

RESEARCH

Open Access



# Holobiont Urbanism: sampling urban beehives reveals cities' metagenomes

Elizabeth Hénaff<sup>1,2\*</sup>, Devora Najjar<sup>3</sup>, Miguel Perez<sup>3</sup>, Regina Flores<sup>4</sup>, Christopher Woebken<sup>5</sup>, Christopher E. Mason<sup>6,7,8</sup> and Kevin Slavin<sup>3</sup>

## Abstract

**Background** Over half of the world's population lives in urban areas with, according to the United Nations, nearly 70% expected to live in cities by 2050. Our cities are built by and for humans, but are also complex, adaptive biological systems involving a diversity of other living species. The majority of these species are invisible and constitute the city's microbiome. Our design decisions for the built environment shape these invisible populations, and as inhabitants we interact with them on a constant basis. A growing body of evidence shows us that human health and well-being are dependent on these interactions. Indeed, multicellular organisms owe meaningful aspects of their development and phenotype to interactions with the microorganisms—bacteria or fungi—with which they live in continual exchange and symbiosis. Therefore, it is meaningful to establish microbial maps of the cities we inhabit. While the processing and sequencing of environmental microbiome samples can be high-throughput, gathering samples is still labor and time intensive, and can require mobilizing large numbers of volunteers to get a snapshot of the microbial landscape of a city.

**Results** Here we postulate that honeybees may be effective collaborators in gathering samples of urban microbiota, as they forage daily within a 2-mile radius of their hive. We describe the results of a pilot study conducted with three rooftop beehives in Brooklyn, NY, where we evaluated the potential of various hive materials (honey, debris, hive swabs, bee bodies) to reveal information as to the surrounding metagenomic landscape, and where we conclude that the bee debris are the richest substrate. Based on these results, we profiled 4 additional cities through collected hive debris: Sydney, Melbourne, Venice and Tokyo. We show that each city displays a unique metagenomic profile as seen by honeybees. These profiles yield information relevant to hive health such as known bee symbionts and pathogens. Additionally, we show that this method can be used for human pathogen surveillance, with a proof-of-concept example in which we recover the majority of virulence factor genes for *Rickettsia felis*, a pathogen known to be responsible for "cat scratch fever".

**Conclusions** We show that this method yields information relevant to hive health and human health, providing a strategy to monitor environmental microbiomes on a city scale. Here we present the results of this study, and discuss them in terms of architectural implications, as well as the potential of this method for epidemic surveillance.

**Keywords** Built environment, Metagenomics, Aerobiome, Holobiont, Honeybee

\*Correspondence:

Elizabeth Hénaff  
ehenaff@nyu.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Introduction

Over half of the world's human population lives in urban areas and, according to the United Nations (UN), nearly 70% of us will live in cities by 2050 [1]. Our cities are built by and for humans, but are also complex, adaptive biological systems involving a diversity of living species [2]. The majority of these species are invisible and constitute the city's microbiome. Our design decisions for the built environment shape these invisible populations, and we interact with them on a constant basis [3, 4]. A growing body of evidence shows us that our health and well-being are dependent on these interactions [5]. Indeed, multicellular organisms owe meaningful aspects of their development and phenotype to interactions with the microorganisms—bacteria or fungi—with which they live in symbiosis [6, 7]. Accumulated evidence confirms that mammalian phenotypes are related to a combination of an individual's genotype as well as that of its microbiota, including disease states such as obesity [8] and influence on neuro-psychiatric disorders as well [9]. Beyond human consequences, plants' flowering time has been found to depend on the soil microbiome [10] and the useful metabolic compounds in medicinal plants are possibly synthesized in conjunction with their symbiont bacteria [11], both traits formerly thought to depend only on the plant's genotype. Metagenomic studies such as these are facilitated by the rapidly decreasing cost of high-throughput DNA sequencing, and support a growing understanding that the phenotype of a multicellular organism depends on both its own genotype and that of its associated microbes. As capacity for gathering and analyzing genomic and metagenomic data grows, our capacity to understand interspecies relationships is growing alongside it, with the potential of elucidating fundamental biological questions of host-symbiont selection and evolution mechanisms such as testing hologenome [12, 13] theories of evolution.

Metagenomics is a rapidly growing field that is well-situated to survey across all domains and kingdoms of life, including city-scale efforts of urban metagenomics. Microbial classification using high-throughput DNA sequencing is faster and more comprehensive than culture-based methods, and has enabled city-wide mapping of microbial populations [14–16]. Mapping indoor environments [3, 17] also provides insights into the relationship between humans and the indoor microbiome, which holds promise for designing buildings that optimize this metric. Thus, we are moving away from the germ-centric paradigm of microbes to the quantification of a ubiquitous, continuous and commensal map of the environmental microbiome within which we live, work, and sleep. While the processing and sequencing of samples can be high-throughput (with automation, hundreds

at a time), gathering samples is still very expensive, labor intensive, and can require mobilizing large numbers of volunteers to get a snapshot of the microbial landscape of a city, such as global City Sampling Day (metasub.org). Moreover, samples collected manually with swabs represent a limited area: 0.1–0.5m<sup>2</sup>. While this scale of resolution is important for applications such as tracking contamination through a hospital, it is not always easily implemented for city-scale studies and leads researchers to look for pinch points where samples might be most meaningful. Examples of this have been MetaSub sampling subways [16], air sampling in indoor environments [18], or sewers [19, 20].

Setting out to collect a more distributed and comprehensive sample of the urban landscape, following conversations with artists Timo Arnall and Jack Schulze, we investigated the potential of using honeybees as proxy sampling mechanisms for the urban microbiome. On average, honeybees forage within a 1–2 mile radius around their hive in rural environments [21] and 0.3–1 miles in urban environments [22], and we hypothesized that their travel would permit them to interact with various microbial environments including air, water, and mammalian sources in addition to their known plant targets. We designed a pilot study to test for geo-specific microbial residues corresponding to all of these environments within material found in a hive.

Here we describe the results of a pilot study conducted with three rooftop beehives in Brooklyn, NY, where we evaluated the potential of various hive materials (honey, debris, hive swabs, bee bodies) to reveal information as to the surrounding metagenomic landscape, and where we conclude that the hive debris are the richest substrate. Based on these results, we profiled four additional cities by collecting hive debris: Sydney, Melbourne, Venice and Tokyo. Here we present the results of this study, and discuss them in terms of architectural implications, as well as the potential of this method for epidemic surveillance.

## Methods

### Hives and collection methods

#### *U.S.A.—Brooklyn*

The hives of three independent beekeepers were sampled in New York City. The first location (AS) were Langstroth hives located in Astoria, Queens, NY. The second location (CH) were Langstroth and Top Bar hives located in Crown Heights, Brooklyn, NY. The third location (FG) were Langstroth hives located in Fort Greene, Brooklyn, NY. Samples of honey, bees, hive debris, and swabs of the inside of the hive were collected using sterile one-time-use scrapers and transferred into sterile 50 ml Falcon tubes. Bee bodies were submerged in isopropyl alcohol for storage.

**Australia—Sydney and Melbourne**

Hive debris from two Langstroth hives in Sydney (SYD1, SYD2) and two in Melbourne (MEL, SH) were sampled. Custom collection trays with self-sealing apertures, designed to be placed under the hives to collect hive debris, were developed and fabricated at MIT, and shipped to Sydney and Melbourne for deployment. Trays were installed for 1 week collections, then removed and hive debris samples were transferred to sterile 50ml Falcon tubes.

**Italy—Venice**

Hive debris from one Langstroth hive at the Palazzo Mora, Venice, Italy was sampled. Debris were collected from the hive using a sterile one-time-use scraper and transferred to 50ml Falcon tube.

**Japan—Tokyo**

Hive debris samples were collected from 12 hives distributed over 4 neighborhoods. Samples were collected with sterile one-time-use scrapers and stored in sterile 50ml Falcon tubes. The locations were Marunouchi (MA), 丸の内千代田区東京 100-0005, Mita (MI), 港区東京 108-0073日本, Marronnier Gate (MR), マロニエゲート銀座1, and Ginza (GK), 銀座中央区東京 104-0061.

**Sample preparation**

The general approach to DNA extraction involved a combination of lysis methods including mechanical, thermal, and enzymatic disruption to try and ensure that DNA from plant, microbe, and human sources would be extracted for sequencing.

**Honey**

The honey samples were diluted in a 1:1 ratio of grams of honey to mL of ultrapure water and then vortexed vigorously. The mixture was then spun down in the centrifuge at 3900 RCF for 20 minutes, the supernatant was discarded and the pellet along with

~ 200  $\mu$ L residual liquid was moved to an Eppendorf, and placed in the  $-20^{\circ}\text{C}$  freezer until the DNA extraction step.

**Bee debris**

The bee debris was diluted in a 1:5 ratio of grams of bee debris to mL of ultrapure water. The mixture was then heated in a water bath at  $70^{\circ}\text{C}$  for 5 minutes in order to soften the debris and have it disperse in the liquid and then spun on the vortex vigorously. The liquid and solids were then separated, and both were placed into Eppendorfs and placed in the  $-20^{\circ}\text{C}$  freezer so that a freeze-thaw cycle would help disrupt the cell

membranes. The bee debris material was then ground with a mortar and pestle to break down any large pieces of bee debris, and resuspended in 1X PBS to bring all of the tubes to a final volume of 20 mL. Then material was then allowed to settle, spun down at 3900 RCF for 20 minutes along with 1–2 grams of  $100\mu\text{m}$  glass beads to further mechanically disrupt the samples. The pellet and a small amount of the supernatant was then used for DNA extraction.

**Bees**

The isopropyl alcohol was drained from the tubes, then bees were placed in a mortar and pestle that was pre-chilled to  $-80^{\circ}\text{C}$  before use. The bees were crushed vigorously into a paste. The paste was then placed in Eppendorf tubes and placed in the  $-20^{\circ}\text{C}$  freezer until the DNA extraction step.

**Swabs**

The swabs, Copan Liquid Amies Elution Swab 481C, were stored in the  $-20^{\circ}\text{C}$  freezer until the DNA extraction step.

**DNA extraction**

The protocol for 3–5 mL of starting material of the Promega Wizard<sup>®</sup> Genomic DNA Purification Kit (A1120) was used, with the following alterations to the standard protocol: one hour incubation at  $37^{\circ}\text{C}$  in a shaker after the neutralization step; the samples were vortexed vigorously for about 1–2 minutes after the lysis and neutralization buffer were added to mechanically disturb the material; following this a phenol/chloroform step was done to remove any remaining organic matter before being placed in the spin column; the DNA was eluted with 20  $\mu$ L of TE buffer warmed to  $65^{\circ}\text{C}$ ; there was a 2 minute incubation time at room temperature before spinning down.

**Library preparation**

The Library preparation protocol was performed at the Mason Lab at Weill Cornell Medicine, using the following kits according to manufacturer's instructions. It was used to prepare libraries for all samples.

Illumina/Qiagen 500bp Prep:

1. Size selection with Agencourt AMPure XP Beads (A63881)
2. End repair and A-tailing: Qiagen GeneRead DNA Library I Core Kit (180,432)
3. Amplification: Qiagen GeneRead DNA Library I Amp Kit (180,455)

4. Illumina TruSeq DNA LT adapter kits A and B for up to 24-plex per sequencing pool.

### Sequencing

**Brooklyn Pilot Study:** The samples were sequenced at the BioMicro Center at MIT. The sequencing requested was a 150bp paired end sequence on one lane of the Illumina MiSeq. **Venice Study:** The sample was sequenced at the CNAG supercomputing center in Barcelona, Spain, with 150bp paired end reads on a Illumina MiSeq lane. **Australia and Tokyo samples:** Sequencing was performed on the Illumina HiSeq platform at Weill Cornell Medicine, with 125bp paired-end reads. See Additional file 6: Table S1 for read counts for all samples.

### Analysis

#### Metagenomic classification

Read quality was assessed with FastQC [23] and read quality was sufficient to not require trimming (see Additional file 7 for sample metadata, and Additional file 8 for MetaQC [24] reports). DIAMOND [25] – MEGAN [26] against the NCBI-nr database was used for read classification, as described in [27].

run diamond:

```
for file in *.fastq.gz; do name=${file/.fastq.gz/}; diamond blastx
-d /path/to/NCBI_nr/nr -q $file -a $name -p 16
convert binary DIAMOND format to BLAST tabular format:
for file in *.daa; do diamond view --daa $file --out
${file/.daa/}.tab --outfmt tab; echo $file; done
perform read-by-read taxonomy classification with MEGAN:
for file in *.tab; do /path/to/programs/megan/tools/blast2lca -- input
$file --format BlastTAB --topPercent 10 --gi2taxa
/path/to/programs/megan/GI_Tax_mapping/gi_taxid-March2015X.bin-- output
$file.read_assignments.txt; done
```

Heatmaps were generated with the script `metaphlan_hclust_heatmap.py` from the MetaPhlan package, displaying the abundances for species only (default `-tax_lev s`), in logarithmic scale (`-s log`). The clustering is performed with "average" linkage (default `-m average`), using "Bray-Curtis" distance for clades (default `-d braycurtis`) and "correlation" for samples (default `-f correlation`).

```
metaphlan_hclust_heatmap.py -in $file
-out $file.Blues.minv0.maxv1.Blues.log.pdf -c Blues -s log -minv 0.0 -maxv 1.
```

#### Diversity quantification

Beta-diversity was calculated according to the Bray-Curtis dissimilarity metric (Bray and Curtis 1957) as implemented by the Qiime2 package [28].

```
$ metaphlan2biom.py merged.samples.metaphlan.out merged.samples.biom
$ beta_diversity.py -i merged.samples.biom -m bray_curtis -o merged.samples.beta_div.bray_curtis
```

*P*-value was calculated based on 100 bootstrapped subsamples of the Brooklyn debris sample, each subsample being of 1 million reads. Bootstrapped samples were classified using the same methods as described above, and pairwise beta-diversity calculated as above. *P*-value was calculated as the number of bootstrap samples with lesser dissimilarity value than the test value.

#### Assembly and contig annotation

Co-assembly of Tokyo samples (assembly of all sequences pooled together) was performed with MegaHit [29] and reads for each individual sample were mapped to contigs with Bowtie2 [30]. Assembly yielded 3207501 contigs with a total of 2802811167 base pairs. Contig length ranged from 200 to 488034 base pairs, with an average of 874bp and an N50 of 1515bp. Contigs were annotated with Anvio [23].

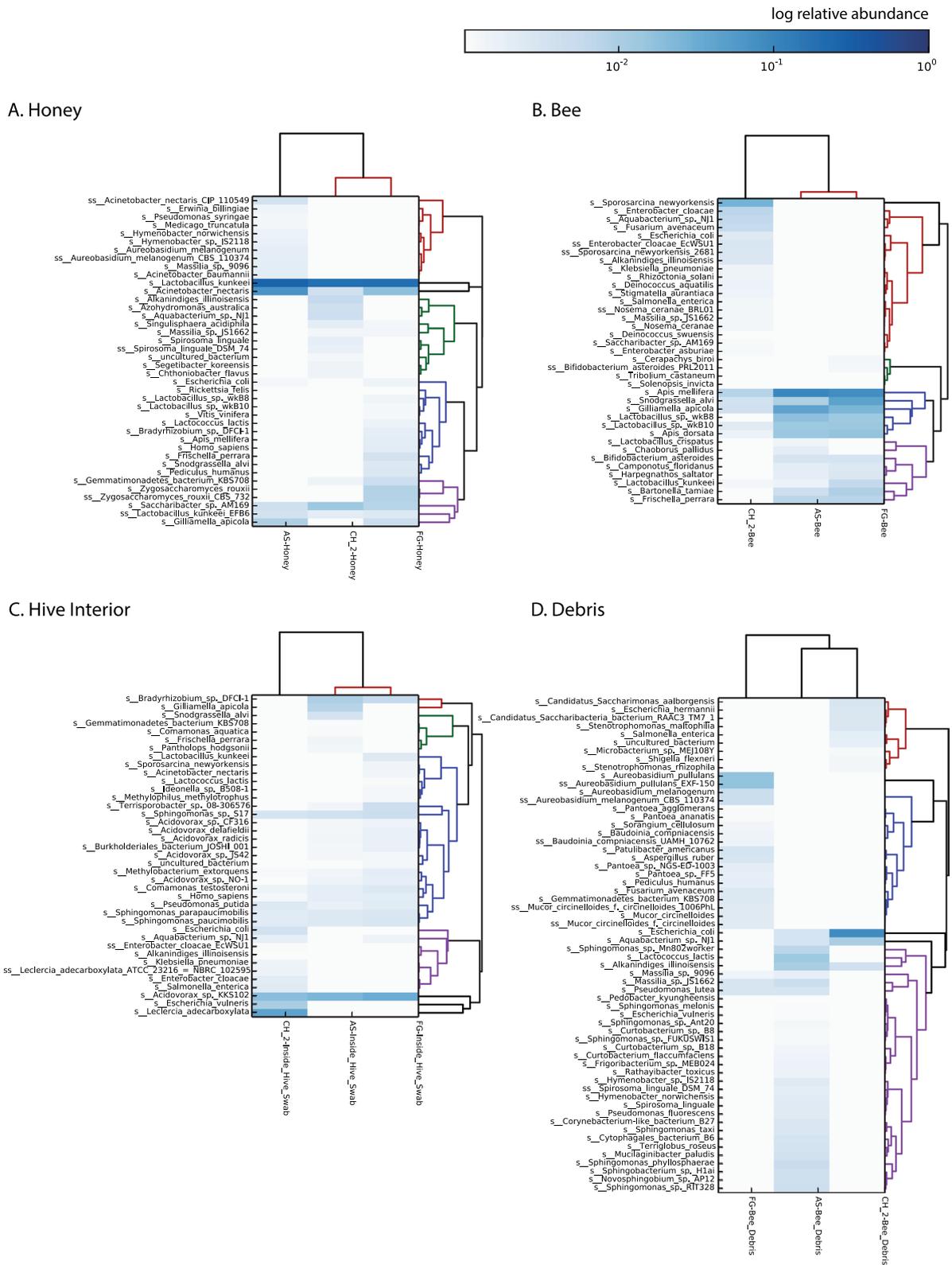
#### Virulence factor identification

Virulence factors for *Rickettsia felis* were downloaded from the Virulence Factors of Pathogenic Bacteria database <http://www.mgc.ac.cn/cgi-bin/VFs/compvfs.cgi>. BLAST [31] was used to align the virulence factor genes to the assembled contigs, reporting the query coverage and percent identity.

### Results

#### Brooklyn pilot study

In order to assess the potential of using honeybees as metagenomic "sample collectors", we designed a pilot study with three Langstroth hives in Brooklyn, wherein we sampled the interior of the hive, the debris at the bottom, bee bodies, and honey. We sequenced the DNA of each sample using a high-throughput shotgun approach, and classified the reads using DIAMOND-MEGAN against the NCBI NR nucleotide database, which includes all kingdoms and domains of life (see Methods



**Fig. 1** Species classification by type of sample in Brooklyn pilot study: **A** Honey, **B** Bee body, **C** Hive interior, **D** Debris. Hives are abbreviated as: AS Astoria, CH Crown Heights, FG Fort Greene. Color map scale corresponds to the log of relative abundance in each sample

**Table 1** Beta-diversity according to sample type (Bray–Curtis dissimilarity)

| Honey |      |      |      |      |         |
|-------|------|------|------|------|---------|
|       | AS   | CH   | FG   | mean | P-value |
| AS    | 0.00 | 0.81 | 0.63 | 0.71 | 0       |
| CH    | 0.81 | 0.00 | 0.70 |      |         |
| FG    | 0.63 | 0.70 | 0.00 |      |         |

| Bee |      |      |      |      |         |
|-----|------|------|------|------|---------|
|     | AS   | CH   | FG   | mean | P-value |
| AS  | 0.00 | 0.87 | 0.21 | 0.65 | 0       |
| CH  | 0.87 | 0.00 | 0.87 |      |         |
| FG  | 0.21 | 0.87 | 0.00 |      |         |

| Hive Interior |      |      |      |      |         |
|---------------|------|------|------|------|---------|
|               | AS   | CH   | FG   | mean | P-value |
| AS            | 0.00 | 0.59 | 0.28 | 0.47 | 0       |
| CH            | 0.59 | 0.00 | 0.54 |      |         |
| FG            | 0.28 | 0.54 | 0.00 |      |         |

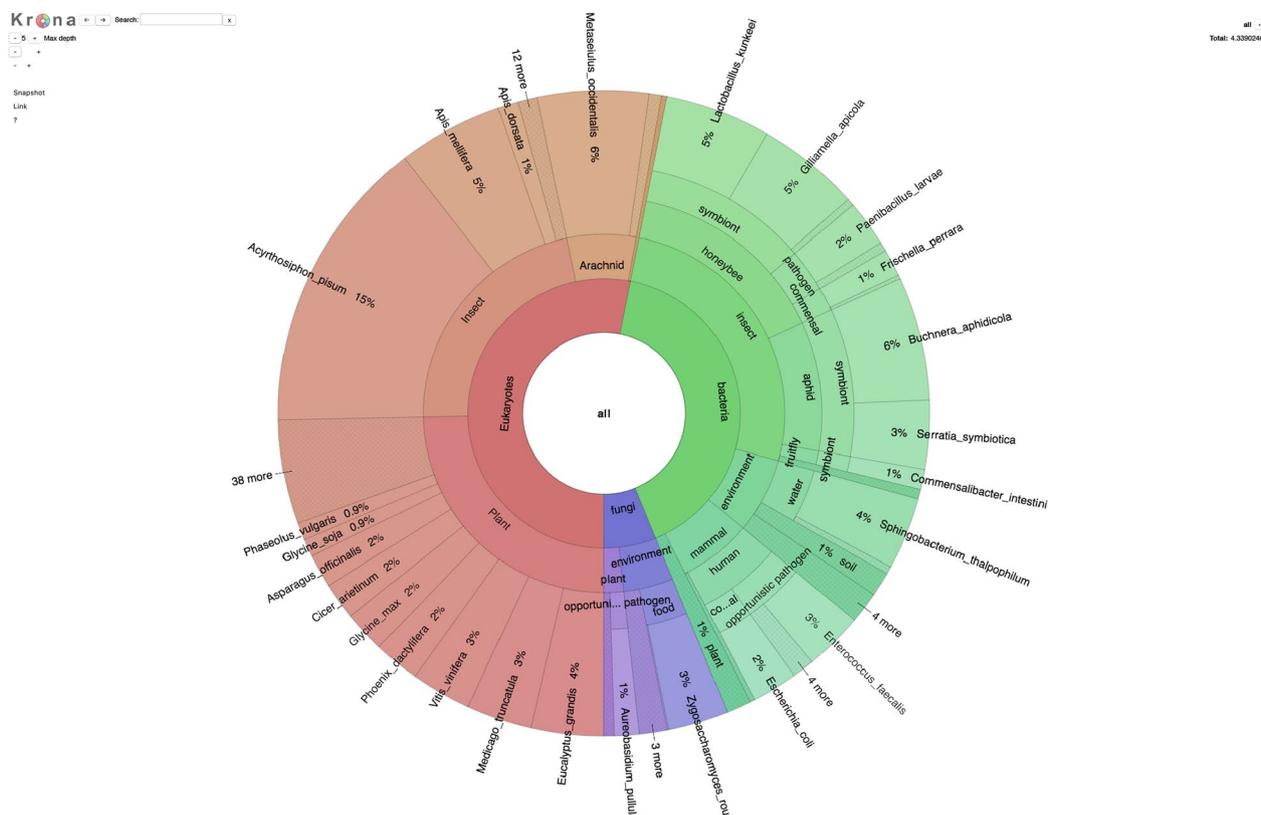
| Bee Debris |      |      |      |      |         |
|------------|------|------|------|------|---------|
|            | AS   | CH   | FG   | mean | P-value |
| AS         | 0.00 | 0.76 | 0.81 | 0.82 | 0       |
| CH         | 0.76 | 0.00 | 0.89 |      |         |
| FG         | 0.81 | 0.89 | 0.00 |      |         |

P-value calculated against 100 random subsamples of a debris sample. Hives are abbreviated as: AS Astoria, CH Crown Heights, FG Fort Greene

for more details) (Fig. 1). The honey of each hive is largely dominated by the species *Lactobacillus kunkeei* (Fig. 1A), an obligate fructophilic lactic acid bacteria found in flowers, wine, and honey [32]. Also of note are *Acinetobacter nectaris*, found in flowers [33], and *Zygosaccharomyces rouxii*, known to thrive under salt or sugar osmotic stress and thus cause food spoilage [34]. Bee gut commensals were found in low abundance in honey, and include the species identified in the bee body samples, described below. Traces of plant DNA were also identified, including *Medicago truncatula* and *Vitis vinifera*. The bee body samples (Fig. 1B) contain sequences representative of both *Apis mellifera* (European honeybee) and *Apis dorsata* (Giant honeybee), indicating the hives are likely hybrids of these two species. The most abundant microbes in the bee body samples include species described as bee commensals such as *Snodgrassella alvi* and *Gilliamella apicola* [35], as well as *Lactobacillus* *wkB8* and *wkB10* [34]. The bees from AS and FG hives display almost identical species distribution, however

the bees from the CH hive show lower abundances of the aforementioned commensals, and present species absent from the other two. These include *Nosema ceranae*, a fungal parasite of the honeybee affecting both larvae and adults [37], as well as various human-related bacteria such as *Sporosarcina newyorkensis*, isolated from clinical samples in New York State [38] and *Enterobacter* species. We hypothesize the colonization of atypical bacteria in this bee is correlated to the dysbiosis caused by *Nosema* infection.

The inside of the hives (Fig. 1C) was quite uniform across locations, and dominated by environmental bacterial species usually described as found in polluted environments. These include *Acidovorax* *sp. KKS102*, known to degrade biphenyl/polychlorinated biphenyls (PCBs) [39], *Sphingomonas* *sp. S17* [40], found in high-altitude Andean lakes and tolerant to high pH and desiccation. The interior of beehives is coated with propolis, a resinous substance including polyphenols from essential



**Fig. 2** Distribution among kingdoms of classified reads across all samples, including most abundant species in each category

**Table 2** Major classes of bacteria across samples

|             | AUSTRALIA         |       |       |       |       |       |       |       |       |       | TOKYO |       |       |       |       |       |       |       |       |       | VENICE |       | BROOKLYN |       |  |
|-------------|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------|-------|----------|-------|--|
|             | SYD1              | SH    | SYD3  | MEL1  | GK_H1 | GK_H2 | GK_H3 | MA_H2 | MA_H3 | MA_H4 | MA_H5 | MA_H6 | MI_H1 | MI_H2 | MR_H1 | MR_H2 | MR_H3 | MR_H4 | TS_H1 | VEN   | BK_AS  | BK_CH | BK_FG    |       |  |
| environment | built             | 0.001 | 0.006 | 0.01  | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.002 | 0      | 0     | 0.001    |       |  |
|             | food              | 0     | 0.004 | 0.002 | 0     | 0     | 0     | 0.001 | 0.036 | 0.046 | 0.019 | 0.002 | 0.002 | 0.026 | 0     | 0     | 0     | 0     | 0     | 0     | 0.013  | 0     | 0        |       |  |
|             | marine            | 0.011 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0      | 0     | 0        | 0     |  |
|             | plant             | 0.011 | 0     | 0.006 | 0.01  | 0     | 0     | 0     | 0     | 0     | 0.001 | 0     | 0.012 | 0.003 | 0     | 0     | 0     | 0     | 0     | 0     | 0.041  | 0.015 | 0        | 0.028 |  |
|             | sludge            | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.02   | 0     | 0        | 0     |  |
|             | soil              | 0.001 | 0.009 | 0.013 | 0     | 0.013 | 0     | 0     | 0     | 0     | 0     | 0     | 0.005 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0      | 0.018 | 0.014    | 0     |  |
| insect      | water             | 0.014 | 0.143 | 0.016 | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0      | 0     | 0        |       |  |
|             | wood              | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.003 | 0.004 | 0     | 0     | 0     | 0     | 0     | 0     | 0.003 | 0      | 0     | 0.005    |       |  |
|             | aphid             | 0     | 0     | 0     | 0     | 0     | 0.006 | 0     | 0.003 | 0.001 | 0.042 | 0.002 | 0     | 0.005 | 0.04  | 0.108 | 0.057 | 0.153 | 0     | 0     | 0      | 0     | 0        |       |  |
|             | fruitfly          | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.043 | 0      | 0     | 0        |       |  |
|             | honeybee          | 0     | 0     | 0     | 0.02  | 0.095 | 0.032 | 0.059 | 0.035 | 0     | 0.004 | 0.034 | 0     | 0.05  | 0.002 | 0.002 | 0.004 | 0     | 0     | 0.316 | 0      | 0     | 0        |       |  |
|             | monarch butterfly | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.016 | 0      | 0     | 0        |       |  |
| mammal      | wasp              | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.002 | 0     | 0     | 0.002 | 0     | 0     | 0.001 | 0     | 0     | 0     | 0      | 0     | 0        |       |  |
|             | dog               | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0     | 0.007 | 0      | 0     | 0        |       |  |
|             | human             | 0.011 | 0     | 0.002 | 0     | 0     | 0.023 | 0.002 | 0.126 | 0     | 0     | 0.009 | 0.002 | 0.004 | 0     | 0.002 | 0     | 0     | 0     | 0.001 | 0      | 0.003 | 0.088    |       |  |

This table summarizes the most abundant bacterial species (accounting for 90% of bacterial contribution to a sample's profile) according to their associated host or environment. Numbers are relative abundance (normalized to 1 for each sample). Colored cells indicate values over 0

oils and with a pH of 8.5 [41]. It is a strong antimicrobial, antifungal and antiviral agent [42] and therefore we hypothesize the presence of extremophile bacteria, and their similar distribution across hives, is a result of selection by the chemical properties of propolis. The species identified in the debris samples (Fig. 1 C) were the most diverse (Table 1), and include several species of plants as well as plant-associated microbes such as the fungus *Aureobasium pullulans*, also an opportunistic human pathogen [43], aquatic microbes such as the

alkane-degrading *Aquabacterium sp. NJ1* [44] and honeybee associated such as *Stenotrophomonas maltophilia* [45] (also known as an opportunistic mammalian pathogen [46]). Taken together, the samples cluster according to sample type, versus sample location (Additional file 1: Fig S1). As a control, we also sampled a beekeeper's hands and hive scraper tool (in one instance) as well as the hive exterior, and these samples were notably different than the debris as well (Additional file 1: Fig S1). The former control indicates that the signatures in the debris

**Table 3** Bee related species: known bee gut species, honey and bee crop species, pathogens, and parasites

|                |                                   | number of samples |
|----------------|-----------------------------------|-------------------|
| honey and crop | <i>Lactobacillus kunkeei</i>      | 5                 |
|                | <i>Saccharibacter</i> sp. AM169   | 2                 |
|                | <i>Frischella perrara</i>         | 5                 |
| bee gut        | <i>Gilliamella apicola</i>        | 9                 |
|                | <i>Lactobacillus</i> sp. wkB10    | 1                 |
|                | <i>Lactobacillus</i> sp. wkB8     | 1                 |
|                | <i>Snodgrassella alvi</i>         | 2                 |
|                | <i>Bifidobacterium asteroides</i> | 1                 |
| parasite       | <i>Varroa destructor</i>          | 3                 |
| pathogen       | <i>Paenibacillus larvae</i>       | 2                 |
|                | <i>Melissococcus plutonius</i>    | 2                 |

collected are not just from manipulation, and the latter indicates that the debris composition is not just from settling of material from the environment immediately exterior to the hive.

While samples from different hives within a sample type are significantly different from each other ( $P=0.0$ ) according to Bray–Curtis dissimilarity (Table 1), we found the debris samples to be the most diverse, as well as have the highest proportion of environmental bacteria. As our interest was to collect metagenomic information of the environment the bees traverse, rather than that of their hive, we concluded that bee debris is the best material for that purpose.

#### Urban metagenomes as seen by bees

We next sampled bee hive debris from four cities across the world: Venice, Italy; Sydney and Melbourne in Australia; several neighborhoods in Tokyo, Japan. Over all of these locations, we recovered DNA from plants, mammals, insects, arachnids, bacteria and fungi. Taken together, 53% of the classified reads were from multicellular organisms, and 47% from microorganisms. (Fig 2).

All metagenomes characterized show different signatures according to cities (Additional file 2: Fig S2), and have particularities that can be related to the identity of the city. The metagenome of the debris collected from the hive in Venice was largely dominated by fungi related to wood rot (Additional file 3: Fig S3), which is a common feature of the buildings, built on submerged wooden pilings, and date palm DNA. Melbourne's sample was dominated by Eucalyptus DNA, while Sydney's showed little plant DNA, but bacteria such as *Gordonia polyisoprenivorans*, which degrades rubber [47] (Additional file 4: Fig S4). Tokyo's metagenome includes plant DNA from Lotus and wild soybean, as well as the soy sauce fermenting yeast *Zygosaccharomyces rouxii* [34] (Additional file 5: Fig S5). Overall, each city has a unique

metagenomic signature as viewed by bees, with microbes coming from a variety of sources: environmental, insect-related, mammalian and aquatic (see Table 2 for relative abundances of bacteria associated with different hosts or environments).

#### Debris as indicator of hive health

As the debris include parts of bees, we looked to the data to see if we could find microbes related to bee health. We found three honey and bee crop related species such as *Lactobacillus kunkeei*, *Saccharibacter* sp. AM169 and *Frischella perrara* and five bee gut species, with *Gilliamella apicola* being found in the most samples (Table 3) [48]. We also identified known bee pathogens, namely *Paenibacillus larvae* and *Melissococcus plutonius*, as well as the parasite *Varroa destructor*. These results indicate that debris may be used to assess overall hive health, or to assess the interaction of bee related species with environmental microbial species.

#### Debris as indicator of human health

As the bees are traversing densely populated urban areas, we tested the hypothesis that they may be able to recover human pathogens and assess their pathogenic capacity by identifying virulence factor genes. Virulence factors are the molecules that enable the specific pathogenicity of the micro-organism [49]. Given the high level of genomic variation within species, asserting the presence of a pathogen through taxonomic classification is not sufficient to assert its pathogenicity. For this, we proceeded by performing de-novo co-assembly of the sequences from a given city, then using a metagenomic-specific classifier targeted to identify bacterial species from the contigs. We identified various opportunistic pathogens as well as some known disease-causing pathogens, including *Shigella dysenteriae* Sd197 (causing bacillary dysentery [50]) and *Rickettsia felis* (causing "cat scratch fever" [51]). We selected the Tokyo dataset for assembly as this location presented the highest number of samples, samples collected at two timepoints, as well as highest sequencing coverage per sample. We chose *Rickettsia felis* as an example to demonstrate the ability to identify a pathogen and its virulence factors with this sample collection method as it was the most represented in the assembled contigs. To go beyond species classification and assess pathogenic potential, we queried the assembled metagenome for *Rickettsia felis* virulence factor genes, as their presence is required for pathogenic capacity. We used *R. felis* as a proof-of-principle example that it is possible to verify pathogenic capacity of classified species with this type of data. In the Tokyo dataset, we recovered 28 of the 31 *Rickettsia felis* virulence genes with high coverage and

**Table 4** Alignment statistics of *Rickettsia felis* virulence factor genes mapped to assembled contigs of Tokyo metagenome

| gene name                 | gene length | Alignment % identity | Alignment % gene coverage | Subject contig | Subject contig length |
|---------------------------|-------------|----------------------|---------------------------|----------------|-----------------------|
| Accession YP_247329adr1   | 870         | 93.02                | 82.07                     | k99_3461957    | 17030                 |
| Accession YP_247330adr2   | 684         | 98.68                | 100.00                    |                |                       |
| Accession YP_246376pat1   | 1650        | 98.37                | 89.33                     | k99_5184779    | 4409                  |
| Accession YP_246759pat2   | 1113        | 98.75                | 101.08                    | k99_6299312    | 21131                 |
| Accession YP_247313pld    | 603         | 97.5                 | 99.50                     | k99_2470097    | 13853                 |
| Accession YP_246601ralF   | 1950        | 97.45                | 100.67                    | k99_802869     | 12965                 |
| Accession YP_246387rickA  | 1590        | 97.92                | 66.42                     | k99_4497423    | 11956                 |
| Accession YP_246483rvhB10 | 1446        | 98.8                 | 98.27                     | k99_2250851    | 6383                  |
| Accession YP_246484rvhB11 | 1005        | 98.31                | 100.00                    |                |                       |
| Accession YP_246819rvhB1  | 840         | 98.1                 | 100.00                    | k99_1949946    | 19589                 |
| Accession YP_247091rvhB2  | 360         | 99.72                | 100.00                    | k99_588303     | 7843                  |
| Accession YP_246103rvhB3  | 288         | 96.88                | 100.00                    | k99_2306863    | 24667                 |
| Accession YP_246104rvhB4a | 2418        | 99.21                | 100.00                    | k99_9338833    | 19842                 |
| Accession YP_247266rvhB4b | 2433        | 99.14                | 100.00                    | k99_758254     | 54842                 |
| Accession YP_246105rvhB6a | 3417        | 98.1                 | 100.00                    | k99_9338833    | 19842                 |
| Accession YP_246106rvhB6b | 2019        | 98.09                | 95.99                     |                |                       |
| Accession YP_246107rvhB6c | 2934        | 98.29                | 99.73                     |                |                       |
| Accession YP_246108rvhB6d | 2673        | 98.13                | 100.22                    |                |                       |
| Accession YP_246109rvhB6e | 3468        | 98.7                 | 100.00                    |                |                       |
| Accession YP_246480rvhB7  | 180         | 99.44                | 100.00                    | k99_7282718    | 13219                 |
| Accession YP_246479rvhB8a | 699         | 98.86                | 100.00                    |                |                       |
| Accession YP_246481rvhB8b | 732         | 99.18                | 100.00                    |                |                       |
| Accession YP_246478rvhB9a | 753         | 98.94                | 100.00                    |                |                       |
| Accession YP_246482rvhB9b | 477         | 99.58                | 100.00                    |                |                       |
| Accession YP_246485rvhD4  | 1776        | 98.48                | 100.00                    | k99_2250851    | 6383                  |
| Accession YP_246741sca4   | 3120        | 94.57                | 96.22                     | k99_796908     | 24425                 |
| Accession YP_247201tlyC   | 900         | 99.11                | 100.00                    | k99_7586727    | 10516                 |

at high similarity on the nucleotide level (Table 4). While co-assembly of these complex metagenomes led to less than optimal N50 values (N50=1515bp), this assembly quality was sufficient for virulence factor gene identification, as the genes tested for *Rickettsia felis* were covered over 97% of their length on average (Table 5) when aligned to the assembled contigs.

We assessed the persistence of virulence factors in the debris by analyzing samples taken at a 1-week interval in the Tokyo hives. After the first sampling, the bottom trays were cleaned and debris was collected after a week. In some cases, no markers were observed in the second samples, indicating that the cleaning was effective. In the Marunouchi hive H2, markers were found again, and more abundantly (Table 5). This indicates virulence markers that are either very abundant in the bee's range or that they can change rapidly in abundance.

## Discussion

Here we show that honeybees are relevant sensors for the urban microbiome, and that the debris collected contain a trace of the microbial clouds the bees are traversing as well as carry indicators of hive health. While these methods are cost prohibitive for amateur or even professional beekeepers as pathogen detection, and existing targeted

methods already exist, these results present a methodology to assess additional dimensions of hive health. Indeed, we show that bees interact with a wide range of microbial species and thus future apiculture research could consider individual hive health in relation to the bees' microbial environment, exploiting for example existing databases and scripts describing bee-associated bacteria [52]. Indeed, these bees recover microbes associated with plants, with which they have physical interactions, but also of mammals and aquatic environments, with which they presumably do not have direct contact. This implies that these microbes were constituents of the respective "microbial clouds" [53] of these entities and that the bees collect a trace of these clouds. Biological content in the atmosphere—the biosphere—was first described in 1978 [54] and has since been characterized as an integral part of ecosystem function [55]. The biosphere is an indicator of climate change, for example, increasing frequency of dust storms from the African continent are carrying plant and aquatic pathogens to the Americas, affecting coral populations [56]. Urban aerosols contain a diverse microbial component including species of potential health and bioterrorism concern. This study demonstrates a novel sampling methodology, with consistent results with a recent study using shotgun

**Table 5** Abundance of virulence factors in samples collected at 1-week interval in Tokyo

| Location   | Hive number | Abundance of virulence marker genes (RPM)  |        |
|------------|-------------|--|--------|
|            |             | Week 1   | Week 2 |
| Ginza Kami | H1          |  1.47 | 0.00   |
|            | H3          | 0.00   | 0.00   |
| Marunouchi | H1          |  2.85 | 0.11   |
|            | H2          | 0.00   | 0.44   |
|            | H4          |  3.17 | 0.07   |
|            | H6          |  2.82 | 0.77   |
| Mita       | H2          |  3.02 | 0.00   |
| Marronnier | H2          |  7.91 | 22.53  |
|            | H3          | 0.29   | 0.33   |
|            | H4          |  1.29 | 0.00   |

Numbers reported are Reads per Million. Colored bars represent numerical values of each cell

sequencing of honey to assess bee core gut microbiomes as well as plant species interaction while foraging [57], while also providing additional environmental microbiome data than the honey substrate. This reveals that different neighborhoods have different clouds just as different humans do, and that the collected microbiome can reveal information about the built environment and its inhabitants. For example, the Venetian bees carried a signature of wood rot and aquatic species, similar to previous work showing how flooded areas of a city can carry a “molecular echo” of the aquatic events of its past [14]. Indeed, it has been shown that microbial communities can serve as quantitative geochemical indicators [58] and the metabolic properties of the recovered communities can yield information about the environment. Furthermore, metagenomic data can be mined for human-health related information [59]. Future uses of data collected in this manner could be assessment of antibiotic resistance gene profiles, and while the molecular and computational methods used here were based on DNA analysis, it is possible they could be used to monitor RNA-based viruses such as Sars-Cov-2 or other future airborne pathogens, as demonstrated by targeted analyses using swab-based collection at hive doors during the COVID19 global pandemic [60].

## Conclusions

Our ability to recover virulence factors associated with human disease indicates that this method can serve for early detection of human-associated pathogens, in a complimentary modality to existing biosurveillance methods such as indoor air or sewage monitoring.

However, this multi-species methodological approach may hold even more hope for a diversified understanding of urban microbiomes, their relationship to the built environment, and their relationship to human and other non-human species. Indeed, insect-based, city-wide microbial monitoring is likely more spatially comprehensive, even if lower resolution, compared to discrete, human-based sampling techniques, such as swabbing or air-sampling. This method offers the capacity to further catalog the urban environmental microbiome, contributing information to our understanding of its impact on humans. Additionally, this methodology offers a framework to understand multispecies interactions in the built environment, namely understanding hive health in the context of the microbiome of the bees’ foraging range.

We have the unique possibility to understand our built environment and therefore design it, not just for ourselves but for all its inhabitants, from environments as common and public as subways [61] to those as specialized and hermetic as space stations [62, 63]. As Jane Jacobs says, “Cities are an immense laboratory of trial and error, failure and success, in city planning and city design” [64]. Through studies such as the one presented here, and using interdisciplinary approaches including art practice [65], we aim to further understand this accidentally engineered multispecies experiment of our built, shared, environment.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40793-023-00467-z>.

**Additional file 1.** Clustered heatmap of Brooklyn pilot samples including hive debris, bee bodies, honey, propolis, swabs of the hive structure as well as the beekeepers’ hands.

**Additional file 2.** Clustered heatmap of hive debris samples from USA, Italy, Australia and Japan.

**Additional file 3.** Heatmap of Venice (Italy) hive debris samples.

**Additional file 4.** Heatmap of Sydney and Melbourne (Australia) hive debris samples.

**Additional file 5.** Heatmap of Tokyo (Japan) hive debris samples.

**Additional file 6.** Read counts for all samples.

**Additional file 7.** Metadata for all samples in MIXS format.

**Additional file 8.** MultiQC reports for samples by country.

## Acknowledgements

We would like to thank all the beekeepers for so generously supporting this project with their time and materials. We would also like to thank Mike Laserwalker and Ben Berman for their assistance in the 2016 Venice Biennale exhibit, as well as Daniela Bezdán for her support in sequencing preparation of the Venice sample.

## Author contributions

DN and MP located the hives and took samples. MP designed custom trays at MIT Media Lab for sampling. DN developed specific DNA extraction protocols for bee material. EH and DN prepared libraries for sequencing at Cooper Union

and Weill Cornell Medicine and EH analyzed the data and generated the data visualizations in the manuscript. For the installation for the Venice Biennale exhibit, RF and MP developed the data visualization and MP and CW designed the physical installation. KS and CM supervised the project and helped design experiments and provided logistical support. EH drafted the manuscript, with help from RF and DN and edits from CM. All authors have read and approved the manuscript.

#### Funding

The authors would like to acknowledge the Mori Building Company for their financial support, and Jun Fujiwara especially for his continued interest in the project. CEM would like to thank the Epigenomics Core Facility, the Vallee Foundation, Igor Tulchinsky and the WorldQuant Foundation, the National Institutes of Health (1R01MH117406), the Bill and Melinda Gates Foundation (OPP1151054), the NSF (1840275), and the Alfred P. Sloan Foundation (G-2015–13964).

#### Availability of data and materials

The datasets generated during the current study are available in the National Center for Biotechnology Information under BioProject accession PRJNA630108 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA630108>) and the Sequence Read Archive repository under the accession SRP259669 (<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP259669>).

#### Declarations

##### Ethical approval and consent to participate

Not Applicable.

##### Consent for publication

Not applicable.

##### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>NYU Tandon School of Engineering, Brooklyn, NY, USA. <sup>2</sup>Center for Urban Science and Progress, NYU, Brooklyn, NY, USA. <sup>3</sup>MIT Media Lab, Cambridge, MA, USA. <sup>4</sup>Parsons School of Design, New York, NY, USA. <sup>5</sup>Extrapolation Factory, New York, NY, USA. <sup>6</sup>Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY, USA. <sup>7</sup>Weill Cornell Medicine, The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, New York, NY, USA. <sup>8</sup>The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY, USA.

Received: 30 May 2022 Accepted: 23 January 2023

Published online: 30 March 2023

#### References

- United Nations, Department of economic and social affairs, and population division, World urbanization prospects: the 2018 revision. 2019.
- J A Puppim de Oliveira et al. Cities, biodiversity and governance: perspectives and challenges of the implementation of the convention on biological diversity at the city level. 2010.
- Kembel SW, et al. Architectural design drives the biogeography of indoor bacterial communities. *PLoS ONE*. 2014;9(1):1–10. <https://doi.org/10.1371/journal.pone.0087093>.
- Ruiz-Calderon JF, Cavallin H, Song SJ, Novoselac A, Pericchi LR, Hernandez JN, Rios R, Branch OH, Pereira H, Paulino LC, Blaser MJ. Walls talk: microbial biogeography of homes spanning urbanization. *Sci Adv*. 2016;2(2):e1501061. <https://doi.org/10.1126/sciadv.1501061>.
- Sharma A, Gilbert JA. Microbial exposure and human health. *Curr Opin Microbiol*. 2018;44:79–87. <https://doi.org/10.1016/j.mib.2018.08.003>.
- McFall-Ngai MJ. Unseen forces: the influence of bacteria on animal development. *Dev Biol*. 2002;242(1):1–14. <https://doi.org/10.1006/dbio.2001.0522>.
- McFall-Ngai M, et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA*. 2013;110(9):3229–36. <https://doi.org/10.1073/pnas.1218525110>.
- Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*. 2006;444(7122):1027–31. <https://doi.org/10.1038/nature05414>.
- Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat Rev Neurosci*. 2012;13(10):701–12. <https://doi.org/10.1038/nrn3346>.
- Wagner MR, Lundberg DS, Coleman-Derr D, Tringe SG, Dangl JL, Mitchell-Olds T. Natural soil microbes alter flowering phenology and the intensity of selection on flowering time in a wild Arabidopsis relative. *Ecol Lett*. 2014;17(6):717–26.
- Köberl M, Schmidt R, Ramadan EM, Bauer R, Berg G. The microbiome of medicinal plants: diversity and importance for plant growth, quality and health. *Front Microbiol*. 2013;4:400. <https://doi.org/10.3389/fmicb.2013.00400>.
- Zilber-Rosenberg I, Rosenberg E. Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiol Rev*. 2008;32(5):723–35. <https://doi.org/10.1111/j.1574-6976.2008.00123.x>.
- Bordenstein SR, Theis KR. Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol*. 2015;13(8):e1002226. <https://doi.org/10.1371/journal.pbio.1002226>.
- Afshinnekoo E, et al. Geospatial resolution of human and bacterial diversity with city-scale metagenomics. *Cell Syst*. 2015. <https://doi.org/10.1016/j.cels.2015.01.001>.
- Hsu T, Joice R, Vallarino J, Abu-Ali G, Hartmann EM, Shafquat A, DuLong C, Baranowski C, Gevers D, Green JL, Morgan XC. Urban Transit system microbial communities differ by surface type and interaction with humans and the environment. *Msystems*. 2016;1(3):e00018–e116. <https://doi.org/10.1128/mSystems.00018-16>.
- Mason C, et al. The metagenomics and metadesign of the subways and urban biomes (MetaSUB) international consortium inaugural meeting report. *Microbiome*. 2016;4(1):24. <https://doi.org/10.1186/s40168-016-0168-z>.
- Lax S, et al. Bacterial colonization and succession in a newly opened hospital. *Sci Transl Med*. 2017;9(391):1–12. <https://doi.org/10.1126/scitranslmed.aah6500>.
- Meadow JF, et al. Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. *Indoor Air*. 2014;24(1):41–8. <https://doi.org/10.1111/ina.12047>.
- Maritz JM, et al. An 18S rRNA workflow for characterizing protists in sewage, with a focus on zoonotic trichomonads. *Microb Ecol*. 2017. <https://doi.org/10.1007/s00248-017-0996-9>.
- Newton RJ, McLellan SL, Dila DK, Vineis JH, Morrison HG, Eren AM, Sogin ML. Sewage reflects the microbiomes of human populations. *MBio*. 2015;6(2):e02574–e2614. <https://doi.org/10.1128/mBio.02574-14>.
- Eckert JE. The flight range of the honeybee. *J Agric Res*. 1931;47(5):257–85. <https://doi.org/10.2307/1541>.
- Garbuzov M, Schürch R, Ratnieks FLW. Eating locally: dance decoding demonstrates that urban honey bees in Brighton, UK, forage mainly in the surrounding urban area. *Urban Ecosyst*. 2015;18(2):411–8. <https://doi.org/10.1007/s11252-014-0403-y>.
- “Babraham bioinformatics - FastQC a quality control tool for high throughput sequence data.” <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed Jan. 20, 2023).
- Kang DD, Sibille E, Kaminski N, Tseng GC. MetaQC: objective quality control and inclusion/exclusion criteria for genomic meta-analysis. *Nucleic Acids Res*. 2012;40(2):e15. <https://doi.org/10.1093/nar/gkr1071>.
- Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using diamond. *Nat Methods*. 2014;12(1):59–60. <https://doi.org/10.1038/nmeth.3176>.
- Huson DH, Auch AF, Qi J, Schuster SC. Megan analysis of metagenomic data. *Genome Res*. 2007;17(3):377–86. <https://doi.org/10.1101/gr.5969107>.
- McIntyre ABR, et al. Comprehensive benchmarking and ensemble approaches for metagenomic classifiers. *Genome Biol*. 2017;18(1):182. <https://doi.org/10.1186/s13059-017-1299-7>.

28. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnol.* 2019;37(8):852–7. <https://doi.org/10.1038/s41587-019-0209-9>.
29. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015;31(10):1674–6. <https://doi.org/10.1093/bioinformatics/btv033>.
30. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012;9(4):357–9. <https://doi.org/10.1038/nmeth.1923>.
31. Camacho C, et al. BLAST+: architecture and applications. *BMC Bioinformatics.* 2009;10:421. <https://doi.org/10.1186/1471-2105-10-421>.
32. Endo A, Irisawa T, Futagawa-Endo Y, Takano K, du Toit M, Okada S, Dicks LM. Characterization and emended description of *Lactobacillus kunkeei* as a fructophilic lactic acid bacterium. *Int J Syst Evolution Microbiol.* 2012;62(Pt3):500–4. <https://doi.org/10.1099/ijs.0.031054-0>.
33. Álvarez-Pérez S, Lievens B, Jacquemyn H, Herrera CM. *Acinetobacter nectaris* sp. nov. and *Acinetobacter boissieri* sp. nov., isolated from floral nectar of wild Mediterranean insect-pollinated plants. *Int J Syst Evol Microbiol.* 2013;63(PART4):1532–9. <https://doi.org/10.1099/ijs.0.043489-0>.
34. Pribylova L, de Montigny J, Sychrova H. Osmoresistant yeast *Zygosaccharomyces rouxii*: the two most studied wild-type strains (ATCC 2623 and ATCC 42981) differ in osmotolerance and glycerol metabolism. *Yeast Chichester Engl.* 2007;24(3):171–80. <https://doi.org/10.1002/yea.1470>.
35. Kwong WK, Moran NA. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family Neisseriaceae of the *Betaproteobacteria*, and *Gilliamella apicola* gen. nov., sp. nov., a memb. *Int J Syst Evol Microbiol.* 2013;63(PART6):2008–18. <https://doi.org/10.1099/ijs.0.044875-0>.
36. W K Kwong, A L Mancenido, and A Moran "Members of the firm-5 clade , from honey bee guts." 2(6), 5–6, 2014, <https://doi.org/10.1128/genomeA.01176-14.Copyright>.
37. Eiri DM, Suwannapong G, Endler M, Nieh JC. *Nosema ceranae* can infect honey bee larvae and reduces subsequent adult longevity. *PLoS ONE.* 2015;10(5):1–17. <https://doi.org/10.1371/journal.pone.0126330>.
38. Wolfgang WJ, et al. *Sporosarcina newyorkensis* sp. nov. from clinical specimens and raw cow's milk. *Int J Syst Evol Microbiol.* 2012;62(2):322–9. <https://doi.org/10.1099/ijs.0.030080-0>.
39. Ohtsubo Y, Maruyama F, Mitsui H, Nagata Y, Tsuda M. Complete genome sequence of *acidovorax* sp. strain KKS102, a polychlorinated-biphenyl degrader. *J Bacteriol.* 2012;194(24):6970–1. <https://doi.org/10.1128/JB.01848-12>.
40. Fariás ME, et al. Genome sequence of *Sphingomonas* sp. S17, isolated from an alkaline, hyperarsenic, and hypersaline volcano-associated lake at high altitude in the Argentinean Puna. *J Bacteriol.* 2011;193(14):3686–7. <https://doi.org/10.1128/JB.05225-11>.
41. Marcucci MC. Propolis : chemical composition, biological properties and therapeutic activity. *Apidologie.* 1994;26(1):83–99. <https://doi.org/10.1051/apido:19950202>.
42. Kujumgieva A, Tsvetkovaa I, Serkedjievaa Y, Bankovav V, Christovb R, Popovb S. Antibacterial, antifungal and antiviral activity of propolis of different geographic origin. *J Ethnopharmacol.* 1999;64(3):235–40.
43. Bolignano G, Criseo G. Disseminated nosocomial fungal infection by *Aureobasidium pullulans* var. *melanigenum*: a case report. *J Clin Microbiol.* 2003;41(9):4483–5. <https://doi.org/10.1128/JCM.41.9.4483-4485.2003>.
44. Masuda H, Shiwa Y, Yoshikawa H, Zylstra GJ. Draft genome sequence of the versatile alkane-degrading bacterium *Aquabacterium* sp. strain NJ1. *Genome Announc.* 2014;2(6):7–8. <https://doi.org/10.1128/genomeA.01271-14>.
45. Evans JD, Armstrong T-N. Antagonistic interactions between honey bee bacterial symbionts and implications for disease. *BMC Ecol.* 2006;6:4. <https://doi.org/10.1186/1472-6785-6-4>.
46. Brooke JS. *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen. *Clin Microbiol Rev.* 2012;25(1):2–41. <https://doi.org/10.1128/CMR.00019-11>.
47. Linos A, Steinbüchel A, Spröer C, Kroppenstedt RM. *Gordonia polyisoprenivorans* sp nov a rubber-degrading actinomycete isolated from an automobile tyre. *Int J Syst Evolution Microbiol.* 1999;49(4):1785–91.
48. Engel P, Martinson VG, Moran NA. Functional diversity within the simple gut microbiota of the honey bee. *Proc Nat Acad Sci.* 2012;109(27):11002–7. <https://doi.org/10.1073/pnas.1202970109>.
49. Sharma AK, et al. Bacterial virulence factors: secreted for survival. *Indian J Microbiol.* 2017;57(1):1–10. <https://doi.org/10.1007/s12088-016-0625-1>.
50. Yang F, et al. Genome dynamics and diversity of *Shigella species*, the etiologic agents of bacillary dysentery. *Nucleic Acids Res.* 2005;33(19):6445–58. <https://doi.org/10.1093/nar/gki954>.
51. Yazid Abdad M, Stenos J, Graves S. *Rickettsia felis*, an emerging flea-transmitted human pathogen. *Emerg Health Threats J.* 2011;4(1):7168. <https://doi.org/10.3402/ehth.v4i0.7168>.
52. Ellegaard KM, Suenami S, Miyazaki R, Engel P. Vast differences in strain-level diversity in the Gut microbiota of two closely related honey bee species. *Curr Biol CB.* 2020;30(13):2520–2531.e7. <https://doi.org/10.1016/j.cub.2020.04.070>.
53. Meadow JF, Altrichter AE, Bateman AC, Stenson J, Brown GZ, Green JL, Bohannan BJ. Humans differ in their personal microbial cloud. *PeerJ.* 2015;22(3):e1258. <https://doi.org/10.7717/peerj.1258>.
54. Imshenetsky AA, Lysenko SV, Kazakov GA. Upper boundary of the biosphere. *Appl Environ Microbiol.* 1978;35(1):1–5.
55. Burrows SM, Elbert W, Lawrence MG, Pöschl U. Bacteria in the global atmosphere – Part 1: review and synthesis of literature data for different ecosystems. *Atmospheric Chem Phys.* 2009;9(23):9263–80. <https://doi.org/10.5194/acp-9-9263-2009>.
56. Behzad H, Mineta K, Gojobori T. Global ramifications of dust and sand-storm microbiota. *Genome Biol Evol.* 2018;10(8):1970–87. <https://doi.org/10.1093/gbe/evy134>.
57. Galanis A, et al. Bee foraging preferences, microbiota and pathogens revealed by direct shotgun metagenomics of honey. *Mol Ecol Resour.* 2022;22(7):2506–23. <https://doi.org/10.1111/1755-0998.13626>.
58. Smith MB, Rocha AM, Smillie CS, Olesen SW, Paradis C, Wu L, Campbell JH, Fortney JL, Mehlhorn TL, Lowe KA, Earles JE. Natural Bacterial communities serve as quantitative geochemical biosensors. *MBio.* 2015;6(3):e00326–e415. <https://doi.org/10.1128/mBio.00326-15.Editor>.
59. Rosenfeld JA, Reeves D, Brugler MR, Narechania A, Simon S, Durrett R, Foon X, Shianna K, Schatz MC, Gandara J, Afshinneko E. Genome assembly and geospatial phylogenomics of the bed bug *Cimex lectularius*. *Nat Commun.* 2016;7(1):10164. <https://doi.org/10.1038/ncomms10164>.
60. Cilia G, Bortolotti L, Albertazzi S, Ghini S, Nanetti A. Honey bee (*Apis mellifera* L.) colonies as bioindicators of environmental SARS-CoV-2 occurrence. *Sci Total Environ.* 2022;20(805):150327. <https://doi.org/10.1016/j.scitotenv.2021.150327>.
61. Danko D, et al. A global metagenomic map of urban microbiomes and antimicrobial resistance. *Cell.* 2021;184(13):3376–3393.e17. <https://doi.org/10.1016/j.cell.2021.05.002>.
62. N K Singh, D Bezdán, A Checinska Sielaff, K Wheeler, C E Mason, and K Venkateswaran "Multi-drug resistant enterobacter bugandensis species isolated from the international space station and comparative genomic analyses with human pathogenic strains." *BMC Microbiol.* <https://doi.org/10.1186/s12866-018-1325-2>.
63. Garrett-Bakelman FE, Darshi M, Green SJ, Gur RC, Lin L, Macias BR, McKenna MJ, Meydan C, Mishra T, Nasrini J, Piening BD. The NASA twins study: a multidimensional analysis of a year-long human spaceflight. *Science.* 2019;364(6436):8650. <https://doi.org/10.1126/science.aau8650>.
64. Jacobs J. *The death and life of great American cities.* New York: Vintage Books; 1992.
65. "Time space existence - biennale architettura 2016, VENEZIA, Italy," google arts & culture. <https://artsandculture.google.com/partner/time-space-existence-biennale-architettura-2016> (accessed Jan. 20, 2023).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.