Metabarcoding of honey to assess differences in plant-pollinator interactions between urban and non-urban sites

Kay LUCEK, Anaïs GALLI, Sabrina GURTEN, Nora HOHMANN, Alessio MACCAGNI, Theofania PATSIU, Yvonne WILLI

Department of Environmental Sciences, University of Basel, Schönbeinstrasse 6, 4056, Basel, Switzerland

Received 5 September 2018 – Revised 15 February 2019 – Accepted 19 March 2019

Abstract – Ever-faster rates of urbanization have led to the substantial loss of agricultural land and natural areas. Furthermore, vegetated land is often fragmented and dominated by nonnative plant species within urban areas, which is likely to impact plant-pollinator interactions. However, the study of plant-pollinator interactions is labor intensive and remains challenging. Here, we compared plant species visited by honeybees in urban versus non-urban sites, using metabarcoding to identify pollen in honey samples. We identified a total of 262 plant genera, 15 of which were represented predominantly. Samples varied greatly in terms of plant composition, the number of nonnative genera and the number of crop plant genera. Even though land cover differed between urban and non-urban sites, plant diversity in the honey did not. We discuss our results in the light of challenges associated with the spatial scale of bee activity, preference behavior in flower visitation, and statistical power.

Apis mellifera / urban ecology / Shannon index / metabarcoding

1. INTRODUCTION

An important fraction of change in land use is caused by the accelerated growth of urban areas at the expense of agricultural land and natural areas (Grimm et al. 2008; Concepción et al. 2015; Harrison and Winfree 2015). One of the consequences is that organismal communities change during or after the transition from a non-urban to an urban state. The latter is for example characterized by the predominance of generalist and often nonnative plant species that are able to ex-

ploit the diverse and novel urban environments more effectively (Lososová et al. 2011; Concepción et al. 2015). A particular challenge that native outcrossing plant species face is the lack of pollinators, which prevents their establishment and spread in urban areas (Cheptou and Avendaño 2006; Geslin et al. 2013). Patterns of plant-pollinator interactions are yet understudied in urban as opposed to more natural or agricultural environments (Baldock et al. 2015; Harrison and Winfree 2015). By using a novel metabarcoding approach and taking advantage of the increasing popularity of urban beekeeping over the last decade, we assessed the utility of genetic pollen analyses of honey to study patterns of plant-pollinator interactions of honeybees (*Apis mellifera*) and plant diversity in urban areas.

Urban environments are generally characterized by an increased loss of habitat, habitat fragmentation, and exposure to higher temperatures or pollutants, all of which are predicted to affect

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s13592-019-00646-3>) contains supplementary material, which is available to authorized users.

Corresponding author: K. Lucek, kay.lucek@unibas.ch
 Manuscript editor: Marina Meixner
 Anaïs Galli, Sabrina Gurten, Nora Hohmann, Alessio Maccagni, and Theofania Patsiou are in alphabetical order

single plant species, plant communities, and plant-pollinator interactions (Pickett et al. 2011; Concepción et al. 2015; Harrison and Winfree 2015). Habitat fragmentation may for example affect plant-pollinator interactions because pollinators avoid travelling to small and isolated habitat fragments that provide less food (Harrison and Winfree 2015). The resulting limitation of pollinators selects for selfing over outcrossing species (Cheptou and Avendaño 2006) and contributes to a loss of diversity in native plant species. However, urban plant communities tend to have a higher proportion of nonnative species, many of which are planted as ornamentals in parks and gardens (Pickett et al. 2011). At high abundance, nonnative species impact plant-pollinator interactions by changing the seasonal availability of foraging resources for pollinators (Harrison and Winfree 2015). The plants that pollinators visit in urban as opposed to non-urban areas are therefore expected to differ given the changes in the composition of the vegetation, driven by the introduction of some plant species and the loss of others.

Traditional methods used to identify the plants visited by pollinators involve either the direct observations of plant visitations in the field or indirect methods, such as analyzing pollen obtained with pollen traps or, for honeybees, extracted from honey. Pollen is either actively collected by bees to provide larval food and energy during the winter season or accumulated passively during nectar collection (Thorp 2000; Grimm et al. 2008; Concepción et al. 2015). Both the direct observation and traditional pollen analyses by light microscopy are time consuming. In addition, pollen analyses require in-depth taxonomic knowledge while yielding often a limited taxonomic resolution, i.e., providing information only for the most common taxa. Moreover, pollen analyses are often restricted to higher taxonomic levels, such as the family or genus level (Lososová et al. 2011; Concepción et al. 2015; Bell et al. 2016a; Sniderman et al. 2018). Novel sequencing technologies, particularly metabarcoding, may overcome these obstacles (Cheptou and Avendaño 2006; Geslin et al. 2013; Bell et al. 2016b), enabling the assessment of biodiversity at an unprecedented scale and efficiency (Deiner et al. 2017; Geldmann and González-

Varo 2018). Pollen metabarcoding has been technically validated (Keller et al. 2014; Sickel et al. 2015; Baldock et al. 2015; Bell et al. 2016b) and found to provide a higher accuracy of pollen identification at all taxonomic levels as well as a higher taxonomic resolution, i.e., allowing the identification of species (Baldock et al. 2015; Smart et al. 2017). Moreover, no a priori taxonomic knowledge is needed and large pollen samples can be analyzed, providing the opportunity for more in-depth studies (Wilson et al. 2010; Hawkins et al. 2015; Richardson et al. 2015; Bell et al. 2016b; Smart et al. 2017).

We investigated whether barcoding of pollen in honey could be used to assess differences in plant-pollinator interactions and plant diversity between urban and non-urban sites. Honeybees have become urban drivers of plant-pollinator interactions due to the recent rise of urban beekeeping, promoting pollination while also negatively impacting wild bee species through competition (Geldmann and González-Varo 2018). Honeybees in urban sites were found to be more productive and to have a broader range of plants to forage on than those in non-urban sites (Baldock et al. 2015). Whereas both classical and metabarcoding approaches for pollen identification are increasingly used to study honeybee foraging in natural or agricultural systems (reviewed in Bell et al. 2016b, 2017), there are relatively few studies that compared foraging of honeybees between urban and non-urban sites and they relied predominantly on classical pollen identification methods (Baldock et al. 2015). Using a next-generation sequencing metabarcoding approach, we assessed the diversity of plants visited by honeybees by sequencing pollen extracted from commercially available honey samples from replicate urban and non-urban sites. We subsequently correlated our estimates of plant diversity with differences in land cover between urban and non-urban sites. We predicted that the diversity of flowering plants visited by honeybees was higher in urban than in non-urban sites, in line with previous studies that had found an increased plant species richness in urban compared to non-urban sites (Kühn et al. 2004) and a greater number of plant species that pollinators visited (Baldock et al. 2015). Given the turnover in plant communities from non-urban to urban sites (Dolan et al. 2011), and the often increased abundance of nonnative plant species in urban areas (Kühn et al. 2004; Dolan et al. 2011;

Baldock et al. 2015; Harrison and Winfree 2015), we further predicted that honeybees visited more nonnative plant species and fewer agriculturally cultivated plant species in urban than in non-urban areas.

2. MATERIAL AND METHODS

2.1. Sample collection

Seven honey samples from urban sites and seven samples from non-urban, predominantly agricultural sites were obtained from local producers of three regions of Switzerland (Figure 1, Table 1). All honey samples had been produced and collected in spring 2016, and for each sample, the locality of the respective hive was obtained from the beekeeper. Given a radius of 2 km around each hive (see Section 3.2 below), sites did not overlap except for the three urban samples from Basel (Figure 1).

2.2. DNA extraction

Honey consists of >80% sugar, which can inhibit PCR amplification (Lallmangaihi et al. 2014). To reduce the sugar content, a sequence of washing steps was employed. For each sample, four times 12.5 g of honey was put in small tubes, dissolved in 45 ml double-distilled H₂O, and subsequently centrifuged for 15 min at 6000 rpm. After discarding of the supernatant, the pellets were re-suspended in 5 ml H₂O, and then pooled, and centrifuged for 10 min at 6000 rpm. The supernatant was discarded and the probe was again suspended in 500 µl H₂O. Adding zirconium beads, extracted pollens were mechanically crushed for 1 min on a swing mill at a frequency of 30/s. One thousand microliters of CTAB extraction buffer and 18 µl RNase A were added, and the solution was incubated at 65 °C for 30 min when 30 µl of proteinase K was added, and each sample was further incubated for 60 min at 65 °C. Finally, DNA was extracted using a standard phenol:chloroform extraction protocol (Green and Sambrook 2017). A negative (blank) control and a positive control were included. The positive control consisted of genomic DNA extracted from North American *Arabidopsis lyrata* subsp.

lyrata, being the main plant species studied in the molecular lab used.

After DNA isolation, the second nuclear ribosomal internal transcribed spacer region (ITS2) was amplified via PCR using published primers: forward: 5'-ATGCGATACTTGGTGTGAAT-3', reverse: 5'-GACGCTTCTCCAGACTACAAT-3' (Richardson et al. 2015). The Phusion High-Fidelity PCR Kit (Bioconcept, Allschwil, Switzerland) was used. The PCR conditions were those suggested by the manufacturer's protocol: initial denaturation at 98 °C for 30 s, followed by 30 cycles of denaturation at 98 °C for 10 s each, annealing at 59 °C for 30 s, extension at 72 °C for 30 s, and a final extension at 72 °C for 10 min. PCR amplicons were cleaned with the QIAquick Kit (Qiagen, Hombrechtikon, Switzerland) following the manufacturer's protocol. Using the NEBNext Ultra DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (Bioconcept, Allschwil, Switzerland), 500 ng purified PCR product from each sample was indexed. Indexed samples were then purified using Agencourt AMPure XP beads (Beckman Coulter, Nyon, Switzerland) and pooled at the end. The pooled library was amplified by nine PCR cycles and paired-end sequenced on an Illumina MiSeq using the MiSeq Reagent Kit v3 for 600 cycles that allows for read lengths of 2 × 300 bp. Sequencing was performed at the Genomic Diversity Center (GDC), ETH Zürich. All sequence data was deposited on the NCBI short-read archive—BioProject ID: PRJNA507718.

2.3. Genetic analyses and plant identity

Demultiplexed reads were assembled using PEAR 0.9.8 (Zhang et al. 2014), allowing a minimal PHRED quality score of 24. Paired reads that did not contain both primer sites were removed. Assembled read pairs were clustered with USEARCH 7.0.1090 (Edgar 2010), clustering reads with more than 97% sequence identity. To reduce sequencing errors, clusters with less than 5 reads were discarded. The identity of each sequence was established by comparing them against the NCBI GenBank database using MegaBlast on the 20th of June 2018. For each sequence, the 20 best matches were retained.

Figure 1. Map showing the locations of the 14 beehives studied. The inset shows a map of Switzerland with the region zoomed in being highlighted. Circles depict a radius of 2 km around each hive. Sample names indicate region (BE, Bern; BS, Basel; ZH, Zürich), followed by an index number and habitat indicator (U, urban; NU, non-urban). Lines depict the borders of Swiss cantons.

Using custom R scripts, each sequence was assigned to a species if the sequence matched a single species in the database with 100% identity. If the sequence matched 100% with more than one species in the database, only the genus was retained. In the absence of a complete match, the genus of matches with > 95% identity was retained. Calculations were performed at sciCORE (<http://scicore.unibas.ch>), the scientific computing core facility of the University of Basel.

Because many sequences could not be confidently assigned to a particular species (see Section 3), all subsequent diversity analyses were performed at the genus level. To further account for variation in read numbers among samples (Table 1), the relative proportions of each genus within a sample were used to calculate Shannon diversity indices. Furthermore, we checked whether each genus occurs naturally north of the Alps and whether it is agriculturally cultivated or used as a crop plant, based on the published Swiss

Flora (Lauber et al. 2018). Ambiguous cases, i.e., where a genus included native and nonnative species, or both crop and non-crop species, were scored as “mixed” and excluded from pairwise comparisons. In addition, data on phenology of the most common plant genera (Landolt et al. 2010) was used to verify the likely time span during which each honey sample was produced by the honeybees.

2.4. Land cover

We estimated land use of an area around the hives that corresponds with the common size of foraging ranges of honeybees in springtime. The spatial range of pollen collecting by honeybees is affected by many factors including the time of the year and landscape complexity. Honeybees forage over a wider spatial area in summer than in spring (Couvillon et al. 2015), and in simple compared to complex landscapes (Steffan-

Table 1. Information on honey samples included in this study. Each sample was assigned with an ID comprising information on its region of origin (two letters), an index number, and the type of environment it came from (U, urban; NU, non-urban). For each sample, further information is given: the coordinates of the hive (in degrees), the number of sequence reads before and after filtering, the number of identified plant genera, the Shannon index based on plant genera, and the likely timing of honey collection based on plant flowering phenology. The information on phenology is the range of the earliest month of flowering that has been reported for the major plant genera detected (Figure 3)

Sample ID	Region	Type	North	East	# Raw reads	# Filtered reads	# Genera	Shannon index	Phenology
BE1_U	Bern	Urban	46.961	7.445	655,947	551,111	143	2.481	3–6
BE2_U	Bern	Urban	46.948	7.391	710,342	567,501	100	2.516	3–6
BE3_NU	Bern	Non-urban	47.009	7.400	392,240	321,987	115	2.899	3–5
BE4_NU	Bern	Non-urban	46.974	7.508	420,881	330,407	25	0.750	4
BS1_U	Basel	Urban	47.554	7.574	476,960	396,088	53	1.720	3–4
BS2_U	Basel	Urban	47.545	7.586	385,921	311,435	88	1.802	3–6
BS3_U	Basel	Urban	47.541	7.594	623,390	505,306	46	1.371	4–6
BS4_NU	Basel	Non-urban	47.462	7.579	527,860	424,712	72	1.422	4
BS5_NU	Basel	Non-urban	47.466	7.464	392,956	319,011	61	1.845	3–6
ZH1_U	Zürich	Urban	47.358	8.577	537,393	441,385	145	2.871	3–5
ZH2_U	Zürich	Urban	47.380	8.487	483,302	402,806	45	1.455	3–5
ZH3_NU	Zürich	Non-urban	47.437	8.640	514,437	409,088	34	0.497	5
ZH4_NU	Zürich	Non-urban	47.436	8.543	544,993	437,620	66	1.845	3–6
ZH5_NU	Zürich	Non-urban	47.280	8.631	430,847	341,436	97	1.984	3–5
Negative control	–	–	–	–	24,622	1003	–	–	–
Positive control	–	–	–	–	455,772	378,478	–	–	–

Dewenter and Kuhn 2003). We chose a radius of 2 km around the hives for the estimation of land cover, which is close to the median estimated pollen collecting range of honeybees during spring in forests and agricultural areas (Steffan-Dewenter and Kuhn 2003) as well as in urban areas (Beekman and Ratnieks 2000; Beekman et al. 2004; Couvillon et al. 2015). Land cover data was extracted from the CORINE Land Cover database, an inventory of land cover classified in 44 categories for Switzerland (Steinmeier 2013). The database has a minimal mapping unit of 25 ha for areal resolution and a minimum width of 100 m for linear phenomena.

2.5. Statistical analysis

In a first step, differences in plant composition among honey samples were correlated with differences in land use. For this, Euclidean distance matrices were established using either the proportions of all identified genera or only the proportions of the genera occurring with > 5% frequency in at least one honey sample, and for land cover. Differentiation among samples based on Euclidean distances was estimated with 1000 bootstrap replicates in the R package *PVCLUST* 2.0 (Suzuki and Shimodaira 2006). The resulting pollen diversity and land cover matrices were further

compared using a Mantel test, whose significance was established with 1000 bootstrap replicates.

All honey samples were included in the analysis to increase statistical power. Given the spatial overlap of urban sites in Basel (Figure 1), Euclidean clustering and Mantel tests were repeated using each of the three urban sites separately. This allowed assessing if their combined inclusion would have resulted in an artificial clustering of urban sites.

In a second step, plant diversity was compared between urban and non-urban honey samples. The total number of identified plant genera, the Shannon diversity index, the percentage of nonnative plant genera, and the percentage of crop plant genera were compared. The Shannon index of diversity was calculated for each honey sample on the genus level with the package *vegan* 2.4-5 (Oksanen et al. 2017). Given the available sample size, statistical testing was performed using two-sample Wilcoxon rank sum tests. Because the urban sample BE2_U clustered with non-urban samples based on land cover data (see Section 3), differentiation between urban and non-urban samples was also tested either removing BE2_U or assigning it to the group of non-urban sites.

Lastly, a sample size calculation was performed to determine the number of samples needed for which the difference between urban and non-urban sites would reach statistical significance (i.e., $p < 0.05$). This was done using the R package *samplesize* 0.2-4 (Scherer 2016), assuming the same variance as observed among the actual samples and a power of 0.8. All statistical analyses were performed in R 3.3.1 (R Core Team 2016).

3. RESULTS

3.1. Meta barcoding

A total of 7,097,469 paired-end raw reads were obtained for the 14 honey samples (mean read number $506,962 \pm 102,076$ SD; Table I). Following assembly and filtering, 5,759,893 reads, i.e., 81.2%, were available for analysis (mean read number $411,421 \pm$

84,472 SD; Table I). After assembly and filtering, the negative control did not contain any sequences and 99.997% of the sequences of the positive control were assigned to *Arabidopsis*. With the exception of sample ZH4_NU, contamination with *Arabidopsis* was low (range 0–29%, average 2.9%) and these sequences were removed. Four samples (BE2_U, BE3_NU, BS1_U, ZH1_U) also contained minor traces of moss belonging to the genera of *Ceratodon*, *Funaria*, and *Orthotrichum*. Moreover, three samples (BE2_U, BS5_NU, ZH4_NU) contained traces of algae (*Vaucheria* sp.). Interestingly, all but three samples (BS3_U, BS5_NU, ZH4_NU) contained reads assigned to the nematode genus *Brugia*. Reads of these five genera were also removed. The remaining sequences could be assigned to a total of 262 genera (Table S1). Moreover, 194 species were identified (Table S2), but most reads could only be identified to the genus level.

The number of genera found per honey sample varied, with a range of 25–145 genera (average 78; Table I). The overlap in plant genera between samples differed among regions (Figure 2). Samples from Bern showed a higher overlap than samples from Basel or Zürich. Many genera in the samples from Basel were uniquely found in a single honey sample, even though the circles with a 2-km radius around hives intersected for the urban sites of Basel. Based on Euclidean distances in plant composition, two significantly differentiated clusters occurred, both including samples from urban and non-urban sites. One cluster had a high abundance of *Brassica*, and the other cluster had high proportions of *Trifolium* and *Rubus* (Figure 3). Two samples were not part of these two clusters: sample BS3_U was dominated by *Hydrangea*, and sample BS1_U showed a high abundance of *Myosotis*. Only 15 genera occurred with more than 5% frequency in at least one of the samples (Figure 3). The time during which honey was collected likely differed among samples as indicated by the phenology of the predominantly occurring genera (Table I). For example, whereas *Trifolium* and *Vicia* species flower as early as March, *Phacelia* and *Tilia* start flowering in June (Table I).

Figure 2. Venn diagrams summarizing the number of shared plant genera among different honey samples for each of the three regions. Samples from urban environments are shaded in red and those from non-urban environments in green.

3.2. Land cover

Land cover fell into two statistically supported clusters based on Euclidean distances between sites, reflecting differentiation between urban and non-urban sites (Figure 4). This was also true when only one urban population from Basel was included (Figure S1). Urban sites were predominantly composed of discontinuous and continuous urban fabric and the respective infrastructure, whereas non-urban sites varied in their composition of forest and arable land (Figure 4). The only exception was sample BE2_U; it was from within the city of Bern but assigned to the non-urban cluster. Unlike other urban sites, BE2_U was surrounded by a substantial fraction of forest.

3.3. Diversity estimates

Pairwise distances in plant composition were not significantly correlated with pairwise distances in land cover using either all genera (Mantel $r = 0.139$, $p = 0.960$) or only the most commonly occurring genera (Mantel $r = 0.140$, $p = 0.952$). This was similarly true when only one urban population from Basel was included (Table S3). Neither the number of genera found in each sample ($W = 31$, $p = 0.456$) nor the Shannon diversity index ($W = 28$, $p = 0.620$) differed significantly between urban and non-urban samples (Table 1; Figure 5). Moreover, there was no statistical difference in the percentages of nonnative plant genera ($W = 12$, $p = 0.128$) or crop

plant genera ($W = 17.5$, $p = 0.406$) between urban and non-urban samples. None of the comparisons were significant when BE2_U was either excluded or treated as a non-urban sample (Table S4).

Sample size calculations suggested that, given the data, at least 101 honey samples ($N_{\text{Urban}} = 58$, $N_{\text{NonUrban}} = 43$) would have been necessary to detect a statistical difference between urban and non-urban sites in the number of genera and 88 honey samples ($N_{\text{Urban}} = 37$, $N_{\text{NonUrban}} = 51$) to detect a difference in the Shannon diversity index. At least 38 ($N_{\text{Urban}} = 19$, $N_{\text{NonUrban}} = 19$) and 212 ($N_{\text{Urban}} = 86$, $N_{\text{NonUrban}} = 126$) honey samples would have been necessary to detect a difference in the percentage of nonnative plant genera and the percentage of crop plant genera respectively between urban and non-urban sites.

4. DISCUSSION

Using metabarcoding of pollen in honey, we aimed for characterizing differences in plant-honeybee interactions and visited plant diversity between urban and non-urban sites. With a total of 262 identified plant genera detected, we successfully uncovered the broad diversity of plants visited by honeybees in both types of environment (Figures 2 and 3). However, while land cover around the hives where the honey was collected did significantly differ between urban and non-urban sites (Figure 4), the plant species composition represented by pollen in honey did not (Figure 3). In line, neither the representation of

Figure 3. Summary of the plant genera identified in each honey sample. The Euclidean distance–based dendrogram (top) summarizes the differentiation among samples (red: urban, green: non-urban). Numbers indicate nodes with > 95% bootstrap support. Bar plots (bottom) summarize the proportion of the 15 most commonly found genera (the proportions of all other genera are pooled and shown in black).

Figure 4. Summary of land cover types surrounding each hive within a 2-km radius. The Euclidean distance-based dendrogram (top) summarizes the differentiation among samples (red: urban, green: non-urban). Numbers indicate nodes with >95% bootstrap support. Bar plots (bottom) show the proportion of land cover types.

Figure 5. Boxplots summarizing variation in plant genera detected by metabarcoding of pollen in honey from non-urban and urban sites: the number of plant genera, the Shannon diversity index, the percentage of nonnative genera, and the percentage of crop plant genera respectively. P values are based on two-sample Wilcoxon rank sum tests.

nonnative plants nor that of crop plants did significantly differ between urban and non-urban honey. Below, we discuss both the potential and the challenges of metabarcoding honey in the light of our results.

Here, we worked with the ribosomal ITS2 gene that had been used for the metabarcoding of pollen before (Keller et al. 2014; Richardson et al. 2015; Sickel et al. 2015; Bell et al. 2017). With 262 genera detected, our taxonomic resolution was higher than the one of traditional methods

(Bell et al. 2016a; Sniderman et al. 2018), i.e., assigning pollen to the genus and often even to the species level (Tables S1 and S2). However, many sequences could not be unambiguously assigned to a particular species. This lack of resolution down to the species level may have three reasons. A first is the lack of available reference DNA barcodes with which sequences can be compared (Deiner et al. 2017). This is the case for many plants occurring in Switzerland (Wyler and Naciri 2016). A second reason is that DNA barcodes can

have a limited phylogenetic resolution, particularly in cases of recently evolved species such as those of the *Brassica* genus (Wyler and Naciri 2016). Lastly, the taxonomic resolution may also be reduced by the potential presence of multiple copies of the ITS region that have not undergone concerted evolution (Feliner and Rosselló 2007). To increase taxonomic resolution, future studies may combine multiple barcode markers or use longer sequences (Parks et al. 2009; Wyler and Naciri 2016).

Despite the broad range of plant taxa visited by honeybees, only 15 genera were predominantly visited among all samples (Figure 3). The overlap between honey samples for less frequently visited plants was modest, with many honeybee colonies visiting unique plant genera, even in cases where the circles with a 2-km radius around hives overlapped (Figures 1 and 2; Table S1). These findings are in concordance with a recent metabarcoding study, which showed that different honeybee colonies within a botanical garden visited only a small fraction of flowering plant species that existed in the garden, with little overlap among hives (de Vere et al. 2017). Such a biased visitation of a few plant genera may be because they provide a relatively high amount of nectar or pollen compared with the other plants, as for example *Trifolium*, or because they exhibit particular flowering signals such as ornamental shrubs of the genus *Hydrangea* (Ohashi et al. 2015), both of which were well represented in several of our honeys. Biased plant visitation could also be a result of flower constancy, i.e., when a pollinator exclusively visits a particular flowering plant species during one collection trip (Waser 1986), which is a characteristic behavior of honeybees (Free 1963; Hill et al. 1997). Also, few scouting honeybees may recruit a majority of a colony to forage on few profitable plants through in-hive communication (Beekman and Ratnieks 2000). One or multiple of these factors may account why some genera were found only in one or a few samples (Figures 2 and 3).

Land cover and thus the habitat that was available for honeybees significantly differed between urban and non-urban sites (Figure 4). Urban sites were characterized by a high abundance of discontinuous and continuous structures, whereas

non-urban sites were dominated by forest and arable land (Figure 4). Despite the difference in land cover around the hives, we did not detect a statistical difference in terms of plant composition or diversity represented in honey between the two environments (Figure 5), unlike to what earlier studies had found (Kühn et al. 2004; Baldock et al. 2015). This could reflect our limited detection power given our sample size. Indeed, our power analyses suggested that many more samples would be needed for the observed differences (Figure 5) to become statistically significant (see Section 3). Apart from selective plant visitation, seasonality may further account for the absence of a difference in plant-pollinator interactions between the two ecologically distinct environments. Seasonality was shown to be very important for honeybee foraging, more than landscape diversity (Danner et al. 2017). While we used commercially available honey that was sold as spring honey, it may have been harvested at different times. Data on first flowering for the major genera (species-averages from across Switzerland) found in each sample showed that some honey was probably collected over a time period from April to May, and some cases as late as June (Table I), which may have blurred patterns between urban and non-urban sites.

Taken together, our study shows that metabarcoding of honey provides a high taxonomic resolution of plant visitations by honeybees. We further showed that metabarcoding is a valuable technology that can help address open questions in urban ecology (Bell et al. 2016b). Despite our predictions based on the literature (Kühn et al. 2004; Baldock et al. 2015; Harrison and Winfree 2015), we did not detect a statistical difference in the composition of flowering plants between two ecological highly distinct environments (Figure 5). Further studies are thus needed to assess if and to which degree the pollen composition found in honey may differ between urban and non-urban sites. Such studies would ideally use honey collected over the same time span and use much larger sample sizes than ours, preferably combined with a good barcode reference database of the local flora.

ACKNOWLEDGMENTS

We thank Harald Burger, Anna Hochreutener, Peter Linder, René Meier, Guy de Morsier, Markus Muntwiler, Manuela Plattner, Andreas Seiler, Tom Scheuer, and Christophe Wicht for providing honey samples and Silvia Kobel from the Genetic Diversity Centre (GDC), ETH Zürich for her assistance with preparing the libraries. Christophe Praz and one anonymous reviewer provided helpful comments on the manuscript.

AUTHOR CONTRIBUTIONS

KL and YW designed the study; KL and NH established the genomic pipeline; TP provided the land cover data; KL, AG, and SG performed the laboratory work and analyses; AM categorized plant species; KL wrote the manuscript with the input of all co-authors. All authors have read and approved the final manuscript.

COMPLIANCE WITH ETHICAL STANDARDS

Conflict of interest The authors declare that they have no conflict of interest.

Le métabarcoding du miel pour évaluer les différences d'interactions plantes-pollinisateurs entre sites urbains et non urbains

Apis mellifera / écologie urbaine / Index de Shannon / Métabarcoding

Metabarcoding von Honig erlaubt die Beurteilung von Unterschieden in der Wechselbeziehung zwischen Pflanzen und Bestäubern an urbanen und nicht-urbanen Standorten

Apis mellifera / Stadtökologie / Shannon Index / Metabarcoding

REFERENCES

- Baldock, K.C.R., Goddard, M.A., Hicks, D.M., Kunin, W.E., Mitschunas, N., *et al.* (2015). Where is the UK's pollinator biodiversity? The importance of urban areas for flower-visiting insects. *P. R. Soc. B.* **282**: 20142849.

- Beekman, M. & Ratnieks, F.L.W. (2000). Long-range foraging by the honey-bee, *Apis mellifera* L. *Funct. Ecol.* **14**: 490–496.
- Beekman, M., Sumpter, D.J.T., Seraphides, N. & Ratnieks, F.L.W. (2004) Comparing foraging behaviour of small and large honey-bee colonies by decoding waggle dances made by foragers. *Funct. Ecol.* **18**: 829–835.
- Bell, K.L., Burgess, K.S., Okamoto, K.C., Aranda, R. & Brosi, B.J. (2016a) Review and future prospects for DNA barcoding methods in forensic palynology. *Forensic. Sci. Int. Genet.* **21**: 110–116.
- Bell, K.L., de Vere, N., Keller, A., Richardson, R.T., Gous, A., *et al.* (2016b) Pollen DNA barcoding: current applications and future prospects. *Genome* **59**: 629–640.
- Bell, K.L., Fowler, J., Burgess, K.S., Dobbs, E.K., Gruenewald, D., *et al.* (2017) Applying pollen DNA metabarcoding to the study of plant-pollinator interactions. *Appl. Plant. Sci.* **5**: 1600124.
- Cheptou, P.-O. & Avendaño, L.G.V. (2006) Pollination processes and the Allee effect in highly fragmented populations: consequences for the mating system in urban environments. *New Phytol.* **172**: 774–783.
- Concepción, E.D., Moretti, M., Altermatt, F., Nobis, M.P. & Obrist, M.K. (2015) Impacts of urbanisation on biodiversity: the role of species mobility, degree of specialisation and spatial scale. *Oikos* **124**: 1571–1582.
- Couvillon, M.J., Riddell Pearce, F.C., Accleton C., Fensome, K.A., Quah, S.K.L., *et al.* (2015) Honey bee foraging distance depends on month and forage type. *Apidologie* **46**: 61–70.
- Danner, N., Keller, A., Härtel, S. & Steffan-Dewenter, I. (2017) Honey bee foraging ecology: Season but not landscape diversity shapes the amount and diversity of collected pollen. *PLoS One* **12**: e0183716.
- de Vere, N., Jones, L.E., Gilmore, T., Moscrop, J., Lowe, A., *et al.* (2017) Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Sci. Rep.* **7**: 42838.
- Deiner, K., Bik, H.M., Mächler, E., Seymour, M., Lacoursière-Roussel, A. *et al.* (2017) Environmental DNA metabarcoding: transforming how we survey animal and plant communities. *Mol. Ecol.* **26**: 5872–5895.
- Dolan, R. W., Moore, M. E. & Stephens J. D. (2011) Documenting effects of urbanization of flora using herbarium records. *J. Ecol.* **99**: 1055–1062.
- Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.
- Feliner, G.N. & Rosselló, J.A. (2007) Better the devil you know? Guidelines for insightful utilization of nrDNA ITS in species-level evolutionary studies in plants. *Mol. Phylogenet. Evol.* **44**: 911–919.
- Free, J.B. (1963) The flower constancy of honeybees. *J. Anim. Ecol.* **32**: 119–131.
- Geldmann, J. & González-Varo, J.P. (2018) Conserving honey bees does not help wildlife. *Science* **359**: 392–393.

- Geslin, B., Gauzens, B., Thébault, E. & Dajoz, I. (2013) Plant pollinator networks along a gradient of urbanisation. *PLoS One* **8**: e63421.
- Green, M.R. & Sambrook, J. (2017) Isolation of high-molecular-weight DNA using organic solvents. *Cold Spring Harb Protoc* **2017**: pdb.prot093450.
- Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J. *et al.* (2008) Global change and the ecology of cities. *Science* **319**: 756–760.
- Harrison, T. & Winfree, R. (2015) Urban drivers of plant-pollinator interactions. *Funct. Ecol.* **29**: 879–888.
- Hawkins, J., de Vere, N., Griffith, A., Ford, C.R., Allainguillaume, J. *et al.* (2015) Using DNA metabarcoding to identify the floral composition of honey: A new tool for investigating honey bee foraging preferences. *PLoS One* **10**: e0134735.
- Hill, P., Wells, P. & Wells, H. (1997) Spontaneous flower constancy and learning in honey bees as a function of colour. *Anim. Behav.* **54**: 615–627.
- Keller, A., Danner, N., Grimmer, G., Ankenbrand, M., von der Ohe, K. *et al.* (2014) Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol.* **17**: 558–66.
- Kühn, I., Brandl, R. & Klotz, S. (2004) The flora of German cities is naturally species rich. *Evol. Ecol. Res.* **6**: 749–764.
- Lalmangaihi, R., Ghatak, S., Laha, R., Gurusubramanian, G. & Kumar, N.S. (2014) Protocol for optimal quality and quantity pollen DNA isolation from honey samples. *J. Biomol. Tech.* **25**: 92–95.
- Landolt, E., Bäumler, B., Erhardt, A., Hegg, O., Klötzli, F. *et al.* (2010) *Flora indicativa*, 1st edition. Haupt Verlag, Bern.
- Lauber, K., Wagner, G., Gygax, A. (2018). *Flora Helvetica - Illustrierte Flora der Schweiz*, 6th edition. Haupt Verlag, Bern.
- Lososová, Z., Chytrý, M., Tichý, L., Danihelka, J., Fajmon, K. *et al.* (2011) Native and alien floras in urban habitats. *Glob. Ecol. Biogeogr.* **21**: 545–555.
- Ohashi, K., Makino, T.T. & Arikawa, K. (2015) Floral colour change in the eyes of pollinators: testing possible constraints and correlated evolution. *Funct. Ecol.* **29**: 1144–1155.
- Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., *et al.* (2017) *vegan*: Community ecology package. R package version 2.4-5. <https://CRAN.R-project.org/package=vegan>
- Parks, M., Cronn, R. & Liston, A. (2009) Increasing phylogenetic resolution at low taxonomic levels using massively parallel sequencing of chloroplast genomes. *BMC Biol.* **7**: 84.
- Pickett, S.T.A., Cadenasso, M.L., Grove, J.M., Boone, C.G., Groffman, P.M. *et al.* (2011) *J. Environ. Manag.* **92**: 331–362.
- R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.
- Richardson, R.T., Lin, C.-H., Sponsler, D.B., Quijia, J.O., Goodell, K. & Johnson, R.M. (2015) Application of ITS2 metabarcoding to determine the provenance of pollen collected by honey bees in an agroecosystem. *Appl. Plant Sci.* **3**: 1400066.
- Scherer, R. (2016) *sampleSize 0.2-4*: Sample size calculation for various t-tests and Wilcoxon-test. <https://CRAN.R-project.org/package=sampleSize>.
- Sickel, W., Ankenbrand, M.J., Grimmer, G., Holzschuh, A., Härtel, S. *et al.* (2015) Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecol.* **15**: 20.
- Smart, M.D., Cornman, R.S., Iwanowicz, D.D., McDermott-Kubeczko, M., Pettis, J.S. *et al.* (2017) A comparison of honey bee-collected pollen from working agricultural lands using light microscopy and ITS metabarcoding. *Environ. Entomol.* **46**: 38–49.
- Sniderman, J.M.K., Matley, K.A., Haberle, S.G. & Cantrill, D.J. (2018) Pollen analysis of Australian honey. *PLoS One* **13**: e0197545.
- Steffan-Dewenter, I. & Kuhn, A. (2003) Honeybee foraging in differentially structured landscapes. *Proc. R. Soc. Lond. B* **270**: 569–575.
- Steinmeier, C. (2013) CORINE Land Cover 2000/2006 Switzerland. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf.
- Suzuki, R. & Shimodaira, H. (2006) Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* **22**: 1540–1542.
- Thorp, R.W. (2000) The collection of pollen by bees. *Plant Syst. Evol.* **222**: 211–223.
- Waser, N.M. (1986) Flower constancy: Definition, cause, and measurement. *Am. Nat.* **127**: 593–603.
- Wilson, E.E., Sidhu, C.S., LeVan, K.E. & Holway, D.A. (2010) Pollen foraging behaviour of solitary Hawaiian bees revealed through molecular pollen analysis. *Mol. Ecol.* **19**: 4823–4829.
- Wyler, S.C. & Naciri, Y. (2016) Evolutionary histories determine DNA barcoding success in vascular plants: seven case studies using intraspecific broad sampling of closely related species. *BMC Evol. Biol.* **16**: 103.
- Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. (2014) PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**: 614–620.

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