 2-3 µm in size. Changing to higher MERV rating filters may help alleviate the problem. Cleanrooms and surgery theatres, for instance, use MERV-18 and MERV-20 filters which have HEPA filtration capabilities in removing particles that are as small as 0.2 µm in size. Nonetheless, these filters are very expensive and are often only used in critically sterile environments.

An alternative solution is to dilute the indoor concentration of airborne microbes by introducing natural ventilation through opened windows or directing air flow using inlet and exhaust fans to ensure sufficient air exchanges¹⁴⁷. Air change rate can be measured using a specialised equipment that is installed over the inlet or exhaust air panels. Table 2.3 shows the recommended air changes per hour for different rooms based on the guidelines from the American Society for Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE)¹⁴⁷.

Table 2.3 | Recommended Air Changes per Hour (ACH) for different room types¹⁴⁷.

Rooms that require sterile environments have a higher air change rates to prevent build-up of indoor microbes.

Room Type	Air Changes per Hour (ACH)
-----------	----------------------------

Operating Room	15
Intensive Care Unit	12
Laboratory	6
Office, Classroom	2
Corridors	1

After source identification, cleaning and disinfection steps should be taken to eliminate and reduce the concentration of these harmful microbes. Ultraviolet radiation (UV) and chemical disinfectants are the two most effective disinfection methods.

UV radiation requires direct exposure and line of sight to inactivate microbes on a surface. UV cleaning procedures have been shown and proven to be effective in reducing the number of healthcare-associated infections (HAIs) cases in hospitals^{148,149}. UV disinfection has also been applied in HVAC systems by deploying UV lamps in ventilation air ducts¹⁵⁰ and cooling coils¹⁵¹ to reduce microbial growth.

Chemical disinfectants, on the other hand, typically contain quaternary ammonium salts, surfactants or oxidizing agents that can kill or inactivate microbes¹⁵². For example, chlorine bleach contains sodium hypochlorite and when added to water can become hypochlorous acid which is toxic to microbes¹⁵³. Cavicide® contains diisobutylphenoxyethoxethyl dimethyl benzyl ammonium chloride, a quaternary ammonium salt that can disrupt cell membranes¹⁵⁴. Virkon® consists of a powerful oxidizing agent, potassium peroxydisulfate which can cause oxidative stress and cell lysis¹⁵¹. They are highly effective but when used in high concentrations, disinfectants can also become potential sources of volatile organic compounds (VOCs) indoors. Secondly, there are also environmental concerns of increased antimicrobial resistance against chemical disinfectants^{155,156}. Disinfection can eliminate or reduce source contamination, but it does not necessarily prevent microbial growth on surfaces. In this case,

antimicrobial coatings can come in useful. For example, studies have shown that silver nanoparticles with antimicrobial properties can be coated on textile materials such as nylon and silk fibers¹⁵⁷, building surfaces like gypsum boards and grout¹⁵⁸ and medical devices¹⁵⁹.

Explanations on the modifications made to source tracker software

The original Bayesian source tracker performs backward calculations of probabilities from the sink to the source states in equation 1.

$$\text{Original Source Tracker:} \quad \text{Source} \xleftarrow{P(\text{Source} | \text{Sink})} \text{Sink} \quad (1)$$

Users must define *a priori* values, specifying if a sample is either a source or a sink. Defining the air as a sink and surfaces as sources and vice versa directs the software to calculate the posterior probabilities. Results of sourcetracker2 were reported either as pie charts or bar graphs (Fig 2.12a). The sourcetracker2 approach is useful especially in cases where flow from sources to sink environment is explicitly known. For example, in the case of water samples collected from river upstream that are sources of the sink sample collected at the downstream estuary. However, in the built environment, no explicit source-sink relationships between surfaces and air can be inferred directly, as source-sink directionality can exist in both ways. Instead, the Bayesian hypothesis can be reframed as shown in (2) and (3) below.

$$\text{Reframed Bayesian Hypothesis:} \quad \text{Location B} \xleftarrow{P(B | A)} \text{Location A} \quad (2)$$

$$\text{Location n} \xleftarrow{P(n | A)} \text{Location A} \quad (3)$$

The new hypothesis is like the original sourcetracker2 algorithm in (1) except states are now defined as locations rather than as source or sink states. In this way, all samples are treated equally as sources where the probabilities of source emissions were examined instead of source contributions. Equation (1) is analogous to a pie chart of sink proportions with slices contributed from various sources while in (2) and (3), the pie chart now represents the proportions found in location A with each slice distributed to various locations (Fig 2.12b).

The calculations are the same. Transition probabilities can now be plotted in a Markov chain transition graph, and according to Markov chain theory, all probabilities leaving each node must sum to 1 (Fig 2.12c).

Fig 2.12 | Reframing the Bayesian hypothesis underlying the sourcetracker2 algorithm.

(a) The original balls in urns probabilistic algorithm used by sourcetracker2 to calculate the posterior probabilities of source contributions to a given sink environment. (b) Reframing of the Bayesian hypothesis where the sourcetracker2 algorithm can be repurposed to calculate

probabilities of source emissions from location A distributed to locations B, to n and an unknown location. (c) The probabilities output of the modified Bayesian hypothesis model can be represented as arrows in a Markov chain transition graph.

In fact, Bayes' theorem can be applied to solve both forward and inverse problems (Fig 2.13)^{107,160}. The modified model takes the forward problem approach as shown in Fig 2.13a. In this case, source tracker is used to predict the future time steps of microbes at possible locations based on the microbiomes observed. Like the “presumption of innocence” in law, every sample in question is first regarded as “innocent until proven guilty”. The modified model assumes that every sample emits a set of microbes to an end location and there is no assumption of precedence whether the surface microbiome existed before the air microbiome or vice versa. In contrast, the original source tracker employs the inverse problem method (Fig 2.13b). Existing source tracking studies often define the air as the end state (sink) and surfaces as the starting states (sources)^{93,161,162}. It assumes that the air microbiome is a consequence of surface microbiomes and that microbes on surfaces existed before the airborne microbes.

Fig 2.13 | Forward and inverse problem modelling using Bayes' Theorem. (a) The cause-and-effect relationship is well defined with finite number of causes in a forward problem model as opposed to (b) the inverse problem model with infinite number of possible causes.

The forward problem can be expressed mathematically as follows:

$$P(B_i) = \sum_i^{N_B} P(B_j^n)P(B_i|B_j^n)$$

Where B is the microbiome and N_B is the total number of species in the microbiome going to location i (end) from location j (start). It can be summarised as the transfer probability between the new location Y given old location X .

$$X \xrightarrow{P} Y$$

Since the air and swab samples were collected at around the same time, it is reasonable to reframe the Bayesian hypothesis such that all samples are treated equally as sources with the same starting point where the probabilities of source emission are quantified as transitional probabilities between locations. The output is also compatible with Markov chain theory where all probabilities leaving each node must sum to 1 (Fig 2.12c). However, it still depends on the analyst to make rational decisions on whether a pairwise relationship exist between two location nodes in the Markov chain transition graph. For example, it does not make sense to connect the floor and ceiling samples together. Therefore, air is placed in the middle of the transition graph as air is the common medium that connects between the various surfaces. This probabilistic model can be supplemented with air flow simulation diagrams for mapping of bioaerosol fluxes in the built environment.

Comparing the obtained results between the modified model and the original source tracker

Results obtained from the modified Bayesian model were compared against the original source tracker. Fig 2.14 shows pie charts of the source contributions of surfaces to the indoor air in office A and office B using the original source tracker. The results suggest that the ceiling is

the main source of the problem, followed by the table and chair in office A. All other samples such as the floor, hallway air and common office air contribute to less than 10% of the indoor air. Notably, the biggest difference between the results from the modified model and the original source tracker model lies in the floor sample in office A. The modified model suggested 63% of source emission from the floor while the original model suggested only 3% of source contribution to the air. It is important to first look at the raw data that was used for modelling. The taxonomic relative abundance chart in Fig 2.5 shows that the floor sample contains 20% of reads assigned to *Aspergillus versicolor* and 15% to *Wallemia mellicola* while the air sample contains 20% of reads assigned to *Aspergillus versicolor* and less than 1% to *Wallemia mellicola*. The metagenomic chart suggests a transfer but there is a 60% discrepancy between the transfer probabilities calculated between the two models.

Fig 2.14 | Results from the original source tracker software without modifications. The main source of the mould problem in office A is the ceiling while the main source of the mould problem in office B comes from the air from the director's office and outdoor air.

The problem with inverse problem modelling is that with a greater number of sources sampled, the percentage contribution becomes smaller and is suppressed by other source samples as the air can be contributed by an infinite number of sources (Fig 2.13b). Conversely, the advantage

of forward problem modelling is that the percentage of surface emission does not depend on the number of other samples as transfer is only considered at the emission site between the surface-air interface. If a surface has a high relative abundance of *Aspergillus* and the air has a high relative abundance of *Aspergillus*, then the source emission will be high. The forward problem approach will be able to highlight and identify contaminated surfaces with higher probabilities of transfer.

For office B, the original source tracker model predicts that the main source of contamination comes from the director's office air and outdoor air. This is driven by the high similarity in microbiomes between air samples as seen in Fig 2.9. However, the other surface samples were not given sufficient weights as it is suppressed by the large number of other source samples considered in the analysis. The relative abundance chart in Fig 2.9 also shows that most of the textile surfaces contain the top fungal species that are also present in air, but surprisingly it only appears to contribute to about 3-6% of the indoor air based on the original source tracker model. On the other hand, the modified model manages to identify the contaminated surfaces with higher source emission rates of 70-90%. The modified model also has merits as direction of transfers can be mapped in a bidirectional manner using a Markov Chain transition graph.

2.5 Conclusions

The microbiome of the built environment is influenced by factors such as temperature, moisture, material, ventilation, airflow, building design and human activity. Office A and B provide two different case studies of mould problem in MDB. Office A is contaminated by *Aspergillus*, *Penicillium* fungal species on building surfaces such as the gypsum ceiling tile, table and *Wallemia mellicola* on the carpet floor. The high relative humidity of 87.1% coupled with poor ventilation rates may have contributed to the increase in airborne microbial concentrations. Office B is also contaminated by *Aspergillus* species and is caused by contaminated textile

materials such as partition boards, chair seats and bags. The high relative humidity of 71.7% in office B also presents itself as a conducive environment for mould growth. Disinfection should be carried out at the identified sources. Susceptible materials such as the partition boards can be replaced with plastic boards or coated with antimicrobial surface coating to prevent re-occurrence of the mould problem.

Most existing studies were conducted either by microbiologists or engineers separately. Microbiologists have isolated microbes in petri dishes and characterised the surface and air microbiomes of buildings while engineers have studied human comfort, air flow and energy efficiency of buildings. It will be helpful to combine both expertise and leverage on the strengths of metagenomics, probabilistic and CFD models for source tracking.

The models used in this study, however, did not consider longitudinal time-series events, human occupancy rates, and the dispersal of volatile organic compounds (VOCs). Future studies can be performed to correlate these environmental factors and metadata with the changes in microbiome over time. Source emission experiments can also be performed artificially by varying concentrations of nebulised microbes from synthetic sources in a controlled chamber and wind tunnel.

Chapter 3. Humans as sources of the air microbiome

3.1 Introduction

Lung diseases are associated with increased bacterial infections due to a consequence of mucus build-up and the dilation of airway¹⁶³. 16S rRNA sequencing studies have unveiled that the healthy lung consists of a diverse microbiota with members primarily from the genera, *Streptococcus*, *Prevotella* and *Veillonella*¹⁶⁴⁻¹⁶⁶. In contrast, lung diseases are often implicated with a dominance of bacteria from the genera, *Haemophilus* and *Pseudomonas* and are likely due to the inefficient discharge of mucus and weak immunity in patients, leading to the regional growth of microbes inhaled from air¹⁶⁷. As a result, respiratory patients are often prescribed with high doses of antibiotics.

Antibiotic therapies may contribute to the dysbiosis of the healthy lung microbiome and selective pressures from such therapies may further lead to the emergence of antimicrobial resistance¹⁶⁸. A study has also shown that the diseased microbiome profiles are heterogeneous where some patients exhibited minimal microbiome perturbations and others having significant changes especially in the relative abundances of *H. influenzae* and *P. aeruginosa*¹⁶⁹. The impact of antimicrobial exposure on the lung microbiome and the development of antimicrobial resistance (AMR) remain challenging to assess to date¹⁷⁰. As sequencing cost becomes cheaper, metagenomics is becoming an emerging molecular tool to characterize microbiomes and the underlying microbial functions for disease diagnosis and treatment¹⁷¹. Metagenomic investigations have only been recently extended to respiratory studies and it has the potential to study the accumulated effects of antibiotic exposures in the microbiomes¹⁷². Additionally, the lung microbiome has also been reported to be involved in antibiotic degradation functions and the diminishing efficacy of antibiotic penetration due to secretion of matrix components¹⁷³. Unlike the targeted 16S amplicon approach, metagenomic shotgun sequencing has been preferentially used in this study because it is less biased to polymerase chain reaction (PCR)

amplifications and is not influenced by copy number variations^{174,175}. No studies-to-date have collectively compared the sputum microbiomes across non-diseased, severe asthma (SA), chronic obstructive pulmonary disease (COPD) and bronchiectasis (BE) individuals.

Poor indoor air quality due to high concentrations of harmful airborne microbes can cause allergies and lead to health problems such as the ‘sick building syndrome’. These indoor microbes can be disseminated via three routes of transmission in the built environment, (1) direct contact transfer with surfaces, (2) bioaerosols release through the human breath or shedding of skin and hair and (3) resuspension of dust³³. Source tracking investigations are therefore warranted to quantify the source emissions of bioaerosols between the lung and potential sources in the built environment that may impact the occupant’s respiratory health.

This chapter seeks to explore the impact of antimicrobial exposure on the healthy and diseased lung microbiomes. The study can help inform and create new strategies in clinical practice to understand the development of antimicrobial resistance. Secondly, fluxes in bioaerosols of the source-sink interactions between the human lung and environment can also be mapped using the probabilistic model developed in Chapter 2 based on the metagenomes of the sputum and swab samples collected from potential sources such as the inhalers, fans, air conditioners and air purifier in residential homes.

3.2 Materials and methods

Sample collection

11 severe asthma (SA), 15 chronic obstructive pulmonary disease (COPD), 15 bronchiectasis (BE) patients and 13 non-diseased volunteers were recruited from Singapore General Hospital (SGH) and Tan Tock Seng Hospital (TTSH) for this study (Table 3.1). The patients were classified into four groups, SA, COPD, BE and ND based on guidelines from the Global Initiative for Asthma guideline treatment ladder for SA, the Global Initiative for Chronic

Obstructive Lung Disease (GOLD) criteria for COPD and dynamic computed tomography for BE. Non-diseased volunteers are healthy individuals with no existing lung medical conditions. Sputum samples were obtained by the huff coughing manoeuvre technique. A swab sample of the inhaler mouthpiece was collected from 31 patients with inhalers, $n = 16$ for SA, $n = 11$ for COPD and $n = 4$ for BE.

Air sampling was also performed using the SASS3100 at the patients' homes and surface swabs were also collected from the fan blades, air conditioner panels and air purifier grilles to perform environmental source tracking. This analysis comprised of a composite of swabs from air conditioner panels ($n = 7$), fan blades ($n = 11$), air purifier grilles ($n = 1$) and air filters from indoor air ($n = 9$) and outdoor air ($n = 9$) samples.

Table 3.1 | Demographics of chronic respiratory patients in the study

Demographic	ND	D			P Value	
		Severe Asthma	COPD	Bronchiectasis	ND vs. D	D vs. D
<i>n</i>	13	11	15	15		
Age, yr	34 ± 8	70 ± 17	70 ± 9	64 ± 15	<0.001	0.472
Sex					0.528	0.012
M	62% (8)	73% (8)	100% (15)	53% (8)		
F	38% (5)	27% (3)	0% (0)	47% (7)		
BMI, kg/m ²	23 ± 3	27.9 ± 8	21.7 ± 8	18.4 ± 6	0.237	0.009
FEV ₁ % predicted	109 ± 14	72 ± 14	45 ± 13	65 ± 20	<0.001	0.005
Disease severity*						
Mild	—	0% (0)	7% (1)	6% (1)	—	0.797
Moderate	—	27% (3)	40% (6)	27% (4)	—	—
Severe	—	73% (8)	53% (8)	67% (10)	—	—
Exacerbations in year preceding study recruitment	0 ± 0	0 ± 1.5	3 ± 3	2 ± 1	0.032	0.013
Smoking status					<0.001	<0.001
Never-smoker	100% (13)	91% (10)	0% (0)	67% (10)		
Ex-smoker	0% (0)	0% (0)	47% (7)	27% (4)		
Current smoker	0% (0)	9% (1)	53% (8)	7% (1)		
Antibiotic use in 6 mo preceding recruitment					0.050	<0.001
Yes	0% (0)	0% (0)	67% (10)	20% (3)		
No	100% (13)	100% (11)	33% (5)	80% (12)		
Inhaled corticosteroid use						
Yes	—	100% (11)	53% (8)	53% (8)	—	0.064
No	—	0% (0)	47% (7)	47% (7)	—	—
Inhaled bronchodilator use						
Yes	—	100% (11)	100% (15)	53% (8)	—	0.009
No	—	0% (0)	0% (0)	47% (7)	—	—

Taxonomic and functional assignment of metagenomic sequence reads

The sputum, swab and air samples were transported back to the laboratory for DNA extraction using the DNeasy PowerWater kit (Qiagen, Germany) and sequenced on the Illumina HiSeq 2500. Sequence reads were then queried using KAIJU, a faster and sensitive taxonomic

classifier¹¹⁵. All output files were visualized in MEGAN v6.12.0. Raw sequence reads were selected, and adapter trimmed using Cutadapt (version 1.14)¹⁷⁶ with a minimum Phred Score of 20 and >30bp in length.

Trimmed reads were mapped against the GRch38 human reference genome using bowtie2 (version 2.3.2)¹⁷⁷. Unmapped non-human reads were selected and aligned against the NCBI non-redundant protein database (7 August 2017) with Diamond (version 0.9.9)¹⁷⁸. Results were then subjected to the Lowest Common Ancestor (LCA) algorithm implemented in MEGAN v6.12.0 with a minimum score of 100 and a support of ≥ 25 ¹⁷⁹. Functional gene annotation was performed in MEGAN using the Kyoto Encyclopedia of Genes and Genomes (KEGG) ‘acc2Kegg-Dec2017’ database. ARGs in the reads were identified using ShortBRED¹⁸⁰. Unique protein markers were created by running the ‘shortbred_identify.py’ script with a 95% cluster identity using ARGs protein sequences from Comprehensive Antibiotic Resistance Database (CARD) and Uniref90 database as a reference. Relative abundance of the unique protein markers was determined by running the ‘shortbred_quantify.py’ script which was normalized based on average read length of 200 bp, expressed in reads per kilobase million mapped reads (RPKM) relative to each ARG. Drug class assignments were manually validated against the CARD database.

Data visualization and statistical analyses

Functional gene assignments were visualized as heatmaps coloured according to Z-scores in MEGAN while RPKM values of ARGs were plotted with ‘heatmap.2’ using ‘gplots’ package in R¹⁸¹. The list of ARG genes was extracted and plotted in a Venn diagram using both the ‘venndiagram’ and ‘upsetR’ packages to examine similarities in ARG profiles across healthy and diseased groups^{182,183}. Principal coordinate analysis (PCoA) plots based on microbiome data were created using ‘vegan’ and ‘phyloseq’ packages to observe clustering patterns among groups^{184,185}. The average distance to centroid in each group was calculated using PERMDISP

tests of homogeneity of dispersions by the ‘betadisper’ function in ‘vegan’. Significance of distances to centroid dispersions between groups was assessed by analysis of variance (ANOVA) with Tukey’s post hoc analysis. All taxonomic assigned reads were visualized in MEGAN as bubble charts. Co-occurrence networks were generated by combining normalized AMR and taxonomic data and assessed several network metrics using generalized boosted linear models (GBLMs) as described by Faust, et al. ¹⁸⁶. Pearson correlation for co-occurring taxa and ARG were visualized in R using the ‘rcorr’ function in the ‘Hmisc’ package, considering only strong correlations and significance of $p > 0.05$ ¹⁸⁷.

Source tracker analysis and modifications used

Metagenomes of the sputum, inhalers, fan blades, air conditioner panels, air purifier grilles and inhalers were averaged into one representative sample for each respective sample type. The metagenomic profiles were then exported to biom format and imported into sourcetracker2 for Monte-Carlo simulations with a burn-in parameter of 100 iterations. Probabilities were first calculated by defining air as sink and surfaces as source and vice versa to get Bayesian estimates of bidirectional transfer. The Bayesian hypothesis were then modified, and results were re-interpreted as follows:

$$\text{Original Source Tracker:} \quad \text{Source} \xleftarrow{P(\text{Source} | \text{Sink})} \text{Sink} \quad (1)$$

$$\text{Reframed Bayesian Hypothesis:} \quad \text{Location B} \xleftarrow{P(B | A)} \text{Location A} \quad (2)$$

$$\text{Location n} \xleftarrow{P(n | A)} \text{Location A} \quad (3)$$

Output probabilities from source tracker were re-interpreted as transfers between location states instead of source-sink states and were plotted as Markov Chain transition graphs.

3.3 Results

Functional metabolic dysbiosis of the airway metagenome

The functional classification of the airway metagenomes was highly variable among chronic obstructive pulmonary disease (COPD) and bronchiectasis (BE) patients (Fig 3.1a). Based on the literature, dysbiosis is defined as the disturbance in microbial metabolism and is usually accompanied by changes in the microbiomes associated with diseased states^{188,189}. In Fig 3.1a, dysbiosis was observed especially among COPD and bronchiectasis with down-regulation of carbohydrate and amino acid related pathways which were shifted towards lipid-associated pathways associated with antibiotic biodegradation and synthesis. This finding is in line with recent studies that suggest that lipid metabolism is altered among chronic respiratory disease patients¹⁹⁰⁻¹⁹². Patients with severely impaired lung function also tend to have higher levels of sphingolipids in their sputum due to increased levels of ceramide needed to mediate alveolar cell apoptosis^{193,194}.

Healthy individuals can be distinguished from the diseased states. The diseased individuals have higher relative abundances of aminoglycoside, bicyclomycin, diaminopyrimidine, multi-drug, peptide, phenicol, sulphonamides, sulfones and triclosan resistance as shown in Fig 3.1b. The black dots in Fig 3.1c indicate antibiotic use in the last 6 months in individual patient represented in each column. Results suggest that recent antibiotic exposure did not necessarily translate to a detectable higher relative abundance of antibiotic resistance genes (ARGs). The detection of β -lactam, fluoroquinolone, macrolide, and tetracycline resistance genes across both healthy and disease states further suggest the presence of a core airway resistome.

Fig 3.1 | Dysbiosis of functional airway metagenomes in COPD and BE patients. Shotgun metagenomics show metabolic dysbiosis with increasing antibiotic resistance gene abundance in chronic respiratory disease states. (a) Heatmap of sequence reads that were classified based on functionally assigned categories of KEGG database with higher read abundance in red that were associated with lipid metabolism, xenobiotic biodegradation, and antibiotic-associated biosynthetic pathways especially among COPD and bronchiectasis patients. (b) Heatmap of specific antibiotic-resistance genes and their relative abundances based on the Comprehensive Antibiotic Resistance Database (CARD). Notably, β -lactam, fluoroquinolone, macrolide, and tetracycline resistance genes were also detectable across diseased and non-diseased individuals.

Nonetheless, COPD and bronchiectasis patients have the highest number of antibiotic resistance genes. (c) Reported antibiotic usage for each respective class of antibiotics in the last six months in each patient before airway sampling are illustrated with black dots.

The core airway resistome

The composition of antibiotic resistance genes that make up the resistome were examined. COPD and bronchiectasis patients were found to have the highest number of antibiotic resistance determinants (Fig 3.2a). The diseased state is distinct with the presence of β -lactam, fluoroquinolone, macrolide, multi-drug, phenicol and tetracycline resistance genes. Healthy individuals and severe asthmatics had lesser antibiotic resistance genes (ARGs). A subset of 18 ARGs was found to be common in all cohorts, including the healthy individuals. This represents the ‘core resistome’, that were predominantly from the β -lactam, fluoroquinolone, macrolide, and tetracycline classes (Fig 3.2). The genes that make up the ‘core resistome’ consisted of mainly macrolide resistant genes including *msrD* (*mel*), *ermB*, *ermX* and *ermF* genes, followed by tetracycline resistant genes *tetW*, *tetA*(46), *tetB* (46), β -lactam (*cfxA2*), and fluoroquinolone (*pmrA*) (Fig 3.2c).

a.

b.

connecting them indicate the intersections between/among the compared groups. (c) Heatmap of the specific antibiotic resistance genes grouped by class along the x-axis and cohort (ND; SA; COPD and BE) in the y-axis. Genes are sorted in the order of detected abundance starting from *msrD* (*mel*), *ErmB*, *ErmF* and *ErmX* macrolide resistance genes followed by other genes encoding for tetracycline, β -lactam, and fluoroquinolone resistance.

The microbial ecology of the core airway resistome

Dysbiosis of the respiratory microbiome was evident especially among chronic respiratory patients and this finding is in line with the existing literature with an observed increase in *Rothia mucilaginosa* and *Pseudomonas aeruginosa* (Fig 3.3a)¹⁹⁵..

α -diversity analysis of the microbiome between groups varied with healthy individuals showing the highest microbial diversity compared to disease based on the Shannon index. Simpson index showed evenness in the microbial communities which were similar across the cohorts while severe asthmatics show reduction in Chao-1 species richness (Fig 3.4). This result suggests that species diversity and richness were both reduced in diseased states. Diseased microbiomes were distinguishable with higher relative amounts of Actinobacteria (*Rothia* spp.) and proteobacteria (*Pseudomonas* and *Haemophilus* spp.) and reductions in *Prevotella* spp., *Treponema* spp., *Fusobacteria* spp as compared to healthy individuals. β -diversity analysis shows that bronchiectasis cohort had the highest heterogeneity, followed by COPD (Fig 3.3b). In contrast, the microbiomes in healthy and severe asthma individuals were more evenly distributed. This can be further confirmed with analysis of average centroid distances, which was observed to be largest in COPD and bronchiectasis patients (Fig 3.3c).

Fig 3.3 | COPD and bronchiectasis patients have the greatest heterogeneity in their microbiomes. (a) Bubble chart of microbial abundance of discriminant taxa in non-diseased vs diseased cohorts based on species-level classification, where bubble size corresponds to read count and colour coding is based on phylum taxonomical classification. (b) Non-metric multi-dimensional scaling plot (nMDS) showing β -diversity between cohorts including ND: Non-Diseased (dark blue), SA: Severe Asthma (light blue), COPD: Chronic Obstructive Pulmonary Disease (purple) and BE: Bronchiectasis (red), each cluster distinguished by coloured ellipses. (c) The average distances to centroid from the nMDS plot was measured using the PERMDISP test of homogeneity and plotted for each respective study group illustrating the heterogeneity of their respective microbiome profiles. Difference in average distance to centroid was validated statistically by ANOVA and Tukey post-hoc analysis; *** $p < 0.001$.

Fig 3.4 | α -diversity analysis of the microbiomes across cohorts. (a) Shannon, (b) Simpson and (c) Chao-1 diversity indices among ND: Non-Diseased, SA: Severe Asthma, COPD: Chronic Obstructive Pulmonary Disease and BE: Bronchiectasis.

Since macrolide resistance genes were the most abundant and key components of the resistome, Macrolide resistance genes were further analysed in a co-occurrence network graph to infer microbe-gene associations. The macrolide resistance gene, *ermX*, is the most highly connected node associated with several microbial taxa such as *Neisseria*, *Veillonella*, *Actinomyces*, *Leptotrichia*, *Granulicatella* and *Streptococcus* spp (Figure 3.5). *ermF* and *msrD* genes have lesser number of microbial associations. There was a strong association between *ermF* and the gut microbe *Bacteroidetes thetaiotaomicron*. Similarly, a strong association between *msrD* and the gut pathogen *Clostridiodes difficile*. It is likely that these microbes acquire the resistance genes through horizontal gene transfer as it may be necessary for the lung resident microflora to maintain these resistance genes to survive the antibiotic-stress environment.

Fig 3.5 | Co-occurrence network graph of gene-microbe associations of the core macrolide resistome. (a) Antibiotic resistance genes were coloured according to antibiotic resistance classes within the co-occurrence network graph. Edges denote the interactions between nodes while line thickness is the strength of the observed interactions. (b) The same co-occurrence network graph as in (a) but coloured according to the species. Nodes of macrolide resistome showing associations between specific microbial species with association to (c) *ermX*, (d) *ermF* and (e) *MsrD*. Arrowheads indicate the directionality of the co-occurrence prediction.

Functional metagenomics show that inhaler devices are antibiotic resistance reservoirs

The swabs collected from inhaler devices were also subjected to metagenomics sequencing. The metagenomes of 31 patients' sputum and inhaler were averaged into representative sample of each sample types. Significant overlap was found between the airway metagenome and inhaler (Fig. 3.6). 80 microbial species were shared between the airway and inhaler and the

relative proportions of species were largely similar, with a larger reduction in the number of reads assigned to *Rothia mucilaginosa* and *Pseudomonas aeruginosa* observed on the inhaler surface as compared to the airway sputum. The airway had 36 unique species while the inhaler had 127 unique species.

Fig 3.6 | Overlap of microbes between airway and inhaler metagenomes. The relative proportions of microbial species were similar with slight reduction in the relative abundances of *Rothia mucilaginosa* and *Pseudomonas aeruginosa* seen in the inhaler relative to the airway.

The inhaler surface harbours higher number of microbes ($n = 207$) which was about twice the amount detected in the airways ($n = 116$). Next, when the antibiotic resistance genes (ARGs) in airway and inhalers were compared, a similar trend was observed (Fig 3.7). There were 53 overlapping genes, where the inhaler had 98 ARGs and airway had 89 ARGs. Based on this result, it is reasonable to suggest that the inhaler swab, when subjected to metagenomics, could be used as a surrogate measure of a patient's airway microbiome.

Fig 3.7 | Overlap of antibiotic resistance genes between airway and inhaler metagenomes.

The relative proportions of antibiotic resistance genes detected were similar between the airway and inhaler device. Reduction in multidrug resistance genes was observed in the inhalers.

Source tracking was performed to identify potential environmental sources in their residential homes that may have an impact on the airway microbiome. Fig 3.8 shows the microbiome profiles of the indoor surface microbiomes, the inhaler sputum, and the indoor and outdoor air. The indoor surfaces, such as the fan blades, air conditioner panels and air purifier grilles bear more resemblance to each other, with higher number of reads assigned to *Hortaea wernicki* (black yeast), *Aspergillus sclerotialis* and *Skermanella aerolata*. These microbes detected are associated with indoor dust^{196,197}. Dust encapsulated on these surfaces may become aerosolised and resuspended when the human occupant switches on the fan, air conditioner or air purifiers. The indoor and outdoor air in Fig 3.8, on the other hand, are enriched with more fungal species found in tropical Singapore air⁹².

Fig 3.8 | Relative abundance of microbial species found on the indoor surfaces, in air and human sputum samples. The swab samples from the fan, air conditioner and air purifier harbour higher concentrations of dust-associated microbes, while the air samples have higher amounts of outdoor air fungi. The human sputum and inhaler show a subset of the species present in both dust and air samples which are highly specific to the airway microbiome.

Unlike the dust and air samples, the sputum contains a unique subset of microbes from both the surface and air and is largely consisted of mainly commensal human lung microflora and potentially opportunistic pathogens. The inhaler microbiome bears high resemblance to the sputum sample, and this was also seen in Fig 3.6 with the overlap between the inhaler and swab metagenomes. Source tracking analysis was performed using the modified Bayesian model and the model suggests that microbes found on the air conditioner panels and air purifier grilles have a probability of source emissions of 10% to the airway microbiome while the fan and indoor air have a probability of source emissions of 5% (Fig 3.9). Notably, the inhaler exhibits the highest bidirectional transfers of more than 80% with the airway microbiome.

Fig 3.9 | Source tracking and identification of potential sources of microbes to the lung microbiome. The inhaler has the highest bidirectional exchange with the lung. The aircon and air purifier exhibit a modest source emission probability of 10%, while the fan and indoor air have probabilities of 5% of source emissions to the lung. Source tracking also suggest that there are outdoor and indoor air exchanges, primarily driven by air flow from outdoor sources.

3.4 Discussion

The dust assemblages on the fan blades, air conditioner panels and air purifier grilles comprise of a mixed community of human-associated and environmental microbes (Fig 3.8). These surfaces are potential sources as they are the outlet interfaces of air circulating devices. The indoor air, on the other hand, consists of mainly sporulating bacterial and fungal spores originating from outdoor sources. These microbes can be aerosolised and inhaled by the human occupants. The respiratory system has physical barriers such as microvilli, mucus layer, and immune cells to deter pathogens and select for beneficial microbes in the respiratory tract^{198,199}.

Multidrug-resistant Gram-negative bacteria such as *Acinetobacter baumannii*, *Escherichia coli*, *Streptococcus pneumoniae* and *Klebsiella pneumoniae* can colonise the airway and cause inflammation and exacerbations in COPD and BE patients^{200,201}. These opportunistic pathogens were also detected in the sputum of the chronic respiratory patients (Fig 3.8). Heterogeneity was observed in the airway microbiome, especially in COPD and BE patients (Fig 3.3). Interestingly, a core resistome of 18 genes, mostly macrolides, but also includes β -lactams and fluoroquinolones was detected. This does not necessarily relate to recent exposure to antibiotics. Importantly, a high abundance of Macrolide-Lincosamide-Streptogramin (MLS) resistance genes were found to dominate the airway resistome and are likely related to the widespread and increasing use of macrolides in the clinics²⁰²⁻²⁰⁵. The dominance of macrolide resistance genes *msrD* (*mel*), *ermB*, *ermX* and *ermF* within the core resistome is related to efflux and rRNA methylases mechanisms which were also found extensively in the urban sewage and human gut^{206,207}. Among the macrolide resistance genes identified, *ermX* represents the most interconnected node in the co-occurrence analysis with associations to several resident airway bacteria such as *Alloprevotella*, *Veillonella* and *Streptococcus*. It has also been shown that these macrolide genes can be aerosolised and were detected in the face masks of COPD patients²⁰⁸. Interestingly, the presence of gut microbes, *Clostridiodes difficile* and *Bacteroides thetaiotaomicron* associated with *ermF* and *msrD* genes were also detected in the airway.

In addition, this study also further demonstrates that the inhaler swab, when subjected to metagenomics, can act as a surrogate of the lung microbiome and resistome. Swabbing the inhalers could be helpful especially among dry and non-sputum producing severe asthmatics who have difficulty in collecting enough sputum for lab analysis. Co-occurrence of microbe-resistance gene analysis further confirms the occurrence of horizontal transfer of resistance genes between the host and the inhaler devices, allowing migration of pathogenic microbes and antibiotic resistant genes (ARGs) between the airway and environment.

The source tracking diagram in Fig 3.9 shows that the inhaler can act as both an environmental source and sink of the lung microbiome. The lung microbiome is distinct from the microbiomes of the other surfaces and the inhaler surface bear high resemblance to the lung microbiome as microbes are transferred via contact through the oral cavity. Similarly, toothbrushes have been shown to harbour oral resident microflora and ARGs like the inhalers seen in this study²⁰⁹. The inhaler may provide a ‘sink’ location for both human-associated and environmental microbes to interact and facilitate horizontal gene transfer activities. Notably, there is also an observed reduction in the multidrug resistance genes detected on the inhaler surface relative to the lung as seen in Fig 3.7. Most of these multi-drug resistance genes encode for either transcriptional regulators or proteins belonging to the multidrug efflux pump²¹⁰. One possible explanation is that microbes in the lungs need to continually maintain these genes to survive in the antibiotic stress environment as chronic respiratory patients are often prescribed with antibiotics. The inhaler, on the other hand, is an antibiotic-free environment and in this case, the microbes may choose to drop these multi-drug resistance genes as there may be increased fitness cost associated with maintaining and expressing them²¹¹.

Conclusions

The dissemination of antibiotic resistance genes (ARGs) should be monitored closely as it is a major clinical concern as bacterial infections become more widespread and increasingly harder to treat with fewer new antibiotics available. To curb the transmission of ARGs, inhalers should be cleaned or disinfected regularly.

There are some limitations to this work. Firstly, the sample size of our study was $n = 85$ individuals which were further subdivided into healthy and diseased groups. Despite the small sample size, we could still see distinct differences in the microbiome in COPD and bronchiectasis patients. A bigger sample size is necessary to infer deeper relationships between the resistome and other specific factors such as exacerbations, ethnicity, and smoking status.

Secondly, although metagenomic shotgun sequencing can give higher taxonomic resolution and functional genomic information, it poses a challenge for real-time diagnostic use in clinical practice. Metagenomics is relatively expensive with slow turnaround times as compared to other sequencing method. In addition, it also requires significant bioinformatics expertise and a high-performance computing facility, which further impedes the adoption of metagenomics in clinical settings.

Chapter 4. Plants as sources of the air microbiome

4.1 Introduction

Leaves constitute a substantial surface area of the phyllosphere, with an estimated total area of combined adaxial (top) and abaxial (bottom) surfaces of 1,017,260,200 km² globally⁴⁹. It has been shown that bacteria of the genera such as *Pantoea*, *Sphingomonas* and *Pseudomonas* and fungi such as *Cryptococcus*, *Cladosporium* and *Botrytis* were highly abundant on leaves were also observed to be enriched in downwind air relative to the upwind air⁷². To investigate the effect of local source emission from plants, swab samples of the adaxial and abaxial leaf surfaces of *Rhapis excelsa* were collected and compared to the nearby air microbiome in tropical Singapore.

The adaxial and abaxial leaf surfaces exhibit distinct anatomical structure and function. The adaxial leaf surface usually has a layer of cuticular wax to prevent water loss and is adapted for efficient light capture during photosynthesis^{212,213}. In contrast, the abaxial surface usually has a higher density of stomata that is adapted for gaseous exchange during respiration. Some plants also evolved epidermal outgrowth of trichomes on their leaves as a form of mechanical defence²¹⁴. As part of the plant's innate immune response, the guard cells surrounding the stoma will close upon detection of microbes⁶¹.

Studies on the phyllosphere microbiome were usually performed sequentially starting from bulk whole-leaf sampling, washing, sonication and filtering, to downstream DNA extraction and amplicon sequencing²¹⁵⁻²¹⁷. By taking bulk whole-leaf samples, this workflow aggregates the total microbial communities found on both adaxial and abaxial leaf surfaces. Amplicon sequencing technique might also give rise to amplification biases and therefore metagenomic shotgun sequencing was used in our study^{218,219}. The major advantage of using metagenomics is the simultaneous detection of both bacteria and fungi without the need for separate 16S, 18S

and internal transcribed spacer (ITS) primers. Hence, swab samples were collected from the adaxial and abaxial surfaces of *R. excelsa* and *C. fruticosa*.

R. excelsa and *C. fruticosa* were chosen as the model plants in this study as they have clusters of broad leaves accessible for swabbing. They are also frequently planted ornamentally along roadsides in Singapore. *R. excelsa* belongs to the *Arecaeae* family and is native of southern China and Southeast Asia²²⁰. *C. fruticosa* belongs to the *Agavaceae* family and is native of tropical Asia, Australia and Pacific Islands, including Hawaii²²¹.

This chapter seeks to examine source-sink relationships between plant leaves and air. Plant leaves have large surface areas and are potential sites of microbial exchange following wind and rain events. A study conducted in California has also reported microbial transfer of up to 50% of plant bacteria in downwind air⁷².

4.2 Materials and methods

Air sample collection

Air samples were collected at the outdoor balcony of N1.3 building in Nanyang Technological University (NTU), Singapore (GPS coordinates N1.346247, E103.679467) during two time periods between 4 December 2016 and 9 December 2016, as well as, between 19 February 2017 and 24 February 2017. Filter-based air samplers, SASS3100 (Research International, USA) were operated at a high flow rate of 300 L/min for 2 hours using the bioaerosol electret filters of diameter 60 mm and 0.5 µm pore size. The air samplers were positioned on a tripod at 1.5 m height. The filters were transported back to the lab and stored at -20 °C.

Leaf swab sample collection

Leaf swab samples were collected at the same time as the air samples. Each sample consisted of swabbing surfaces of ten leaves of the same plant to improve the DNA yield. The leaf

surfaces were swabbed with 4N6Floq swabs (Copan, USA) dipped in Gibco phosphate buffered saline (PBS), pH 7.2 (Life Technologies, USA) with 0.1% Triton-X100 (Sigma-Aldrich, USA). This experiment was conducted to estimate the probability of microbial transfer from leaf surface to nearby air. The complete experimental processing pipeline is illustrated in Fig 4.1.

Fig 4.1 | Flowchart of processing pipeline. A pictorial outline of how samples were collected and processed starting from sampling, DNA extraction, metagenomic shotgun sequencing to bioinformatic data analyses.

Swab collection of surfaces at the balcony

Swabs of the steel door panels, floor, paint coated pillar, clipboard, and the leaf surfaces of *Rhapis excelsa* and *Cordyline fruticosa* were collected for source tracking. The sampling collection was performed at the same location as the air samples at the outdoor balcony of N1.3 building in NTU.

Fig 4.2 | Swab sampling locations at the N1.3 balcony. Swabs were collected from outdoor surfaces such as the steel doors, floor and paint coated pillar. Swab samples were also obtained from the leaves of *Rhapis excelsa* and *Cordyline fruticosa*. A dust sample was collected by placing a clipboard disinfected with 70% ethanol on a table for passive dust collection.

DNA extraction, library preparation and sequencing

After sampling, the swabs were snapped into a DNeasy PowerWater kit (Qiagen, Germany) bead tube and immediately transported to the laboratory for DNA extraction following manufacturer's instructions with the addition of 0.1 mg/ml proteinase K (Sigma-Aldrich, USA) and overnight sonication at 65 °C²²². Extracted DNA samples were quantitated on a Qubit 2.0 fluorometer, using the Qubit dsDNA High Sensitivity Assay Kit (Invitrogen, USA) Sample quantitation was repeated on a Promega QuantiFluor fluorometer, using Invitrogen's Picogreen assay.

Next-generation sequencing libraries were prepared using the Swift Biosciences' Accel-NGS 2S Plus DNA kit, following the manufacturer's instructions. DNA was sheared to ~450 bp on either a Covaris S220 or E220 focused-ultrasonicator (Covaris). All libraries were dual-

barcoded, using Swift Biosciences' 2S Dual Indexing kit. and then sequenced using the HiSeq 2500 sequencer (Illumina, USA).

Taxonomical identification of plants

The taxonomical identification of the plant species at N1.3 balcony was validated using DNA barcoding and anatomical structures^{223,224}. Plant leaves were cut, and DNA was extracted using DNeasy PowerWater kit (Qiagen, Germany) according to manufacturer's protocol. Extracted DNA was then amplified with polymerase chain reaction (PCR) targeting the *rbcL* (RUBISCO large subunit) and *MatK* (Maturase K) genes. The *rbcL* primer sets used were *rbcLa-F* (5'-ATGTCACCACAAACAGAGACTAAAGC-3')²²⁵ and *rbcLa-R* (5'-GTAAAATCAAGTCCACCRCG-3')²²⁶, whereas the *MatK* primer sets used were *MatK-xF* (5'-TAATTTACGATCAATTCATTC-3')²²⁷ and *MatK-MALP* (5'-ACAAGAAAGTCGAAGTAT-3')²²⁸. PCR conditions were as follows: 2× KAPA HiFi Hotstart Ready Mix (12.5 µl) (Kapa Biosystems, USA), primers (for both forward and reverse primers each 0.75 µl of 10 µM), template DNA (1 µl of 5 ng/µl) and H₂O (10 µl) for each reaction (25 µl). PCR cycles for *rbcL* were performed as follows: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 52 °C for 40 s, 60 °C for 1 min with a final elongation step at 60 °C for 5 min while PCR cycles for *MatK* were as follows: 98 °C for 45 s, 35 cycles of 98 °C for 10 s, 54 °C for 30 s, 72 °C for 40 s with a final elongation step at 72 °C for 10 min. The PCR products were sent to AITbiotech Singapore for Sanger sequencing. Read quality was analysed with 4Peaks (Mekentosj, Amsterdam) and Chromas (Technelysium, Australia) where sites below Phred score 20 were filtered and trimmed. Forward and reverse reads were overlapped using BioEdit v7.2.6.1²²⁹. Sequence similarity results correspond to *Rhapis* and *Cordyline* genus using blastn²³⁰ and BOLD²³¹ at 99% identity and 100% query cover respectively. Further taxonomic plant species identification was matched to *Rhapis excelsa* and *Cordyline fruticosa* respectively based on their anatomical structure descriptions^{223,224}.

Adaxial and abaxial leaf swabs

A subsequent follow-up experiment was then performed to investigate if the adaxial (upper) and abaxial (lower) leaf surfaces of *R. excelsa* and *C. fruticosa* had different microbiomes. Leaf swabs were collected across eight locations in NTU on 31 October 2017 at midnight and 22 November 2017 at noon on the map illustrated in Fig 4.2. All plants sampled were planted outdoors. A duplicate swab sample was taken from each site.

Fig 4.2 | Map of leaf swab sampling locations. Map depicting the locations of A to H where *R. excelsa* and *C. fruticosa* leaves represented in green and red squares were swabbed. The relative distances between each location is shown. The global positioning system (GPS) coordinates of the sampling locations are: Site A (1°21'02.7"N 103°40'49.0"E), B (1°21'15.2"N 103°41'18.7"E), C (1°20'45.1"N 103°41'15.0"E), D (1°20'41.0"N 103°41'10.0"E), E (1°20'43.8"N 103°40'46.0"E), F (1°20'36.3"N 103°40'44.7"E), G (1°20'42.5"N 103°40'49.3"E) and H (1°21'11.6"N 103°40'54.8"E).

Bioinformatics and taxonomic assignment

Fastq files after sequencing were examined for quality using FastQC (version 0.11.5)²³². Raw sequence reads with a Phred score of at least 20 and minimum read length of 30 bp with an error probability of less than 0.05 were selected. Adapters were trimmed using Cutadapt (version 1.15)¹⁷⁶. To speed up the computing timing, two million reads were randomly sub-sampled to represent the entire population of reads as rarefaction curve showed that no new species were observed when sampling beyond two million reads. The reads were queried across NCBI non-redundant (nr) protein database using RapSearch2²³³.

All output files including the swab and reagent blank controls were visualised in MetaGenome ANalyzer (MEGAN) v6.12.0¹⁷⁹ in absolute read counts. Fungi and bacteria species-level taxonomy tables were exported as text files for subsequent removal of contaminating taxa that could have originated from the swab itself or the reagent blank using the *decontam*²³⁴ package in R²³⁵. Microbial species with less than 25 read counts were removed. After decontamination, the files were re-imported into MEGAN for comparative metagenomic analyses.

Source tracker and modifications used

Metagenomes of the air and swabs were exported in the biom format and imported into sourcetracker2 for Monte-Carlo simulations with a burn-in parameter of 100 iterations. Probabilities were first calculated by defining air as sink and surfaces as source and vice versa. The Bayesian hypothesis were modified, and results obtained were re-interpreted as follows:

$$\text{Original Source Tracker:} \quad \text{Source} \xleftarrow{P(\text{Source} | \text{Sink})} \text{Sink} \quad (1)$$

$$\text{Reframed Bayesian Hypothesis:} \quad \text{Location B} \xleftarrow{P(B | A)} \text{Location A} \quad (2)$$

$$\text{Location n} \xleftarrow{P(n | A)} \text{Location A} \quad (3)$$

Statistical analyses

Statistical analyses were performed in R using *vegan*²³⁶ and *phyloseq*²³⁷ packages. Beta-diversity was assessed using the Bray-Curtis dissimilarity distance. Principal coordinate analysis (PCoA) plots were constructed to pick up clusters associated with spatiotemporal factors such as the adaxial/abaxial surface, time of the day and locations. Alpha-diversity was scored based on Shannon and Chao-1 indices. The Chao-1 index was used to estimate the number of observed species and the differences between adaxial and abaxial surface were plotted as scatter plots. Mann-Whitney test was then performed in PRISM v7.03²³⁸ to validate if the differences observed were statistically significant.

Co-occurrence network graph of microbial species was constructed with Gephi v0.9.2²³⁹ with a Spearman rank correlation cut-off of 0.8 and p-value of 0.01 adapted from a R script written by Li, et al.²⁴⁰.

Processing of leaf samples for scanning electron microscopy (SEM)

The leaves were fixed in 2.5% glutaraldehyde in phosphate buffer. Samples were washed gently with distilled water, placed onto an aluminum stub sample mount using double sided carbon tape and snap frozen in liquid nitrogen. After metal block cooling, equilibrated metal blocks were freeze-dried. The metal block with samples was immediately transferred to Leica SCD050 (Leica, Germany) sputter coater stage lined with a layer of insulating foam and the chamber was promptly evacuated. Samples were slowly freeze dried in the evacuated chamber overnight, then removed and sputter coated with 4 nm layer of platinum. Scanning electron microscopy (SEM) analysis was then performed using a Jeol JSM-5690LV SEM.

Reactive oxygen species (ROS) assay

Swabs were collected using the same procedure as the metagenomic samples but only with PBS and not Triton-X100 as surfactants may interfere with the cellular ROS activity²⁴¹. The

swab tips were resuspended in 500 µl of PBS in 1.5 ml amber tubes and samples were kept on ice and processed on the same day of sample collection. An *in vivo* ROS assay protocol developed by Nemzer, et al.²⁴² was modified and adapted for the leaf swabs. Samples were briefly vortexed and centrifuged at 10,000 rpm at 4 °C for 3 min. 240 µl of supernatant were transferred into a clean Eppendorf tube and 720 µl of methanol (Fisher Scientific, USA) was added to supernatant to precipitate proteins. Samples were briefly vortexed for another 30 s and centrifuged at 10,000 rpm at 4 °C for 5 min. Supernatants were transferred into clean tubes and the methanol was evaporated to dryness using Eppendorf Concentrator plus (Eppendorf, Germany). The evaporated samples were reconstituted in 60 µl PBS. 40 µl of the reconstituted samples was loaded into each well of a black flat bottom 96 well plate (Nunc, Denmark) and 60 µl of dihydrorhodamine 123 dye (Invitrogen, USA) diluted with PBS to a working concentration of 202 µM with PBS were added to the samples. The fluorescence intensity was measured using Synergy H1 microplate reader (BioTek, USA) with an excitation wavelength of 485 nm and emission wavelength of 545 nm after 5 minutes. Bar charts were plotted in PRISM v7.03 for visualisation.

Surface water contact angle measurement

The sessile drop method was performed at room temperature using the OCA-20 contact angle meter (DataPhysics, USA) to measure the surface wettability of the leaves. The adaxial and abaxial leaves of both plants were cut into 25 mm² squares and mounted onto microscope slides before placing on the stage. 3 µl of a water droplet was dispensed and suspended on the needle tip of the robotic syringe. The stage was then manually adjusted upwards until the specimen was in contact with the water droplet. The stage was then lowered slightly, and the water droplet-on-leaf image was captured. Using the SCA software (DataPhysics, USA), a base line was drawn, and an ellipse was fitted over the water droplet for calculation of the contact angle.

4.3 Results

Plant leaves contribute to less than 5% of the tropical outdoor air microbiome

The Singapore outdoor air microbiome is largely dominated by fungi, and it is hypothesized that the airborne fungi were contributed by outdoor plants at the balcony. Using the original Bayesian source tracker model, only 4% of the outdoor air microbiome was found to have originated from leaf surfaces within 50 metres of the air sampling site (Fig 4.4a). A large fraction of the air microbiome was mapped to unknown sources. Results suggest that airborne microbes could have originated from either the abiotic balcony surfaces or from remote sites such as the forest adjacent to NTU or even as far as Malaysia.

A closer look at the microbiome profiles of both plant leaves in Fig 4.4b reveal that their microbiomes were largely similar and were dominated by leaf-inhabiting Ascomycetes, with varying proportions of some species such as *Aureobasidium melanogenum* and *Jaminalia angkorensis* that may prefer colonising the *Cordyline* leaves over the *Rhapis* leaves.

To tackle this question on the relative contributions of local and remote sources, the abiotic balcony surfaces such as the steel pillar, doors, floor, and dust accumulated on a sterilised clipboard were swabbed. The floor and door surfaces were constantly exposed to wet and dry outdoor weather conditions with the presence of cyanobacteria. A sterilised clipboard was also placed on a table for dust sampling where the corresponding dust microbiome profile showed an even accumulation of microbes with microbial community members like those found in the floor and door samples. The air, however, had more Ascomycota and Basidiomycota fungi which were not found in any of the localised abiotic surfaces and were likely to have drifted across long distances, originating from remote sources.

Fig 4.4 | Relative contributions of nearby vegetation to the air microbiome. (a) An estimate of about 4% of the outdoor air microbiome originated from the plant leaves within 50 metres of the air sampling site whereas the vast majority were mapped to unknown sources. (b) Both plant leaves share similar microbes in the list of the top 40 most abundant species, which consist of mainly leaf-inhabiting Ascomycetes, with notable varying proportions of some species.

Fig 4.5 | Biotic and abiotic surface microbiomes at the outdoor balcony versus air. *Rhapis* and *Cordyline* plants each have their own microbiome, comprising of mainly leaf-inhabiting ascomycetes. The pillar was located at 1 metre away from the *Cordyline* plant and high similarities in the microbiomes were observed between the pillar and plant. The floor and door surfaces were exposed to sun and rain, where higher relative abundance of cyanobacteria was observed. The dust sample accumulated on a sterilised clipboard showed a relatively even accumulation of microbes settling on the surface. In contrast, the air consisted of more Ascomycota and Basidiomycota fungi.

The modified Bayesian source tracking model was used to find out the source emissions from the balcony surfaces and plant leaves to the outdoor air. Plant leaves emit less than 5% of their microbiomes to the outdoor air (Fig 4.6). The floor, steel door and steel pillar each release less than 1% of microbes to the outdoor air. The microbiome of the dust-settled surface on the

clipboard bears more resemblance to the air microbiome and has a probability of 10% of source emission to the air. This is probably driven by wind stimulating the resuspension of dust.

Fig 4.6 | Source tracking investigation performed at the N1.3 balcony. Microbes on the plant leaf surface has a probability of source emissions of less than 5% to the outdoor air. The floor, steel door and the steel pillar each exhibits transfer probability of about 1% or less. The dust settled surface on the clipboard shows the highest source emission probability of 10%.

Significant reduction in bacterial species observed on abaxial leaf surfaces

Next, the microbiomes of the adaxial and abaxial leaf surfaces of both plants were investigated. There was a significant decrease in bacterial read counts observed on the abaxial leaf surface of both plants at different sampling locations and two different times of the day (Fig 4.7). Bacterial read counts were in the range of 8,000 to 50,000 counts on the adaxial surface, in contrast to the abaxial surface that were in the range of 100 to 3,000 reads.

a. *Rhapis excelsa*

b. *Cordyline fruticosa*

Fig 4.7 | SEM images and bubble plot of microbes, plants, and arthropod-related reads on the leaf surfaces. Reduction in bacterial read counts was evident on the abaxial surface relative to Fungi, Viridiplantae and Arthropoda reads on (a) *R. excelsa* and (b) *C. fruticosa*. The moon and sun symbols represent the time of sampling at midnight and noon, the letters refer to the sampling sites that were shown in Fig 4.2, while the bubble sizes correspond to the relative abundance of reads assigned to groups of taxa.

Bacterial reduction becomes more evident when samples were plotted as dot plots with the species richness estimate indicator, Chao-1 index in Fig 4.8. The differences in bacterial species richness between the adaxial and abaxial surface of both plants were tested to be statistically significant using the Mann Whitney test (Table 4.1).

Fig 4.8 | Dot plot of species richness expressed as Chao-1 index on the adaxial and abaxial leaf surface. An estimate of 33.4-fold fewer bacterial species observed on the abaxial leaf surface of (a) *R. excelsa* and 18.5-fold reduction in (b) *C. fruticosa*. There was only a modest 1.1-fold decrease in fungal species observed on (c) *R. excelsa* and 18.5-fold decrease on (d) *C. fruticosa* abaxial surface.

Table 4.1 | Mann Whitney test of Chao-1 index between adaxial and abaxial surface.

Using Hodges-Lehmann to estimate the differences in median at 92 and 137 bacterial species for *R. excelsa* and *C. fruticosa* respectively. The reduction was statistically significant. Decrease in fungal species was also observed to be significant but to a smaller extent at 10 and 34 species.

Mann Whitney test	Bacteria		Fungi	
Plant Species	<i>R. excelsa</i>	<i>C. fruticosa</i>	<i>R. excelsa</i>	<i>C. fruticosa</i>
<i>p</i> -value	<0.0001	<0.0001	0.0191	<0.0001
Median of Adaxial, n=20	93	147	82.5	75
Median of Abaxial, n=20	1	6	71.5	34.5
Difference: Actual	-92	-141	-11	-40.5
Difference: Hodges-Lehmann	-92	-137	-10	-34

Co-occurrence network reveals nine clusters of microbial species

Nine co-occurrence clusters were obtained from Gephi using the modularity class algorithm with parameters of Spearman rank correlation coefficient of 0.8 and *p*-value of 0.01 (Fig 5.9). Clusters A and B have highly interconnected nodes, where cluster A is largely comprised of soil and leaf bacteria whereas cluster B is a collection of wood-rotting fungi, photosynthetic, thermophilic, and radiation-tolerant bacteria. Further in-depth analyses showed that the species within these two clusters were present on the adaxial surface but were significantly reduced on the abaxial surface (Figure 4.10 and 4.11). The seven remaining clusters were classified according to commonalities in phylogeny and phytopathogenic characteristics. For instance, cluster E comprises of fungi that cause leaf blotch disease²⁴³⁻²⁴⁵, cluster H consists of a group of saprotrophic fungi that belong to the Herpotrichiellaceae family, living on rotting wood and leaves²⁴⁶. Cluster C is made up of a group of smut fungi from the Ustilaginaceae family²⁴⁷, whereas cluster G consists of *Aurebasidium spp* that are indigenous microbes of the leaf niche habitat^{248,249}. Cluster F is a group of endophytic fungi²⁵⁰⁻²⁵² and clusters I and D are agricultural pathogens of apple, wheat, brassica and tomatoes²⁵³⁻²⁵⁵.

a. Co-occurrence Clusters

Fig 4.9 | Co-occurrence network analysis of microorganisms on the adaxial and abaxial leaf surfaces. Of the nine clusters, clusters A and B were being identified as the biggest clusters with the most highly interconnected nodes. Cluster A is comprised of soil and leaf bacteria whereas cluster B is composed of wood-rotting fungi, photosynthetic, thermophilic, and radiation-tolerant bacteria. The network graph was constructed using cut-off parameters of Spearman rank correlation coefficient of 0.8 and *p*-value of 0.01.

Cluster A

Cluster A

Fig 4.10 | Relative abundance of cluster A species. Cluster A comprises of only bacterial species. Significant reduction in the relative abundances was observed on the abaxial leaf surface in contrast to the adaxial surface.

a. *Rhapis excelsa*

Cluster E fungi were twice as abundant on the abaxial leaf surfaces. Comparing the read counts of Bacteria, Ascomycota, and Basidiomycota among nine clusters, the leaf blotch fungi from cluster E dominated both the adaxial and abaxial leaf surfaces (Fig 4.12).

Fig 4.12 | Read counts of microbes from each co-occurrence cluster. Among the nine

clusters A to I, cluster E consists of leaf blotch fungi that dominate the leaf surfaces of (a) *Rhapis excelsa* and (b) *Cordyline fruticosa*. *R. excelsa* has about twice the number of reads of cluster E fungi on the abaxial surface as compared to *C. fruticosa*.

It is likely that these Ascomycete fungi from cluster E prefer to colonise the abaxial surface where there are higher densities of stomata as these stomatal openings could be vulnerable ports of entry into the host plant. From Fig 4.12, clusters A and B microbes were present mainly on the adaxial surface and the highly interconnected network of nodes from these clusters suggests the presence of a microbial community on the adaxial surface. Additionally, Fig 4.4b showed that the two plants harbour different microbial cloud with distinct members in the microbial community. Likewise, the same trend can be observed in Fig 4.12, where *C. fruticosa* have more microbes in clusters G, H, and I.

Leaf surface chemical and physical properties

Swabs of the adaxial and abaxial leaf surfaces from both plants were assessed for both chemical and physical properties. The leaf swabs were tested for reactive oxygen species (ROS). The leaf surface structures were also examined under the scanning electron microscope (SEM) and the hydrophobicity of the leaves was measured using a contact angle meter.

The abaxial leaf surface of *R. excelsa* showed significantly elevated levels of ROS activity across all sampling sites (Fig 4.13). However, ROS activity was absent on both adaxial and abaxial surface of *C. fruticosa*. The SEM images further validated that the abaxial leaf surfaces have higher densities of stomata than the adaxial side. For *R. excelsa*, there was no structural difference between the adaxial and abaxial surface apart from the stomata. Leaves of *C. fruticosa*, on the other hand, have distinctively patterned abaxial surface with hierarchical structures of micrometre-scale wrinkles and nanometre-sized scales (Fig 4.14).

a. *Rhapis excelsa*

b. *Cordyline fruticosa*

Legend:

■ Adaxial ■ Abaxial ■ Control

Fig 4.13 | Reactive oxygen species detection on the leaf surface. Significantly elevated levels of ROS activity were detected on the abaxial leaf surface of *R. excelsa* at all sampling sites from A to E. The ROS activity, however, was absent on both adaxial and abaxial leaf surfaces of *C. fruticosa*.

Surface wettability of the leaves was measured using a water contact angle meter. Results in Fig 4.15 showed that the adaxial leaf surface of *R. excelsa* has a contact angle of 83° and is likely hydrophilic, whereas the abaxial surface is 95° and is slightly hydrophobic. Conversely, *C. fruticosa* adaxial leaf surface has a contact angle of 102° and is slightly hydrophobic, while its abaxial leaf surface has a contact angle of 140°, which is very hydrophobic. Based on the SEM images in Fig 4.14, the surface hierarchical structures make *C. fruticosa* leaves more hydrophobic. The hydrophobic effect was further observed by dropping a 10 µl water droplet and spreading it on each leaf surface with a micropipette tip. Only the water droplet on the abaxial surface of *C. fruticosa* stayed spherical and glided across the surface.

Fig 4.14 | Levels of reactive oxygen species and structural differences of the leaf surfaces.

More than 2-fold increase in ROS activity was detected on the abaxial leaf surface of *R. excelsa* but not in *C. fruticosa*. Based on the SEM images, both plants have higher densities of stomata on the abaxial surface. Interestingly, *C. fruticosa* have hierarchical structures with micrometer-scaled wrinkles and nanometer-sized scales unlike *R. excelsa*. This surface patterned structure makes the abaxial leaf surface of *C. fruticosa* more hydrophobic.

Fig 4.15 | Surface wettability of leaf surfaces. The adaxial and abaxial leaf surfaces of (a) *R. excelsa* have contact angles of 83° and 95°, respectively. The adaxial and abaxial leaf surfaces of (b) *C. fruticosa* have contact angles of 102° and 140°, respectively. *C. fruticosa* has an abaxial leaf surface that is hydrophobic and repels water from its surface.

4.4 Discussion

Source tracking at N1.3 balcony in tropical Singapore showed that plant leaves within 50 metres contribute to only less than 5% of the outdoor air. This finding conflicts a study conducted in California where they found up to 50% of airborne bacteria in downwind samples originating from plants within 50 metres upstream⁷². It can be probably explained by the low average wind speed of less than 2.5 m/s and the hot humid climate in Singapore unlike the coastal areas in California.

The outdoor building surfaces such as the floor, steel door and steel pillar were estimated to have source emission rates of less than 1% to the outdoor air microbiome. These surfaces were exposed to hot sunny and cool rainy weather events causing higher fluctuations of surface

temperatures and were largely colonised by cyanobacteria as seen in Fig 4.5. Cyanobacteria are photosynthetic microbes, and they only need light and water to survive. This observation was also seen in another study that found cyanobacteria and algae on historic stone building surfaces and they are hardy microbes that can withstand tough weather conditions²⁵⁶. The pillar was located within 1 metre distance from the *Cordyline fruticosa* plant, and the pillar was observed to have a very similar microbiome to the plant and is likely a consequence of the plant shedding its microbiome to its surroundings (Fig 4.5).

Interestingly, the abaxial leaf surface was found to have fewer bacteria than the adaxial leaf surface. Preliminary experiments suggest that *R. excelsa* deploys a chemical defence approach by producing ROS on its abaxial leaf surface and could potentially be associated with the pathogen-associated molecular patterns (PAMPs)-trigger immunity (PTI) mechanism. PTI is well-described in *Arabidopsis thaliana* where the plant senses the presence of bacteria through its *fls2* flagellin receptor²⁵⁷. Upon bacteria detection, the plant produces ethylene, ROS and fortifies its cell walls to inhibit bacterial colonization. Despite the higher levels of ROS detected on *Rhapis excelsa* leaves, fungal species that are associated with leaf blotch diseases still persist. *Aureobasidium* causes brown spots on leaves²⁵⁸, *Dothistroma* causes red band on leaves²⁵⁹, and *Pseudocercospora* causes Sigatoka disease in banana leaves²⁶⁰. *Pseudocercospora fijiensis* first colonizes the leaf as an epiphyte and then extend its hyphae, penetrating through the stomata and into the leaf tissue²⁶¹. A separate study has also shown that *Pseudocercospora* have hyphae that can secrete catalase and superoxide dismutase, that allow it to survive the oxidative burst response²⁶². There are also other fungi like *Zymoseptoria tritici* that release more ROS and use the plant's ROS response to their own advantage by inducing plant cell necrosis and triggering the onset of infection²⁶³. It is likely that leaf blotch fungi are more resistant to ROS than bacteria and have adaptations to survive the stressful conditions on abaxial surface.

C. fruticosa may have evolved a different defence strategy by having hydrophobic abaxial leaf surface adaptations with low surface energy. This surface structure prevents microbes from adhering and limits the water available for bacterial growth²⁶⁴. The hydrophobic defence strategy seems to be more effective in minimising microbial colonisation than the ROS defence as the relative abundance of leaf blotch fungi in *C. fruticosa* abaxial surface was observed to be halved of *R. excelsa* (Fig 4.12). It appears that these leaf blotch fungi are persistent colonisers even in environment with low water availability.

4.5 Conclusions

In summary, plant leaves within 50 metres of the outdoor air releases only up to 5% of its microbiome to the outdoor air in tropical Singapore. Interestingly, the abaxial leaf surface has a lower microbial diversity as compared to the adaxial surface. It is likely a result of plant-microbe interactions which could be explained by two hypothetical mechanisms, namely reactive oxidative burst response and hydrophobic self-cleaning surface structures. However, further experiments are needed to draw conclusions on the underlying molecular mechanism and function for this observation.

Since the *Cordyline fruticosa* leaf has near superhydrophobic and self-cleaning properties due to its unique hierarchical micrometre wrinkles and nanometre trichomes structure, it may be of interest to produce a biomimetic surface with similar properties. A nanoimprint capture of the leaf was performed using both epoxy and polydimethylsiloxane (PDMS) resin with 10:1 resin:hardener ratio at room temperature. The leaf was then peeled off from the cured resin. SEM images of the nanoimprinted leaf surface were captured at 250x and 1000x magnification (Fig 4.16 and 4.17). However, during the initial attempts, only the microstructure was captured in the imprint. The nano-scaled features such as the scales and trichomes were too fine and were not captured by the resin. Future leaf replicas can be made, probably by varying the ratio of the

resin and hardener or exploring other means of nanofabrication such as femtosecond laser ablation.

Fig 4.16 | SEM image of the nanoimprinted abaxial leaf surface at 250x magnification.

The structure was successfully captured at the micrometre scale on the PDMS resin.

Fig 4.17 | SEM image of the nanoimprinted abaxial leaf surface at 1000x magnification.

Higher magnification reveals that the nanoimprinted PDMS resin, however, did not capture the shapes of the nanometre trichomes in the replica.

Chapter 5. Conclusions and future perspectives

Source identification and mapping of bioaerosol fluxes are the crucial first steps towards understanding of the transport processes of airborne microorganisms. This will help identify the source of the problem, limit the spread of pathogens, and calls for efforts to disinfect contaminated surfaces, or fumigation using chemical disinfectants or ultraviolet radiation. This thesis adds to the current body of bioaerosols studies by proposing a modification of the Bayesian source tracker tool, where results can be visualised in Markov Chain transition graphs. The study of sources and sinks of the air microbiome is a highly complex problem that requires an interdisciplinary solution. Potential sources were identified in the built environment by harnessing the combined strengths of metagenomics, probabilistic and computational fluid dynamics models.

Based on the findings of this thesis, the likely reservoirs of indoor air microbes are cellulosic, gypsum and textile materials. These materials are prone to microbial contamination especially in moist environments. Secondly, the inhaler has shown to be an environmental sink for antibiotic-resistant microbes via contact transfer and breath of chronic respiratory patients. Lastly, leaf surfaces are minor source contributors to the air microbiome in the tropics as plant leaves only release 5% of the microbes to the outdoor air.

Metagenomics is a powerful tool for profiling the relative abundances of microbes in the environment. However, metagenomics relies on the detection of environmental DNA, and it does not distinguish between live and dead cells. To make full assessments on the viable fraction of airborne microbes, future experiments can be supplemented with adenosine triphosphate (ATP) measurements and meta-transcriptomics analyses. Metagenomics and microbial source-tracking if widely adopted has the potential to disrupt and revolutionise the cleaning industry. However, at this moment, the technology is too expensive and slow. It costs about S\$200 to sequence a sample and the turnaround time from sample to answer takes at least

a week. Therefore, it is unsustainable for businesses to offer metagenomic source tracking services as a solution. Customers want their homes or offices disinfected in a cheap and fast manner. An attractive and low-cost alternative solution is to use quantitative PCR for real time detection of a panel of pathogens such as fungi, bacteria, and viruses. However, this requires optimisation and design of primers and probes. There is also a challenge in multiplex PCR as the number of fluorophores and detector channel of the instrument can be a limiting factor on the panel size.

Future source tracking experiments could be performed by nebulising microbes from a synthetic source within a chamber or wind tunnel. This should also be coupled with environmental measurements such as temperature and relative humidity of air, as well as the temperature and equilibrium relative humidity of surfaces. Higher-level statistical analyses can be conducted to correlate these environmental factors with the artificial source emission experiments under varying induced environmental conditions.

References

- 1 Matthias-Maser, S. & Jaenicke, R. The size distribution of primary biological aerosol particles with radii $> 0.2 \mu\text{m}$ in an urban/rural influenced region. *Atmospheric Research* **39**, 279-286 (1995).
- 2 Kim, K.-H., Kabir, E. & Kabir, S. A review on the human health impact of airborne particulate matter. *Environment international* **74**, 136-143 (2015).
- 3 Stetzenbach, L. Airborne Infectious Microorganisms. *Encyclopedia of Microbiology*, 175 (2009).
- 4 Burrows, S. M., Elbert, W., Lawrence, M. & Pöschl, U. Bacteria in the global atmosphere—Part 1: Review and synthesis of literature data for different ecosystems. *Atmospheric Chemistry and Physics* **9**, 9263-9280 (2009).
- 5 Rieux, A. *et al.* Long-Distance Wind-Dispersal of Spores in a Fungal Plant Pathogen: Estimation of Anisotropic Dispersal Kernels from an Extensive Field Experiment. *PLOS ONE* **9**, e103225, doi:10.1371/journal.pone.0103225 (2014).
- 6 Li, H., Karney, G., O'Toole, D. & Crawford, T. B. Long distance spread of malignant catarrhal fever virus from feedlot lambs to ranch bison. *The Canadian Veterinary Journal* **49**, 183 (2008).
- 7 Yu, I. T. *et al.* Evidence of airborne transmission of the severe acute respiratory syndrome virus. *New England Journal of Medicine* **350**, 1731-1739 (2004).
- 8 Cole, C. R. & McCullough, K. A. The Mold Crisis. *Journal of Insurance Regulation* **23** (2005).
- 9 Ma, J. *et al.* COVID-19 patients in earlier stages exhaled millions of SARS-CoV-2 per hour. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* (2020).

- 10 Xie, W. *et al.* The source and transport of bioaerosols in the air: A review. *Frontiers of Environmental Science & Engineering* **15**, 1-19 (2021).
- 11 Mayol, E. *et al.* Long-range transport of airborne microbes over the global tropical and subtropical ocean. *Nature Communications* **8**, 201, doi:10.1038/s41467-017-00110-9 (2017).
- 12 Lighthart, B. Mini-review of the concentration variations found in the alfresco atmospheric bacterial populations. *Aerobiologia* **16**, 7-16, doi:10.1023/A:1007694618888 (2000).
- 13 Whitman, W. B., Coleman, D. C. & Wiebe, W. J. Prokaryotes: The unseen majority. *Proceedings of the National Academy of Sciences* **95**, 6578, doi:10.1073/pnas.95.12.6578 (1998).
- 14 Behzad, H., Gojobori, T. & Mineta, K. Challenges and Opportunities of Airborne Metagenomics. *Genome Biology and Evolution* **7**, 1216-1226, doi:10.1093/gbe/evv064 (2015).
- 15 Mbareche, H., Veillette, M., Bilodeau, G. J. & Duchaine, C. Bioaerosol sampler choice should consider efficiency and ability of samplers to cover microbial diversity. *Applied and environmental microbiology* **84**, e01589-01518 (2018).
- 16 Dybwad, M., Skogan, G. & Blatny, J. M. Comparative testing and evaluation of nine different air samplers: end-to-end sampling efficiencies as specific performance measurements for bioaerosol applications. *Aerosol Science and Technology* **48**, 282-295 (2014).
- 17 Gómez-Domenech, M. *et al.* Evaluation of the efficiency of the Coriolis air sampler for pollen detection in South Europe. *Aerobiologia* **26**, 149-155 (2010).
- 18 Langer, V., Hartmann, G., Niessner, R. & Seidel, M. Rapid quantification of bioaerosols containing *L. pneumophila* by Coriolis® μ air sampler and

- chemiluminescence antibody microarrays. *Journal of aerosol science* **48**, 46-55 (2012).
- 19 Tseng, C.-C. & Li, C.-S. Collection efficiencies of aerosol samplers for virus-containing aerosols. *Journal of Aerosol Science* **36**, 593-607 (2005).
 - 20 Yao, M. Bioaerosol: A bridge and opportunity for many scientific research fields. *Journal of Aerosol Science* **115**, 108-112, doi:<https://doi.org/10.1016/j.jaerosci.2017.07.010> (2018).
 - 21 Staley, J. T. & Konopka, A. MEASUREMENT OF IN SITU ACTIVITIES OF NONPHOTOSYNTHETIC MICROORGANISMS IN AQUATIC AND TERRESTRIAL HABITATS. *Annual Review of Microbiology* **39**, 321-346, doi:10.1146/annurev.mi.39.100185.001541 (1985).
 - 22 Claesson, M. J. *et al.* Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions. *Nucleic acids research* **38**, e200, doi:10.1093/nar/gkq873 (2010).
 - 23 Handelsman, J. Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**, 669-685, doi:10.1128/MMBR.68.4.669-685.2004 (2004).
 - 24 Emerson, J. B. *et al.* Schrödinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome* **5**, 86, doi:10.1186/s40168-017-0285-3 (2017).
 - 25 National Academies of Sciences, E., and Medicine *et al.* *Microbiomes of the Built Environment: A Research Agenda for Indoor Microbiology, Human Health, and Buildings*. (National Academies Press (US), 2017).
 - 26 Chenari, B., Dias Carrilho, J. & Gameiro da Silva, M. Towards sustainable, energy-efficient and healthy ventilation strategies in buildings: A review. *Renewable and*

- Sustainable Energy Reviews* **59**, 1426-1447,
doi:<https://doi.org/10.1016/j.rser.2016.01.074> (2016).
- 27 Liu, D.-L. & Nazaroff, W. W. Modeling pollutant penetration across building envelopes. *Atmospheric Environment* **35**, 4451-4462,
doi:[https://doi.org/10.1016/S1352-2310\(01\)00218-7](https://doi.org/10.1016/S1352-2310(01)00218-7) (2001).
 - 28 American Society of Heating, R. a. A.-C. E. ASHRAE Standard : Standards for Natural and Mechanical Ventilation. (2009).
 - 29 Acerbi, E. *et al.* Ecological succession of the microbial communities of an air-conditioning cooling coil in the tropics. *Indoor air* **27**, 345-353 (2017).
 - 30 Bennett, D. H. *et al.* Ventilation, temperature, and HVAC characteristics in small and medium commercial buildings in California. *Indoor Air* **22**, 309-320,
doi:<https://doi.org/10.1111/j.1600-0668.2012.00767.x> (2012).
 - 31 Li, S., Yang, Z., Hu, D., Cao, L. & He, Q. Understanding building-occupant-microbiome interactions toward healthy built environments: A review. *Frontiers of environmental science & engineering* **15**, 1-18 (2021).
 - 32 Luckey, T. D. (Oxford University Press, 1972).
 - 33 Meadow, J. F. *et al.* Humans differ in their personal microbial cloud. *PeerJ* **3**, e1258 (2015).
 - 34 Hospodsky, D. *et al.* Human occupancy as a source of indoor airborne bacteria. *PloS one* **7**, e34867 (2012).
 - 35 Flores, G. E. *et al.* Microbial Biogeography of Public Restroom Surfaces. *PLOS ONE* **6**, e28132, doi:10.1371/journal.pone.0028132 (2011).
 - 36 Kembel, S. W. *et al.* Architectural design influences the diversity and structure of the built environment microbiome. *The ISME journal* **6**, 1469-1479 (2012).

- 37 Lindsley, W. G. *et al.* Measurements of airborne influenza virus in aerosol particles from human coughs. *PloS one* **5**, e15100 (2010).
- 38 Vejerano, E. P. & Marr, L. C. Physico-chemical characteristics of evaporating respiratory fluid droplets. *Journal of The Royal Society Interface* **15**, 20170939 (2018).
- 39 Van Doremalen, N. *et al.* Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *New England journal of medicine* **382**, 1564-1567 (2020).
- 40 Dickson, R. P., Erb-Downward, J. R. & Huffnagle, G. B. The role of the bacterial microbiome in lung disease. *Expert review of respiratory medicine* **7**, 245-257 (2013).
- 41 Segal, L. N. *et al.* Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation. *Microbiome* **1**, 1-12 (2013).
- 42 Charlson, E. S. *et al.* Topographical continuity of bacterial populations in the healthy human respiratory tract. *American journal of respiratory and critical care medicine* **184**, 957-963 (2011).
- 43 Moffatt, M. F. & Cookson, W. O. The lung microbiome in health and disease. *Clinical Medicine* **17**, 525 (2017).
- 44 Buist, A. Similarities and differences between asthma and chronic obstructive pulmonary disease: treatment and early outcomes. *European respiratory journal* **21**, 30s-35s (2003).
- 45 Athanazio, R. Airway disease: similarities and differences between asthma, COPD and bronchiectasis. *Clinics* **67**, 1335-1343 (2012).
- 46 Corrigan, C. & Kay, A. CD4 T-lymphocyte activation in acute severe asthma. *Am Rev Respir Dis* **141**, 970 (1990).
- 47 O'Shaughnessy, T. C., Ansari, T. W., Barnes, N. C. & Jeffery, P. K. Inflammation in bronchial biopsies of subjects with chronic bronchitis: inverse relationship of CD8+ T

- lymphocytes with FEV1. *American journal of respiratory and critical care medicine* **155**, 852-857 (1997).
- 48 George, L. & Brightling, C. E. Eosinophilic airway inflammation: role in asthma and chronic obstructive pulmonary disease. *Therapeutic advances in chronic disease* **7**, 34-51 (2016).
- 49 Vorholt, J. A. Microbial life in the phyllosphere. *Nature Reviews Microbiology* **10**, 828-840, doi:10.1038/nrmicro2910 (2012).
- 50 Kamo, T. *et al.* Methylobamine, a UVA-absorbing compound from the plant-associated bacteria *Methylobacterium* sp. *Natural Product Communications* **13**, 1934578X1801300208 (2018).
- 51 Morris, C. E. & Kinkel, L. Fifty years of phyllosphere microbiology: significant contributions to research in related fields. *Phyllosphere microbiology*, 365-375 (2002).
- 52 Lindow, S. E. & Brandl, M. T. Microbiology of the phyllosphere. *Applied and environmental microbiology* **69**, 1875-1883 (2003).
- 53 Schönherr, J. Characterization of aqueous pores in plant cuticles and permeation of ionic solutes. *Journal of Experimental Botany* **57**, 2471-2491 (2006).
- 54 Vacher, C. *et al.* The phyllosphere: microbial jungle at the plant–climate interface. *Annual review of ecology, evolution, and systematics* **47**, 1-24 (2016).
- 55 Burch, A. Y., Zeisler, V., Yokota, K., Schreiber, L. & Lindow, S. E. The hygroscopic biosurfactant syringafactin produced by *Pseudomonas syringae* enhances fitness on leaf surfaces during fluctuating humidity. *Environmental microbiology* **16**, 2086-2098 (2014).
- 56 Schreiber, L. *et al.* Plant–microbe interactions: identification of epiphytic bacteria and their ability to alter leaf surface permeability. *New Phytologist* **166**, 589-594 (2005).

- 57 Burch, A. Y., Browne, P. J., Dunlap, C. A., Price, N. P. & Lindow, S. E. Comparison of biosurfactant detection methods reveals hydrophobic surfactants and contact-regulated production. *Environmental microbiology* **13**, 2681-2691 (2011).
- 58 Fall, R. & Benson, A. A. Leaf methanol—the simplest natural product from plants. *Trends in Plant Science* **1**, 296-301 (1996).
- 59 Sy, A., Timmers, A. C., Knief, C. & Vorholt, J. A. Methylophilic metabolism is advantageous for *Methylobacterium extorquens* during colonization of *Medicago truncatula* under competitive conditions. *Applied and environmental microbiology* **71**, 7245-7252 (2005).
- 60 Zeng, W., Melotto, M. & He, S. Y. Plant stomata: a checkpoint of host immunity and pathogen virulence. *Current opinion in biotechnology* **21**, 599-603 (2010).
- 61 Melotto, M., Underwood, W., Koczan, J., Nomura, K. & He, S. Y. Plant stomata function in innate immunity against bacterial invasion. *Cell* **126**, 969-980 (2006).
- 62 Arnaud, D. & Hwang, I. A sophisticated network of signaling pathways regulates stomatal defenses to bacterial pathogens. *Molecular plant* **8**, 566-581 (2015).
- 63 Xin, X.-F., Kvitko, B. & He, S. Y. *Pseudomonas syringae*: what it takes to be a pathogen. *Nature Reviews Microbiology* **16**, 316 (2018).
- 64 Bigeard, J., Colcombet, J. & Hirt, H. Signaling mechanisms in pattern-triggered immunity (PTI). *Molecular plant* **8**, 521-539 (2015).
- 65 Bi, G. & Zhou, J.-M. MAP kinase signaling pathways: A hub of plant-microbe interactions. *Cell host & microbe* **21**, 270-273 (2017).
- 66 Mersmann, S., Bourdais, G., Rietz, S. & Robatzek, S. Ethylene signaling regulates accumulation of the FLS2 receptor and is required for the oxidative burst contributing to plant immunity. *Plant physiology* **154**, 391-400, doi:10.1104/pp.110.154567 (2010).

- 67 Guan, R. *et al.* Multilayered regulation of ethylene induction plays a positive role in Arabidopsis resistance against *Pseudomonas syringae*. *Plant physiology* **169**, 299-312 (2015).
- 68 Helliwell, E. E., Wang, Q. & Yang, Y. Ethylene biosynthesis and signaling is required for rice immune response and basal resistance against *Magnaporthe oryzae* infection. *Molecular Plant-Microbe Interactions* **29**, 831-843 (2016).
- 69 Bodenhausen, N., Bortfeld-Miller, M., Ackermann, M. & Vorholt, J. A. A synthetic community approach reveals plant genotypes affecting the phyllosphere microbiota. *PLoS Genet* **10**, e1004283 (2014).
- 70 Elmore, J. M., Lin, Z.-J. D. & Coaker, G. Plant NB-LRR signaling: upstreams and downstreams. *Current opinion in plant biology* **14**, 365-371 (2011).
- 71 Peng, Y., van Wersch, R. & Zhang, Y. Convergent and divergent signaling in PAMP-triggered immunity and effector-triggered immunity. *Molecular plant-microbe interactions* **31**, 403-409 (2018).
- 72 Lymperopoulou, D. S., Adams, R. I. & Lindow, S. E. Contribution of Vegetation to the Microbial Composition of Nearby Outdoor Air. *Applied and environmental microbiology* **82**, 3822-3833, doi:10.1128/AEM.00610-16 (2016).
- 73 Mahnert, A., Moissl-Eichinger, C. & Berg, G. Microbiome interplay: plants alter microbial abundance and diversity within the built environment. *Frontiers in Microbiology* **6**, doi:10.3389/fmicb.2015.00887 (2015).
- 74 Uetake, J. *et al.* Seasonal changes of airborne bacterial communities over Tokyo and influence of local meteorology. *Frontiers in microbiology* **10**, 1572 (2019).
- 75 Fan, H. *et al.* Time-Dependent Size-Resolved Bacterial and Fungal Aerosols in Beijing Subway. *Aerosol and Air Quality Research* **17**, 799-809, doi:10.4209/aaqr.2016.03.0114 (2017).

- 76 Bowers, R. M., McLetchie, S., Knight, R. & Fierer, N. Spatial variability in airborne bacterial communities across land-use types and their relationship to the bacterial communities of potential source environments. *The ISME journal* **5**, 601-612 (2011).
- 77 Knights, D. *et al.* Bayesian community-wide culture-independent microbial source tracking. *Nature Methods* **8**, 761-763, doi:10.1038/nmeth.1650 (2011).
- 78 Qi, Y. *et al.* Temporal-spatial variations of fungal composition in PM_{2.5} and source tracking of airborne fungi in mountainous and urban regions. *Science of the total environment* **708**, 135027 (2020).
- 79 Mu, F. *et al.* Source identification of airborne bacteria in the mountainous area and the urban areas. *Atmospheric Research* **231**, 104676 (2020).
- 80 Andrews, J. H. & Hirano, S. S. *Microbial Ecology of Leaves*. 43-59 (Springer New York, 1991).
- 81 Legg, B. J. & Powell, F. A. Spore dispersal in a barley crop: A mathematical model. *Agricultural Meteorology* **20**, 47-67, doi:[https://doi.org/10.1016/0002-1571\(79\)90050-5](https://doi.org/10.1016/0002-1571(79)90050-5) (1979).
- 82 Aylor, D. E. & Ferrandino, F. J. Dispersion of spores released from an elevated line source within a wheat canopy. *Boundary-Layer Meteorology* **46**, 251-273, doi:10.1007/BF00120842 (1989).
- 83 Lee, I.-B. *et al.* The past, present and future of CFD for agro-environmental applications. *Computers and electronics in agriculture* **93**, 168-183 (2013).
- 84 Thatiparti, D. S., Ghia, U. & Mead, K. R. Computational fluid dynamics study on the influence of an alternate ventilation configuration on the possible flow path of infectious cough aerosols in a mock airborne infection isolation room. *Science and technology for the built environment* **23**, 355-366 (2017).

- 85 Salmela, A. *et al.* Measurement and Simulation of Biocontamination in an Enclosed Habitat. *Aerosol Science and Engineering* **4**, 101-110, doi:10.1007/s41810-020-00057-3 (2020).
- 86 Ma, M., Zhen, Y. & Mi, T. Characterization of bacterial communities in bioaerosols over Northern Chinese Marginal Seas and the Northwestern Pacific Ocean in spring. *Journal of Applied Meteorology and Climatology* **58**, 903-917 (2019).
- 87 Cáliz, J., Triadó-Margarit, X., Camarero, L. & Casamayor, E. O. A long-term survey unveils strong seasonal patterns in the airborne microbiome coupled to general and regional atmospheric circulations. *Proceedings of the National Academy of Sciences* **115**, 12229-12234 (2018).
- 88 Yeoh, G. H. & Yuen, K. K. *Computational fluid dynamics in fire engineering: theory, modelling and practice*. (Butterworth-Heinemann, 2009).
- 89 Adams, R. I., Miletto, M., Taylor, J. W. & Bruns, T. D. The Diversity and Distribution of Fungi on Residential Surfaces. *PLOS ONE* **8**, e78866, doi:10.1371/journal.pone.0078866 (2013).
- 90 Konya, T. & Scott, J. A. Recent Advances in the Microbiology of the Built Environment. *Current Sustainable/Renewable Energy Reports* **1**, 35-42, doi:10.1007/s40518-014-0007-4 (2014).
- 91 Frankel, M. *et al.* Seasonal Variations of Indoor Microbial Exposures and Their Relation to Temperature, Relative Humidity, and Air Exchange Rate. *Applied and Environmental Microbiology* **78**, 8289, doi:10.1128/AEM.02069-12 (2012).
- 92 Gusareva, E. S. *et al.* Microbial communities in the tropical air ecosystem follow a precise diel cycle. *Proceedings of the National Academy of Sciences* **116**, 23299-23308, doi:10.1073/pnas.1908493116 (2019).

- 93 Prussin, A. J. & Marr, L. C. Sources of airborne microorganisms in the built environment. *Microbiome* **3**, 78, doi:10.1186/s40168-015-0144-z (2015).
- 94 Jarvis, B. B. in *Mycotoxins and Food Safety* Vol. 504 43-52 (Springer, Boston, 2002).
- 95 Shoemaker, R. C. & House, D. E. Sick building syndrome (SBS) and exposure to water-damaged buildings: Time series study, clinical trial and mechanisms. *Neurotoxicology and Teratology* **28**, 573-588, doi:<https://doi.org/10.1016/j.ntt.2006.07.003> (2006).
- 96 Hashemi, A., Liyanage, C., Hadjri, K. & Khatami, N. Assessing the risks of dampness and mould growth in renovated properties. (2014).
- 97 Howieson, S. Are our homes making us ill? The impact of energy efficiency on indoor air quality. *Perspectives in public health* **134**, 318-319, doi:10.1177/1757913914551916 (2014).
- 98 Howieson, S., Sharpe, T. & Farren, P. Building tight–ventilating right? How are new air tightness standards affecting indoor air quality in dwellings? *Building Services Engineering Research and Technology* **35**, 475-487 (2014).
- 99 Raw, G. *Sick building syndrome: a review of the evidence on causes and solutions*. (HM Stationery Office, 1992).
- 100 Howieson, S. *et al.* Domestic ventilation rates, indoor humidity and dust mite allergens: are our homes causing the asthma pandemic? *Building Services Engineering Research and Technology* **24**, 137-147 (2003).
- 101 Straus, D. C., Cooley, J. D., Wong, W. C. & Jumper, C. A. Studies on the role of fungi in Sick Building Syndrome. *Archives of environmental health* **58**, 475-478, doi:10.3200/aeoh.58.8.475-478 (2003).

- 102 Cooley, J. D., Wong, W. C., Jumper, C. A. & Straus, D. C. Correlation between the prevalence of certain fungi and sick building syndrome. *Occup Environ Med* **55**, 579-584, doi:10.1136/oem.55.9.579 (1998).
- 103 Burge, H. A. Fungi: toxic killers or unavoidable nuisances? *Annals of Allergy, Asthma & Immunology* **87**, 52-56, doi:[https://doi.org/10.1016/S1081-1206\(10\)62342-3](https://doi.org/10.1016/S1081-1206(10)62342-3) (2001).
- 104 Joshi, S. M. The sick building syndrome. *Indian J Occup Environ Med* **12**, 61-64, doi:10.4103/0019-5278.43262 (2008).
- 105 Van Gelder, L., Janssen, H. & Roels, S. in *Proceedings of the 5th International Building Physics Conference (IBPC)*. 112-112.
- 106 Marincioni, V., Marra, G. & Altamirano-Medina, H. Development of predictive models for the probabilistic moisture risk assessment of internal wall insulation. *Building and Environment* **137**, 257-267, doi:<https://doi.org/10.1016/j.buildenv.2018.04.001> (2018).
- 107 Smirnov, A., Rowan, S. & McCormick, J. Aerosols dispersion modelling using probabilistic particle tracking. *International journal for numerical methods in fluids* **54**, 295-311 (2007).
- 108 Oladokun, M. O., Ali, M., Osman, S. B. & Lin, Z. Indoor mould growth prediction using coupled computational fluid dynamics and mould growth model. *Building Simulation* **10**, 551-562, doi:10.1007/s12273-016-0343-y (2017).
- 109 Rosenstock, L. The future of intervention research at NIOSH. *American Journal of Industrial Medicine* **29**, 295-297, doi:10.1002/(sici)1097-0274(199604)29:4<295::aid-ajim3>3.0.co;2-o (1996).

- 110 Adams, R. I. *et al.* Ten questions concerning the microbiomes of buildings. *Building and Environment* **109**, 224-234, doi:<https://doi.org/10.1016/j.buildenv.2016.09.001> (2016).
- 111 Leung, M. H. Y. & Lee, P. K. H. The roles of the outdoors and occupants in contributing to a potential pan-microbiome of the built environment: a review. *Microbiome* **4**, 21, doi:10.1186/s40168-016-0165-2 (2016).
- 112 Miletto, M. & Lindow, S. E. Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. *Microbiome* **3**, 61, doi:10.1186/s40168-015-0128-z (2015).
- 113 Chase, J. *et al.* Geography and Location Are the Primary Drivers of Office Microbiome Composition. *mSystems* **1**, e00022-00016, doi:10.1128/mSystems.00022-16 (2016).
- 114 Sharma, A. *et al.* Longitudinal homogenization of the microbiome between both occupants and the built environment in a cohort of United States Air Force Cadets. *Microbiome* **7**, 70, doi:10.1186/s40168-019-0686-6 (2019).
- 115 Menzel, P., Ng, K. L. & Krogh, A. Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications* **7**, 11257, doi:10.1038/ncomms11257 (2016).
- 116 Community, B. O. *Blender - a 3D modelling and rendering package*, <<http://www.blender.org>> (2018).
- 117 Hess, R. *Blender Foundations: The Essential Guide to Learning Blender 2.5*. (Taylor & Francis, 2013).
- 118 Jasak, H., Jemcov, A. & Tukovic, Z. in *International workshop on coupled methods in numerical dynamics*. 1-20 (IUC Dubrovnik Croatia).

- 119 Ayachit, U. *The paraview guide: a parallel visualization application*. (Kitware, Inc., 2015).
- 120 Agency, U. S. E. P. *Moisture Control, Part of Indoor Air Quality Design Tools for Schools*, <<https://www.epa.gov/iaq-schools/moisture-control-part-indoor-air-quality-design-tools-schools>> (n.d.).
- 121 Frankel, M. *et al.* Seasonal variations of indoor microbial exposures and their relation to temperature, relative humidity, and air exchange rate. *Applied and environmental microbiology* **78**, 8289-8297 (2012).
- 122 Dannemiller, K. C., Gent, J. F., Leaderer, B. P. & Peccia, J. Influence of housing characteristics on bacterial and fungal communities in homes of asthmatic children. *Indoor air* **26**, 179-192 (2016).
- 123 Tang, J. W. The effect of environmental parameters on the survival of airborne infectious agents. *Journal of the Royal Society Interface* **6**, S737-S746 (2009).
- 124 Jump, R. L., Pultz, M. J. & Donskey, C. J. Vegetative *Clostridium difficile* survives in room air on moist surfaces and in gastric contents with reduced acidity: a potential mechanism to explain the association between proton pump inhibitors and *C. difficile*-associated diarrhea? *Antimicrobial agents and chemotherapy* **51**, 2883-2887 (2007).
- 125 Zhao, Y. *et al.* Effects of temperature, relative humidity, absolute humidity, and evaporation potential on survival of airborne Gumboro vaccine virus. *Applied and environmental microbiology* **78**, 1048-1054 (2012).
- 126 Lopez, G. U. *et al.* Transfer efficiency of bacteria and viruses from porous and nonporous fomites to fingers under different relative humidity conditions. *Applied and environmental microbiology* **79**, 5728-5734 (2013).
- 127 Dannemiller, K. C., Weschler, C. J. & Peccia, J. Fungal and bacterial growth in floor dust at elevated relative humidity levels. *Indoor air* **27**, 354-363 (2017).

- 128 Frankel, M., Hansen, E. W. & Madsen, A. M. Effect of relative humidity on the aerosolization and total inflammatory potential of fungal particles from dust-inoculated gypsum boards. *Indoor Air* **24**, 16-28 (2014).
- 129 Hyytiäinen, H. K. *et al.* Crawling-induced floor dust resuspension affects the microbiota of the infant breathing zone. *Microbiome* **6**, 1-12 (2018).
- 130 Wu, T. *et al.* Infant and adult inhalation exposure to resuspended biological particulate matter. *Environmental science & technology* **52**, 237-247 (2018).
- 131 Wolkoff, P. Indoor air humidity, air quality, and health—An overview. *International journal of hygiene and environmental health* **221**, 376-390 (2018).
- 132 Pasanen, A. L., Kalliokoski, P., Pasanen, P., Jantunen, M. J. & Nevalainen, A. Laboratory studies on the relationship between fungal growth and atmospheric temperature and humidity. *Environment International* **17**, 225-228, doi:[https://doi.org/10.1016/0160-4120\(91\)90006-C](https://doi.org/10.1016/0160-4120(91)90006-C) (1991).
- 133 Stekelenburg, F. K. & Labots, H. Measurement of water activity with an electric hygrometer. *International Journal of Food Science & Technology* **26**, 111-116, doi:<https://doi.org/10.1111/j.1365-2621.1991.tb01146.x> (1991).
- 134 Troller, J. A. Water Activity Measurements with A Capacitance Manometer. *Journal of Food Science* **48**, 739-741, doi:<https://doi.org/10.1111/j.1365-2621.1983.tb14886.x> (1983).
- 135 Pasanen, A.-L., Juutinen, T., Jantunen, M. & Kalliokoski, P. Occurrence and moisture requirements of microbial growth in building materials. *International biodeterioration & biodegradation* **30**, 273-283 (1992).
- 136 Stefanowski, B., Curling, S. & Ormondroyd, G. A rapid screening method to determine the susceptibility of bio-based construction and insulation products to

- mould growth. *International Biodeterioration & Biodegradation* **116**, 124-132 (2017).
- 137 Polizzi, V. *et al.* JEM Spotlight: Fungi, mycotoxins and microbial volatile organic compounds in mouldy interiors from water-damaged buildings. *Journal of Environmental Monitoring* **11**, 1849-1858 (2009).
- 138 Andersen, B., Frisvad, J. C., Søndergaard, I., Rasmussen, I. S. & Larsen, L. S. Associations between fungal species and water-damaged building materials. *Applied and environmental microbiology* **77**, 4180-4188 (2011).
- 139 Hyvärinen, A., Meklin, T., Vepsäläinen, A. & Nevalainen, A. Fungi and actinobacteria in moisture-damaged building materials—concentrations and diversity. *International Biodeterioration & Biodegradation* **49**, 27-37 (2002).
- 140 Gao, P., Korley, F., Martin, J. & Chen, B. T. Determination of unique microbial volatile organic compounds produced by five *Aspergillus* species commonly found in problem buildings. *AIHA Journal* **63**, 135-140 (2002).
- 141 Ezeonu, I. *et al.* Effect of relative humidity on fungal colonization of fiberglass insulation. *Applied and Environmental Microbiology* **60**, 2149-2151 (1994).
- 142 Lugauskas, A., Levinskait, L. & Pečiulyt, D. Micromycetes as deterioration agents of polymeric materials. *International biodeterioration & biodegradation* **52**, 233-242 (2003).
- 143 Gravesen, S., Nielsen, P. A., Iversen, R. & Nielsen, K. F. Microfungal contamination of damp buildings--examples of risk constructions and risk materials. *Environmental Health Perspectives* **107**, 505-508 (1999).
- 144 Salo, M. J. *et al.* *Penicillium expansum* strain isolated from indoor building material was able to grow on gypsum board and emitted guttation droplets containing

- chaetoglobosins and communesins A, B and D. *Journal of applied microbiology* **127**, 1135-1147 (2019).
- 145 Grant, C., Hunter, C., Flannigan, B. & Bravery, A. The moisture requirements of moulds isolated from domestic dwellings. *International Biodeterioration* **25**, 259-284 (1989).
- 146 Lax, S. *et al.* Microbial and metabolic succession on common building materials under high humidity conditions. *Nature communications* **10**, 1-12 (2019).
- 147 Memarzadeh, F. Literature Review of the Effect of Temperature and Humidity on Viruses. *ASHRAE transactions* **118** (2012).
- 148 Levin, J., Riley, L. S., Parrish, C., English, D. & Ahn, S. The effect of portable pulsed xenon ultraviolet light after terminal cleaning on hospital-associated *Clostridium difficile* infection in a community hospital. *American journal of infection control* **41**, 746-748 (2013).
- 149 Miller, R., Simmons, S., Dale, C., Stachowiak, J. & Stibich, M. Utilization and impact of a pulsed-xenon ultraviolet room disinfection system and multidisciplinary care team on *Clostridium difficile* in a long-term acute care facility. *American journal of infection control* **43**, 1350-1353 (2015).
- 150 Foarde, K. *et al.* Biological Inactivation Efficiency by HVAC In-Duct Ultraviolet Light Systems. *Environmental Protection Agency* (2006).
- 151 Wang, Y., Sekhar, C., Bahnfleth, W. P., Cheong, K. W. & Firrantello, J. Effectiveness of an ultraviolet germicidal irradiation system in enhancing cooling coil energy performance in a hot and humid climate. *Energy and Buildings* **130**, 321-329 (2016).
- 152 Finnegan, M. *et al.* Mode of action of hydrogen peroxide and other oxidizing agents: differences between liquid and gas forms. *The Journal of antimicrobial chemotherapy* **65**, 2108-2115, doi:10.1093/jac/dkq308 (2010).

- 153 Bockmühl, D. P., Schages, J. & Rehberg, L. Laundry and textile hygiene in healthcare and beyond. *Microbial Cell* **6**, 299 (2019).
- 154 Severa, J. & Klaban, V. [DESIDENT CaviCide a new disinfectant]. *Cas Lek Cesk* **148**, 269-270, 272 (2009).
- 155 Sundheim, G., Langsrud, S., Heir, E. & Holck, A. L. Bacterial resistance to disinfectants containing quaternary ammonium compounds. *International Biodeterioration & Biodegradation* **41**, 235-239, doi:[https://doi.org/10.1016/S0964-8305\(98\)00027-4](https://doi.org/10.1016/S0964-8305(98)00027-4) (1998).
- 156 Jennings, M. C., Minbiole, K. P. & Wuest, W. M. Quaternary ammonium compounds: an antimicrobial mainstay and platform for innovation to address bacterial resistance. *ACS infectious diseases* **1**, 288-303 (2015).
- 157 Dubas, S. T., Kumlangdudsana, P. & Potiyaraj, P. Layer-by-layer deposition of antimicrobial silver nanoparticles on textile fibers. *Colloids and Surfaces A: Physicochemical and Engineering Aspects* **289**, 105-109 (2006).
- 158 da Silva, G. *et al.* Silver nanoparticles in building materials for environment protection against microorganisms. *International Journal of Environmental Science and Technology* **16**, 1239-1248 (2019).
- 159 Eby, D. M., Luckarift, H. R. & Johnson, G. R. Hybrid Antimicrobial Enzyme and Silver Nanoparticle Coatings for Medical Instruments. *ACS Applied Materials & Interfaces* **1**, 1553-1560, doi:10.1021/am9002155 (2009).
- 160 Tarantola, A. Popper, Bayes and the inverse problem. *Nature Physics* **2**, 492-494, doi:10.1038/nphys375 (2006).
- 161 Miletto, M. & Lindow, S. E. Relative and contextual contribution of different sources to the composition and abundance of indoor air bacteria in residences. *Microbiome* **3**, 1-14 (2015).

- 162 Adams, R. I., Bateman, A. C., Bik, H. M. & Meadow, J. F. Microbiota of the indoor environment: a meta-analysis. *Microbiome* **3**, 49, doi:10.1186/s40168-015-0108-3 (2015).
- 163 Flume, P. A., Chalmers, J. D. & Olivier, K. N. Advances in bronchiectasis: endotyping, genetics, microbiome, and disease heterogeneity. *The Lancet* **392**, 880-890 (2018).
- 164 Dickson, R. P., Erb-Downward, J. R. & Huffnagle, G. B. Homeostasis and its disruption in the lung microbiome. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **309**, L1047-L1055 (2015).
- 165 Hilty, M. *et al.* Disordered microbial communities in asthmatic airways. *PloS one* **5** (2010).
- 166 Dickson, R. P. The microbiome and critical illness. *The Lancet Respiratory Medicine* **4**, 59-72 (2016).
- 167 Dickson, R. P., Erb-Downward, J. R., Martinez, F. J. & Huffnagle, G. B. The microbiome and the respiratory tract. *Annual review of physiology* **78**, 481-504 (2016).
- 168 Miravittles, M. & Anzueto, A. Antibiotics for acute and chronic respiratory infection in patients with chronic obstructive pulmonary disease. *American journal of respiratory and critical care medicine* **188**, 1052-1057 (2013).
- 169 Cox, M. J. *et al.* Longitudinal assessment of sputum microbiome by sequencing of the 16S rRNA gene in non-cystic fibrosis bronchiectasis patients. *PloS one* **12** (2017).
- 170 Taylor, S. L. *et al.* Understanding the impact of antibiotic therapies on the respiratory tract resistome: a novel pooled-template metagenomic sequencing strategy. *Multidisciplinary respiratory medicine* **13**, 30 (2018).

- 171 Malla, M. A. *et al.* Exploring the human microbiome: The potential future role of
next-generation sequencing in disease diagnosis and treatment. *Frontiers in*
immunology **9**, 2868 (2019).
- 172 Relman, D. A. & Lipsitch, M. Microbiome as a tool and a target in the effort to
address antimicrobial resistance. *Proceedings of the National Academy of Sciences*
115, 12902-12910 (2018).
- 173 Vandeplasseche, E., Tavernier, S., Coenye, T. & Crabbé, A. Influence of the lung
microbiome on antibiotic susceptibility of cystic fibrosis pathogens. *European*
Respiratory Review **28**, 190041 (2019).
- 174 Pinto, A. J. & Raskin, L. PCR biases distort bacterial and archaeal community
structure in pyrosequencing datasets. *PloS one* **7** (2012).
- 175 Hong, S., Bunge, J., Leslin, C., Jeon, S. & Epstein, S. S. Polymerase chain reaction
primers miss half of rRNA microbial diversity. *The ISME journal* **3**, 1365-1373
(2009).
- 176 Martin, M. Cutadapt removes adapter sequences from highthroughput sequencing
reads. *EMBnet J* **17**, 10-12 (2011).
- 177 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nature*
Methods **9**, 357-359, doi:10.1038/nmeth.1923 (2012).
- 178 Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using
DIAMOND. *Nature Methods* **12**, 59-60, doi:10.1038/nmeth.3176 (2015).
- 179 Huson, D. H., Auch, A. F., Qi, J. & Schuster, S. C. MEGAN analysis of metagenomic
data. *Genome Res* **17**, 377-386, doi:10.1101/gr.5969107 (2007).
- 180 Kaminski, J. *et al.* High-Specificity Targeted Functional Profiling in Microbial
Communities with ShortBRED. *PLOS Computational Biology* **11**, e1004557,
doi:10.1371/journal.pcbi.1004557 (2015).

- 181 Warnes, G. R. *et al.* gplots: Various R programming tools for plotting data. (2015).
- 182 Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-
customizable Venn and Euler diagrams in R. *BMC bioinformatics* **12**, 35 (2011).
- 183 Conway, J. R., Lex, A. & Gehlenborg, N. UpSetR: an R package for the visualization
of intersecting sets and their properties. *Bioinformatics* **33**, 2938-2940 (2017).
- 184 Oksanen, J. *et al.* (2016).
- 185 McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive
analysis and graphics of microbiome census data. *PloS one* **8** (2013).
- 186 Faust, K. *et al.* Microbial Co-occurrence Relationships in the Human Microbiome.
PLOS Computational Biology **8**, e1002606, doi:10.1371/journal.pcbi.1002606 (2012).
- 187 Harrell Jr, F. E. & Harrell Jr, M. F. E. Package ‘Hmisc’. *CRAN2018*, 235-236 (2015).
- 188 Sitkin, S. I., Tkachenko, E. I. & Vakhitov, T. Y. METABOLIC DYSBIOSIS OF THE
GUT MICROBIOTA AND ITS BIOMARKERS. *Ekspierimental'naia i klinicheskaia
gastroenterologiiia = Experimental & clinical gastroenterology* **12**, 6-29 (2016).
- 189 Shreiner, A. B., Kao, J. Y. & Young, V. B. The gut microbiome in health and in
disease. *Current opinion in gastroenterology* **31**, 69 (2015).
- 190 Alvarado, A. & Arce, I. Metabolic functions of the lung, disorders and associated
pathologies. *Journal of clinical medicine research* **8**, 689 (2016).
- 191 Lamonaca, P. *et al.* Metabolic disorder in chronic obstructive pulmonary disease
(COPD) patients: towards a personalized approach using marine drug derivatives.
Marine drugs **15**, 81 (2017).
- 192 Kilk, K. *et al.* Phenotyping of chronic obstructive pulmonary disease based on the
integration of metabolomes and clinical characteristics. *International journal of
molecular sciences* **19**, 666 (2018).

- 193 Petrache, I. *et al.* Ceramide causes pulmonary cell apoptosis and emphysema: a role for sphingolipid homeostasis in the maintenance of alveolar cells. *Proceedings of the American Thoracic Society* **3**, 510-510 (2006).
- 194 Petrache, I. *et al.* Ceramide upregulation causes pulmonary cell apoptosis and emphysema-like disease in mice. *Nature medicine* **11**, 491-498 (2005).
- 195 Budden, K. F. *et al.* Functional effects of the microbiota in chronic respiratory disease. *The Lancet Respiratory Medicine* **7**, 907-920, doi:10.1016/S2213-2600(18)30510-1 (2019).
- 196 Humphries, Z., Seifert, K. A., Hirooka, Y. & Visagie, C. M. A new family and genus in Dothideales for *Aureobasidium*-like species isolated from house dust. *IMA Fungus* **8**, 299-315, doi:10.5598/ima fungus.2017.08.02.05 (2017).
- 197 Mousavi, B., Hedayati, M. T., Hedayati, N., Ilkit, M. & Syedmousavi, S. *Aspergillus* species in indoor environments and their possible occupational and public health hazards. *Current medical mycology* **2**, 36 (2016).
- 198 LeMessurier, K. S., Tiwary, M., Morin, N. P. & Samarasinghe, A. E. Respiratory barrier as a safeguard and regulator of defense against influenza A virus and *Streptococcus pneumoniae*. *Frontiers in immunology* **11**, 3 (2020).
- 199 Kumpitsch, C., Koskinen, K., Schöpf, V. & Moissl-Eichinger, C. The microbiome of the upper respiratory tract in health and disease. *BMC biology* **17**, 1-20 (2019).
- 200 Chen, L., He, X., Li, T., Luo, L. & Chen, Y. Pathogenic bacteria for exacerbation of Chronic Obstructive Pulmonary Disease. *European Respiratory Journal* **52**, PA4099, doi:10.1183/13993003.congress-2018.PA4099 (2018).
- 201 Rodrigo-Troyano, A. & Sibila, O. The respiratory threat posed by multidrug resistant Gram-negative bacteria. *Respirology* **22**, 1288-1299 (2017).

- 202 Herath, S. C. & Poole, P. Prophylactic antibiotic therapy for chronic obstructive pulmonary disease (COPD). *The Cochrane database of systematic reviews*, Cd009764, doi:10.1002/14651858.CD009764.pub2 (2013).
- 203 Laska, I. F., Crichton, M. L., Shoemark, A. & Chalmers, J. D. The efficacy and safety of inhaled antibiotics for the treatment of bronchiectasis in adults: a systematic review and meta-analysis. *The Lancet. Respiratory medicine* **7**, 855-869, doi:10.1016/s2213-2600(19)30185-7 (2019).
- 204 Taylor, S. L. *et al.* Long-Term Azithromycin Reduces Haemophilus influenzae and Increases Antibiotic Resistance in Severe Asthma. *American Journal of Respiratory and Critical Care Medicine* **200**, 309-317, doi:10.1164/rccm.201809-1739OC (2019).
- 205 Welte, T. Azithromycin: The Holy Grail to Prevent Exacerbations in Chronic Respiratory Disease? *Am J Respir Crit Care Med* **200**, 269-270, doi:10.1164/rccm.201903-0706ED (2019).
- 206 Hendriksen, R. S. *et al.* Global monitoring of antimicrobial resistance based on metagenomics analyses of urban sewage. *Nature Communications* **10**, 1124, doi:10.1038/s41467-019-08853-3 (2019).
- 207 Ohashi, Y. & Fujisawa, T. Detection of antibiotic resistance genes in the feces of young adult Japanese. *Biosci Microbiota Food Health* **36**, 151-154, doi:10.12938/bmfh.17-004 (2017).
- 208 Kennedy, M. *et al.* Face mask sampling reveals antimicrobial resistance genes in exhaled aerosols from patients with chronic obstructive pulmonary disease and healthy volunteers. *BMJ Open Respiratory Research* **5**, e000321, doi:10.1136/bmjresp-2018-000321 (2018).

- 209 Blaustein, R. A. *et al.* Toothbrush microbiomes feature a meeting ground for human oral and environmental microbiota. *Microbiome* **9**, 32, doi:10.1186/s40168-020-00983-x (2021).
- 210 Sun, J., Deng, Z. & Yan, A. Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochemical and biophysical research communications* **453**, 254-267 (2014).
- 211 Vogwill, T. & MacLean, R. C. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evolutionary applications* **8**, 284-295 (2015).
- 212 Kidner, C. A. & Timmermans, M. C. Signaling sides adaxial-abaxial patterning in leaves. *Current topics in developmental biology* **91**, 141-168, doi:10.1016/s0070-2153(10)91005-3 (2010).
- 213 Kirkwood, R. C. Recent developments in our understanding of the plant cuticle as a barrier to the foliar uptake of pesticides†. *Pesticide Science* **55**, 69-77, doi:10.1002/(sici)1096-9063(199901)55:1<69::aid-ps860>3.0.co;2-h (1999).
- 214 Tian, N. *et al.* The molecular basis of glandular trichome development and secondary metabolism in plants. *Plant Gene* **12**, 1-12 (2017).
- 215 Finkel, O. M., Burch, A. Y., Lindow, S. E., Post, A. F. & Belkin, S. Geographical Location Determines the Population Structure in Phyllosphere Microbial Communities of a Salt-Excreting Desert Tree. *Applied and Environmental Microbiology* **77**, 7647-7655, doi:10.1128/aem.05565-11 (2011).
- 216 Kim, M. *et al.* Distinctive Phyllosphere Bacterial Communities in Tropical Trees. *Microbial Ecology* **63**, 674-681, doi:10.1007/s00248-011-9953-1 (2012).
- 217 Williams, T. R. & Marco, M. L. Phyllosphere Microbiota Composition and Microbial Community Transplantation on Lettuce Plants Grown Indoors. *mBio* **5**, doi:10.1128/mBio.01564-14 (2014).

- 218 Hong, S., Bunge, J., Leslin, C., Jeon, S. & Epstein, S. S. Polymerase chain reaction
primers miss half of rRNA microbial diversity. *Isme j* **3**, 1365-1373,
doi:10.1038/ismej.2009.89 (2009).
- 219 Pinto, A. J. & Raskin, L. PCR Biases Distort Bacterial and Archaeal Community
Structure in Pyrosequencing Datasets. *PLOS ONE* **7**, e43093,
doi:10.1371/journal.pone.0043093 (2012).
- 220 Henny, R. J., Chen, J. & Janick, J. Cultivar Development of Ornamental Foliage
Plants. *Plant Breeding Reviews* **23**, 281 (2003).
- 221 Little, E. L. & Skolmen, R. G. Common forest trees of Hawaii (native and
introduced). *Agricultural Handbook No. 679*, 321 (1989).
- 222 Luhung, I. *et al.* Protocol Improvements for Low Concentration DNA-Based
Bioaerosol Sampling and Analysis. *PLOS ONE* **10**, e0141158,
doi:10.1371/journal.pone.0141158 (2015).
- 223 Hastings, L. H. A Revision of *Rhapis*, the Lady Palms. *Palms* **47** (2003).
- 224 Gilbert, V. *Foliage for Florists*. 54, 150 (The Society of Floristry Limited, 2006).
- 225 Levin, R. A. *et al.* Family-level relationships of Onagraceae based on chloroplast *rbcl*
and *ndhF* data. *American journal of botany* **90**, 107-115, doi:10.3732/ajb.90.1.107
(2003).
- 226 Kress, W. J. *et al.* Plant DNA barcodes and a community phylogeny of a tropical
forest dynamics plot in Panama. *Proceedings of the National Academy of Sciences of
the United States of America* **106**, 18621-18626, doi:10.1073/pnas.0909820106
(2009).
- 227 Ford, C. S. *et al.* Selection of candidate coding DNA barcoding regions for use on
land plants. *Botanical Journal of the Linnean Society* **159**, 1-11, doi:10.1111/j.1095-
8339.2008.00938.x (2009).

- 228 T., D. L. & Savolainen, V. Broad-scale amplification of matK for DNA barcoding plants, a technical note. *Botanical Journal of the Linnean Society* **164**, 1-9, doi:doi:10.1111/j.1095-8339.2010.01071.x (2010).
- 229 Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**, 95-98, doi:citeulike-article-id:691774 (1999).
- 230 Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids research* **25**, 3389-3402 (1997).
- 231 Ratnasingham, S. & Hebert, P. D. N. BOLD: The Barcode of Life Data System. *Molecular ecology notes* **7**, 355-364, doi:10.1111/j.1471-8286.2007.01678.x (2007).
- 232 Andrews, S. FastQC A Quality Control tool for High Throughput Sequence Data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, doi:citeulike-article-id:11583827.
- 233 Zhao, Y., Tang, H. & Ye, Y. RAPSearch2: a fast and memory-efficient protein similarity search tool for next-generation sequencing data. *Bioinformatics (Oxford, England)* **28**, 125-126, doi:10.1093/bioinformatics/btr595 (2012).
- 234 Davis, N. M., Proctor, D., Holmes, S. P., Relman, D. A. & Callahan, B. J. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *bioRxiv*, doi:10.1101/221499 (2017).
- 235 Development Core Team, R. *R: A Language and Environment for Statistical Computing*. Vol. 1 (2013).
- 236 Oksanen, J. *et al.* *Community Ecology Package 'vegan' version 2.4-6.* (2018).
- 237 McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLOS ONE* **8**, e61217, doi:10.1371/journal.pone.0061217 (2013).

- 238 Swift, M. L. GraphPad Prism, Data Analysis, and Scientific Graphing. *Journal of Chemical Information and Computer Sciences* **37**, 411-412, doi:10.1021/ci960402j (1997).
- 239 M., B., S., H. & M., J. Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*. (2009).
- 240 Li, B. *et al.* Metagenomic and network analysis reveal wide distribution and co-occurrence of environmental antibiotic resistance genes. *The ISME Journal* **9**, 2490-2502, doi:10.1038/ismej.2015.59 (2015).
- 241 Lee, J., Fortner, J. D., Hughes, J. B. & Kim, J.-H. Photochemical Production of Reactive Oxygen Species by C60 in the Aqueous Phase During UV Irradiation. *Environmental Science & Technology* **41**, 2529-2535, doi:10.1021/es062066l (2007).
- 242 Nemzer, B., Pietrkowski, Z., Chang, T. & Ou, B. A Novel Approach for Measurement of Total Reactive Oxidant Species (ROS) In Vivo by A Fluorometric Method. *American Journal of Biomedical Sciences* (2013).
- 243 Foster, A. J., Pelletier, G., Tanguay, P. & Séguin, A. Transcriptome Analysis of Poplar during Leaf Spot Infection with *Sphaerulina* spp. *PLOS ONE* **10**, e0138162, doi:10.1371/journal.pone.0138162 (2015).
- 244 Arango Isaza, R. E. *et al.* Combating a Global Threat to a Clonal Crop: Banana Black Sigatoka Pathogen *Pseudocercospora fijiensis* (Synonym *Mycosphaerella fijiensis*) Genomes Reveal Clues for Disease Control. *PLOS Genetics* **12**, e1005876, doi:10.1371/journal.pgen.1005876 (2016).
- 245 Fabre, B., Ioos, R., Piou, D. & Marcais, B. Is the emergence of *Dothistroma* needle blight of pine in France caused by the cryptic species *Dothistroma pini*? *Phytopathology* **102**, 47-54, doi:10.1094/phyto-02-11-0036 (2012).

- 246 Teixeira, M. M. *et al.* Exploring the genomic diversity of black yeasts and relatives (Chaetothyriales, Ascomycota). *Studies in mycology* **86**, 1-28, doi:10.1016/j.simyco.2017.01.001 (2017).
- 247 Wang, Q. M. *et al.* Multigene phylogeny and taxonomic revision of yeasts and related fungi in the Ustilaginomycotina. *Studies in Mycology* **81**, 55-83 (2015).
- 248 Andrews, J. H., Spear, R. N. & Nordheim, E. V. Population biology of *Aureobasidium pullulans* on apple leaf surfaces. *Canadian journal of microbiology* **48**, 500-513 (2002).
- 249 Peterson, S. W., Manitchotpisit, P. & Leathers, T. D. *Aureobasidium thailandense* sp. nov. isolated from leaves and wooden surfaces. *International journal of systematic and evolutionary microbiology* **63**, 790-795, doi:10.1099/ij.s.0.047613-0 (2013).
- 250 Cai, R. *et al.* Depsidones from *Talaromyces stipitatus* SK-4, an endophytic fungus of the mangrove plant *Acanthus ilicifolius*. *Phytochemistry Letters* **20**, 196-199 (2017).
- 251 Walker, A. K. *et al.* Full Genome of *Phialocephala scopiformis* DAOMC 229536, a Fungal Endophyte of Spruce Producing the Potent Anti-Insectan Compound Rugulosin. *Genome announcements* **4**, e01768-01715, doi:10.1128/genomeA.01768-15 (2016).
- 252 Chen, L. *et al.* Genomics-driven discovery of the pneumocandin biosynthetic gene cluster in the fungus *Glarea lozoyensis*. *BMC genomics* **14**, 339, doi:10.1186/1471-2164-14-339 (2013).
- 253 Yang, H. *et al.* Mapping and screening of the tomato *Stemphylium lycopersici* resistance gene, Sm, based on bulked segregant analysis in combination with genome resequencing. *BMC Plant Biol* **17**, 266, doi:10.1186/s12870-017-1215-z (2017).

- 254 Howlett, B. J., Idnurm, A. & Pedras, M. S. C. *Leptosphaeria maculans*, the Causal Agent of Blackleg Disease of Brassicas. *Fungal Genetics and Biology* **33**, 1-14, doi:<https://doi.org/10.1006/fgbi.2001.1274> (2001).
- 255 Downie, R. C. *et al.* Assessing European Wheat Sensitivities to *Parastagonospora nodorum* Necrotrophic Effectors and Fine-Mapping the Snn3-B1 Locus Conferring Sensitivity to the Effector SnTox3. *Frontiers in Plant Science* **9**, doi:10.3389/fpls.2018.00881 (2018).
- 256 C Gaylarde, C. Influence of environment on microbial colonization of historic stone buildings with emphasis on cyanobacteria. *Heritage* **3**, 1469-1482 (2020).
- 257 O'Brien, J. A., Daudi, A., Butt, V. S. & Bolwell, G. P. Reactive oxygen species and their role in plant defence and cell wall metabolism. *Planta* **236**, 765-779, doi:10.1007/s00425-012-1696-9 (2012).
- 258 Crous, P. W. *et al.* Fungal pathogens of Proteaceae. *Persoonia* **27**, 20-45, doi:10.3767/003158511X606239 (2011).
- 259 Boateng, K. & Lewis, K. J. Spore Dispersal by *Dothistroma septosporum* in Northwest British Columbia. *Phytopathology* **105**, 69-79, doi:10.1094/phyto-06-13-0175-r (2015).
- 260 Chang, T.-C., Salvucci, A., Crous, P. W. & Stergiopoulos, I. Comparative Genomics of the Sigatoka Disease Complex on Banana Suggests a Link between Parallel Evolutionary Changes in *Pseudocercospora fijiensis* and *Pseudocercospora eumusae* and Increased Virulence on the Banana Host. *PLOS Genetics* **12**, e1005904, doi:10.1371/journal.pgen.1005904 (2016).
- 261 Churchill, A. C. *Mycosphaerella fijiensis*, the black leaf streak pathogen of banana: progress towards understanding pathogen biology and detection, disease development,

- and the challenges of control. *Molecular plant pathology* **12**, 307-328, doi:10.1111/j.1364-3703.2010.00672.x (2011).
- 262 Beltran-Garcia, M. J. *et al.* Oxidative stress response of *Mycosphaerella fijiensis*, the causal agent of black leaf streak disease in banana plants, to hydrogen peroxide and paraquat. *Canadian journal of microbiology* **55**, 887-894, doi:10.1139/w09-023 (2009).
- 263 Choi, Y.-E., Lee, C. & Goodwin, S. B. Generation of Reactive Oxygen Species via NOXa Is Important for Development and Pathogenicity of *Mycosphaerella graminicola*. *Mycobiology* **44**, 38-47, doi:10.5941/MYCO.2016.44.1.38 (2016).
- 264 Stosch *et al.* *Efficiency of self-cleaning properties in wheat (Triticum aestivum L.)*. Vol. 81 (2007).