Honey bee pollen foraging ecology across an urbanization gradient

Rodney T. Richardson^{1*}, Tyler D. Eaton², Chia-Hua Lin², Garrett Cherry², Reed M. Johnson²
and Douglas B. Sponsler³

¹ Department of Biology, York University

² Department of Entomology, Ohio State University

³ Department of Entomology, Penn State University

* Author for correspondence:

Name: Rodney Trey Richardson

Email: rtr87@yorku.ca

Abstract

Understanding animal foraging ecology requires large samples sizes spanning broad environmental and temporal gradients. For pollinators, this has been hampered by the laborious nature of morphologically identifying pollen. Metagenetic pollen analysis is a solution to this issue, but the field has struggled with poor quantitative performance. Building upon prior laboratory and bioinformatic methods, we applied quantitative multi-locus metabarcoding to characterize the foraging ecology of honey bee colonies situated along an urban-agricultural gradient in central Ohio, USA. In cross-validating a subset of our metabarcoding results using microscopic palynology, we find strong concordance between the molecular and microscopic methods. Our results show that, relative to the agricultural environment, urban and suburban environments were associated with higher taxonomic diversity and temporal turnover of honey bee pollen forage. This is likely reflective of the fine-grain heterogeneity and high beta diversity of urban floral landscapes at the scale of honey bee foraging. Our work also demonstrates the power of honey bees as environmental samplers of floral community composition at large spatial scales, aiding in the distinction of taxa characteristically associated with urban or agricultural land use from those distributed ubiquitously across our landscape gradient.

Keywords: Quantitative metabarcoding; urban landscape ecology; urban biotic homogenization; pollinator nutrition; community ecology; quantitative ecology; molecular ecology

Introduction

As a perennial, eusocial and cosmopolitan florivore, the western honey bee (*Apis mellifera* L.) faces unique demands for resource acquisition. These demands are met by a sophisticated colony-level foraging biology characterized by extreme dietary generalism, energetic optimization, and the integration of private and shared information about landscape-scale floral resources (Seeley 1995, Grüter and Farina 2009). These traits make honey bee foraging responsive to daily and seasonal turnover of floral resources within a given landscape (Visscher and Seeley 1982, Couvillon et al. 2014b) and adaptable to the floristic variation across landscapes ranging from deserts to rainforests in every continent but Antarctica. Nevertheless, honey bee fitness responds to landscape composition (Lecocq et al. 2015, Sponsler and Johnson 2015, Smart et al. 2016, Döke et al. 2018), indicating that behavioral sophistication cannot fully compensate for variation in landscape-scale floral resource availability. This inference is corroborated by empirical and theoretical studies demonstrating that the adaptive value of the waggle dance—the pinnacle of honey bee foraging sophistication—varies with the spatiotemporal distribution of floral resources (Donaldson-Matasci and Dornhaus 2012, Schürch and Grüter 2014).

During the Anthropocene, honey bee biogeography has come to include the novel landscapes created by human agriculture and urbanization, both of which generate floristic conditions that are far-removed from their nearest natural analogs. In typical agricultural landscapes, large crop monocultures are the predominant feature, intercalated with marginal non-crop habitat features like field margins, roadsides, and riparian buffers. In contrast, urban landscapes are characterized by an extensive impervious matrix finely intermixed with relatively small patches of greenspace composed of a mixture of native and exotic species, in some cases

resulting in a net species richness equal to or greater than that of nearby non-urban areas, though typically with lower species density (McKinney 2008, Aronson et al. 2014, Aronson et al. 2015). Ultimately, agricultural and urban landscapes diverge substantially both from each other and from the historic landscapes they displace.

The question of how honey bee foraging behavior responds to the unique challenges of urban and agricultural landscapes is salient from both basic and applied perspectives (Couvillon et al. 2014a, Härtel and Steffan-Dewenter 2014). With respect to urban land use, recent empirical studies have described honey bee (Lau et al. 2019, Lucek et al. 2019) diet and foraging distance (Garbuzov et al. 2015). The only study, however, to articulate a theoretical relationship between honey bee foraging behavior and suburban land use is that of Waddington et al. (1994), where the authors suggest that the fine-grain pattern of floral patch distribution typical of suburban landscapes favors independent patch discovery over the use of waggle dance recruitment by foraging honey bees. Under this hypothesis, researchers would expect a greater number of floral patches to be simultaneously exploited in suburban areas. With respect to agricultural landscapes, Steffan-Dewenter and Kuhn (2003) suggest that highly simplified agricultural landscapes with large monoculture patches result in fewer patches being simultaneously exploited, and Nürnberger et al. (2019) conclude that patch discovery would be more dependent on waggle dance recruitment in this circumstance. Together, these studies formulate essentially reciprocal hypotheses regarding the effects of landscape on optimal foraging strategy. Thus, a preliminary synthesis of existing literature would be the hypothesis that urban and agricultural land use have divergent effects on honey bee spatial foraging patterns, the former generating more diffuse foraging and the latter generating a more clumped recruitment-driven foraging. A logical extension of this spatial hypothesis would be the prediction that the taxonomic

composition of honey bee foraging would follow similar patterns, being more diversified in urban landscapes and more homogenized in agricultural landscapes.

Areas where agricultural and urban landscapes occur in close proximity present ideal opportunities to study the behavioral response of honey bee foraging to the respective challenges presented by these landscapes while controlling for regional climate and floristics (Sponsler et al. 2017). In the present study, we employ DNA metabarcoding to identify the botanical origins of honey bee-collected pollen sampled along an urban-agricultural gradient in central Ohio, USA. By characterizing the taxonomic composition, diversity, and turnover of these pollen samples, we revisit the hypotheses of Waddington et al. (1994) and Steffan-Dewenter and Kuhn (2003), testing the prediction that, along a gradient spanning from fine-grain urban land use to highly simplified agricultural land use, pollen diversity and sample-to-sample turnover will correlate positively with increasing urbanization.

Methods

Site selection and sample collection

In the spring of 2013, we established apiaries at four sites spanning an urban-agricultural gradient (Figure 1). Two of these sites (BK and GH) were within the metropolitan area of Columbus Ohio, USA, with a third apiary (CC) residing on the edge of a western suburb and the fourth apiary (FSR) residing approximately 23 km from the western edge of the city. Each apiary contained two to six eight-frame Langstroth hives, two of which were fitted with Sundance I bottom-mounted pollen traps (Ross Rounds, Albany, NY, USA). From late spring to late fall of both 2013 and 2014, we continuously trapped corbicular pollen from one hive per apiary on a

weekly basis, collecting 109 bulk pollen samples in total. To avoid causing pollen starvation, we alternated trapping between colonies from one sampling event to the next.

GIS landscape analysis

We analyzed landscape composition using the 2011 National Land Cover Database (NLCD) land cover layer (Homer et al. 2015). For each site, we extracted from the NLCD layer a 3 km radius and tabulated its composition using QGIS (Team 2015) and R (Team 2014). As an approximate measure of the degree of urbanization for each site, we summed all four categories of developed land classifications of the NLCD layer into a single ‘urban’ category. This measure of urbanization ranged from 7 to 98 percent of the total land area within the landscapes surrounding our sites.

Microscopic palynology

For a subset of our samples, 9 samples collected at site CC between July 31st and September 26th of 2014, microscopic palynology was used to characterize pollen taxonomic composition. These microscopy results were published in a previous study (Sponsler et al. 2017), otherwise independent from the present work. Briefly, bulk pollen samples were subsampled to 10 percent of their total weight up to a maximum of 10 g for samples with mass of greater than 100 g. Subsamples were then sorted by color and texture, mounted with fuchsin-stained glycerin jelly, and identified by comparison to a local reference collection. Proportional abundance of taxa within each sample were calculated using the volumetric approach of O’Rourke and Buchmann (1991).

Molecular pollen analysis laboratory procedures

Broadly, our molecular pollen analysis methods follow those of Richardson et al. (2019). Prior to DNA extraction, corbicular pollen pellets were subsampled and homogenized using

previously described methods. For each sample, the lesser amount of either 20 g or 10 percent by mass of each bulk sample was added to distilled water to yield a final concentration of 0.1 g/mL of pollen. The pollen mixture was placed in a blender (Hamilton Beach #54225, Southern Pines, NC, USA) and homogenized for 2.5 min. Pollen homogenate, 1.4 mL, was then placed in a 2 mL bead beater microfuge tube (Fisherbrand Free- $\square$ Standing Microcentrifuge Tubes; Fisher Scientific, Hampton, NH, USA). To facilitate mechanical pollen disruption during bead beating, 3,355 mg of 0.7 mm zirconia beads (Fisher Scientific, Hampton, NH, USA) were added to each tube. Tubes were placed in a bead-beater (Mini- $\square$ BeadBeater- $\square$ 16; BioSpec Products, Bartlesville, OK, USA) for 5 minutes at the highest setting. Samples were briefly agitated with a vortex mixer prior to removing 400 μ l of homogenate for DNA extraction using the DNeasy Plant Minikit (QIAGEN Inc., Valencia, CA) according to the manufacturer's protocol.

Three loci were selected for amplicon library preparation: *rbcL*, *trnL* and ITS2. For compatibility with the Illumina Nextera sequencing protocol, we conducted a 3-step PCR protocol in which the product of each step was used as the template in the next step. Since this protocol was necessary for each locus, each sample required a total of nine PCR reactions. The following templates were used sequentially in the three PCR steps: 1) 100–150 ng of pollen DNA extract; 2) 1 μ l of unpurified PCR product from the first PCR reaction; 3) 1 μ l of unpurified PCR product from the second PCR reaction. Universal primers (White 1990, Taberlet et al. 1991, Kress and Erickson 2007, Palmieri et al. 2009, Chen et al. 2010) with no 5-prime fusion oligos were used in the first PCR reaction. The primers used in the second and third PCR reactions were modified to allow template priming, sample indexing, and lane hybridization during Illumina sequencing. Each of the three PCR steps were performed in a standard 20 μ l reaction containing 10 μ l of Phusion Flash High-Fidelity PCR Master Mix (Thermo Fisher

Scientific, Inc), template DNA, and a pair of DNA primers. After the third PCR, amplicon products were purified and normalized using the SequelPrep Normalization Plate Kit (Thermo Fisher Scientific, Waltham, MA, USA), pooled equimolarly and sequenced using Illumina MiSeq (2 × 300 cycles). Further details on primer sequences, PCR conditions, and sample dual-indexing design are provided at <https://github.com/RTRichar/UrbanRuralPollen>.

Taxonomic annotation of sequence data

During bioinformatic processing, mate-paired sequences were first merged using Pear (v.0.9.1; (Zhang et al. 2013)). While *rbcL* and ITS2 were short enough for mate-pair merging, the *trnL* amplicon we used often exceeded 600 bp in length and few sequences could be merged for this locus. Thus, we used the Metaxa2 Database Builder (v1.0 beta 4; (Bengtsson-Palme et al. 2018)) along with the Metaxa2 classifier (v2.2; (Bengtsson-Palme et al. 2015)) to identify *trnL* sequences, reverse-complement the reverse reads and concatenate the mate-pairs into contiguous sequences. Semi-global top-hit alignment was then performed with VSEARCH (v2.8.1; (Rognes et al. 2016)) to assign taxonomy to all merged or concatenated sequences using custom reference sequence databases, geographically filtered to only include sequences belonging to genera known to be present in Ohio according to the USDA Plants Database. These databases were curated using MetaCurator (v1.0.1; (Richardson et al. 2019)), which utilizes HMMER (v3.1; (Eddy 2011)), MAFFT (v7.270 (Kato et al. 2002)) and VSEARCH to extract reference sequences of the amplicon barcode of interest from all available sequences, including whole genome assemblies.

Statistical analysis of sequencing results

Upon taxonomically annotating our sequence data, we employed the consensus-filtered, median-based analysis implemented in Richardson et al. (2019). Briefly, for each sample, we

excluded genera identified with only one of the three barcode loci. We then calculated the median proportional abundance of these consensus-filtered genera while excluding zeros as non-detections. We then regressed the results of each individual locus, as well as the metabarcoding median against the microscopic data using least squares regression to characterize their quantitative performance. For these regressions, genus-level proportions were used for all genera except those from the family Asteraceae, which were summed to the family level due to the difficulty of morphologically distinguishing these genera.

In addition to testing the molecular data against the microscopic data for a subset of our samples, we investigated the relationship between ITS2-derived taxonomic proportions and metabarcoding median proportions at the genus level for all samples. Given previous studies (Keller et al. 2015, Richardson et al. 2015, Smart et al. 2017, Bell et al. 2019), it was surprising that ITS2 showed the best agreement with the microscopic validation data, and we sought to test if the experiment-wide ITS2 proportions were in strong disagreement with the three-marker median estimates.

We chose to conduct ecological inferences using 3,000 randomly sampled ITS2 sequences per sample as opposed to using the taxonomic proportions generated using the consensus-filtered metabarcoding median results for the following reasons. First, the ecological indices estimated were dependent upon normalized sampling effort, an assumption which was not feasible across all three markers since several *trnL* libraries failed to yield sufficient coverage for analysis, despite the fact that we analyzed all libraries with gel electrophoresis to ensure the presence of the appropriately sized product prior to sequencing. Such library dropout is not uncommon among plant metabarcoding literature (Sickel et al. 2015), and likely points to amplicon clustering issues or an insufficient number of PCR amplification cycles for this

particular marker. Second, genus-level ITS2 taxonomic proportions were strongly correlated with both the microscopic palynology results as well as the three-marker median proportions, indicating that they provided sufficient characterization of our samples both quantitatively and qualitatively.

Using the standardized ITS2 dataset, we estimated genus-level Shannon diversity using the Vegan R package (Dixon 2003). We also estimated two indices of temporal turnover in forage at the genus level, mean rank shift (MRS), a non-parametric index of the degree of quantitative change from one sampling to the next, and qualitative turnover, the sum of taxon appearances and disappearances between sampling events. Qualitative turnover was calculated with a custom R script, while MRS was estimated using the Codyn R package (Hallett et al. 2016). The commands, R and python code, and reference sequence databases used for bioinformatic analysis can be found at <https://github.com/RTRichar/UrbanRuralPollen>.

With ITS2-based estimates of sample diversity, MRS and qualitative turnover, we calculated the median of each ecological index for each site and year. We then regressed these annual index medians against the percent developed land within the surrounding landscape using ordinary least squares regression. For these regressions, we calculated one-tailed *p*-values as we expected higher diversity and temporal turnover at the urban extreme of the gradient at the outset of our experiment. Further, we conducted non-metric multidimensional scaling (NMDS) of these standardized ITS2 datasets to visualize differences in forage composition by site and year. For this analysis, the ggrepel R package (Kamil Slowikowski 2019) was used to dodge text elements.

Results

Sequencing run statistics

After sequencing, we obtained 4,024,124 mate-paired reads, of which 3,094,447 were clearly identifiable as belonging to Viridiplantae after VSEARCH alignment. On average, individual amplicon libraries were sequenced to a depth of $9,463 \pm 286$ SEM, although there was significant variation in the number of sequences per library across the three markers ($P < 0.001$, one-way ANOVA), with the mean and standard error of sequences obtained per library for *rbcL*, *trnL* and ITS2 being $9,862 \pm 422$, $7,439 \pm 626$ and $11,088 \pm 340$, respectively. A follow-up Tukey's HSD test revealed a significant decrease in sequencing depth for *trnL* libraries in particular ($P < 0.005$ for both the *trnL-rbcL* and *trnL-ITS2* comparisons).

Quantitative evaluation of pollen metabarcoding data

Regressing the cube-root squared metabarcoding proportions against the cube-root transformed microscopic palynology proportions (Figure 2), we found strong and significant agreement between the two methods for ITS2 ($P < 0.001$, $R^2 = 0.72$) and the three-marker metabarcoding median ($P < 0.001$, $R^2 = 0.61$). Further, significant relationships were found for *trnL* and *rbcL*, though these relationships were considerably weaker ($P < 0.001$ for both markers, $R^2 = 0.41$ for *rbcL*, $R^2 = 0.21$ for *trnL*). Notably, while the untransformed molecular and microscopic proportions were also highly correlated ($P < 0.001$ for all four tests), the differential transformations we report above resulted in considerably improved model fit, with 4 to 16 percent more variance explained across the four model comparisons, indicating that the molecular results exhibited a slight non-linear bias. Lastly, analysis of the relationship between ITS2 and the three-marker median proportions revealed a strong correlation ($P < 0.001$, $R^2 = 0.71$).

Pollen forage composition across the urban-agricultural gradient

Our results indicate that plant species within Fabaceae and Asteraceae, predominantly *Trifolium*, *Chamaecrista*, *Solidago*, *Symphotrichum* and *Eurybia*, provision the majority of the summer and fall honey bee pollen diet at our study sites (Figure 3 and Figure 4). On average, these two families accounted for a mean and standard error of 73 ± 2 percent across all samples analyzed, with the genus *Trifolium* comprising an average and standard error of 36 ± 2 percent of each sample.

Across the landscape gradient, we detected honey bee foraging on taxa characteristically associated with either urban or agricultural landscapes (Figure 4). Apiaries with access to rural agricultural land cover collected small proportions of *Zea mays* pollen and large proportions of *Glycine*, *Brassica* and *Chamaecrista*. Urban and suburban apiaries appeared to be associated with common ornamental and urban-naturalized plants such as *Ulmus*, *Hedera*, *Hydrangea* and *Buddleja*. We also observed pollen foraging on disturbance-associated taxa, such as *Aralia*, *Rhus*, *Arctium* and *Phytolacca*, toward the urbanized end of the gradient.

Ecological inference across the urban-agricultural gradient

Analyzing median estimates of diversity, mean rank shift (MRS), and qualitative turnover per site and per year, revealed interesting trends (Figure 5). For 2013, diversity, MRS and qualitative turnover were apparently higher in the urban relative to the agricultural site, with the highest median of each index observed in the intermediate half-suburban and half-agricultural site, CC. However, none of these ecological indices were significantly linearly related to the proportion of developed land within the landscape surrounding each apiary ($P > 0.1$ for all tests). In 2014, we observed the same trend, wherein higher ecological index median values were observed toward the urban end of the gradient. For 2014, each of the three indices were

significantly linearly related to the proportion of developed land surrounding each apiary ($P = 0.027$ for diversity, $P = 0.004$ for MRS and $P = 0.005$ for qualitative turnover). Additionally, the two-year average of median diversity per site was significantly related to the proportion of developed land surrounding each apiary ($P = 0.003$). The two-year averages of median MRS and qualitative turnover were not significantly linearly related to the proportion of developed land surrounding each apiary ($P = 0.097$ and $P = 0.123$ for MRS and qualitative turnover, respectively).

NMDS analysis revealed similarity in forage composition within sites across years (Figure 6). Moreover, the arrangement of sites along NMDS axis one reflected their positions on the urbanization gradient, with highly urban sites GH and BK roughly coinciding, followed by the mixed urban and agricultural site CC, and finally the strictly agricultural site FSR.

Discussion

The most salient pattern in our data was that, consistent with our prediction, pollen diversity was greatest at the urban and suburban sites in our study system. It is important, however, not to jump to the interpretation that our more urban sites were more suitable for honey bees than the agricultural sites. To the contrary, prior studies in this system found that urban land use was negatively associated with colony productivity (Sponsler and Johnson 2015) and that honey bees situated at the agricultural-suburban divide of site CC dedicated significantly more foraging activity toward rural agricultural landcover components (Sponsler et al. 2017). While Ohio urban landcover apparently contains a diverse floral assemblage, a trait which likely makes it amenable to generalist solitary bees and bumble bees (Samuelson et al. 2018, Baldock et al. 2019), it may be depauperate in the high-volume floral resource patches which drive honey bee colony growth. Thus, future research contributions which seek to explicitly characterize the

nutritional currency value of honey bee forage diversity may benefit from considering the implications of urban landscapes through the lens of optimal foraging theory and the life history of the honey bee. Investigations of whether elevated urban forage diversity is associated with decreased resource volume and patchiness is particularly warranted.

To the degree that honey bee foraging preferences are highly generalized, colonies containing thousands of foraging individuals enable efficient sampling of the surrounding floral community. The composition of the pollen samples collected from this study highlight the unique floral communities present in different landcover types. Particularly with respect to urban and suburban landscapes, small and heterogeneously managed parcels of land appear to result in greater overall plant community diversity relative to rural agricultural landscapes, at least for the spatial scales corresponding to honey bee foraging. This observation is consistent with the insight of Aronson et al. (2015), which suggests that the effect of urbanization on biological diversity depends strongly on the scale of investigation. At the scale of individual patches, urban land use may be associated with lower diversity due to the dominance of invasive species or the stringent filtering of species traits. At the opposite extreme of scale, geographically distant cities may exhibit low beta diversity due to the characteristic urban ecological conditions shared by all cities. Within cities, however, or along urban-rural gradients, urban land use may be associated with increased inter-patch biological diversity due to the extreme heterogeneity of current and historical land use found in urban areas. We suggest that the inter-patch scale of honey bee foraging combined with the inter-patch diversity of urban landscapes explains the increased diversity of pollen samples collected at our urban sites.

It is evident in Figure 4 that certain genera occur in our system as characteristically urban or agricultural taxa, while others are ubiquitous across the whole landscape gradient. The most

striking example of the latter is *Trifolium*, which was found in high abundance in every pollen sample. *Plantago*, while less dominant, was similarly ubiquitous across sites. In both cases, this reflects the commonness of these genera both in agricultural interstices and in urban lawns and roadsides. In contrast, *Chamaecrista* was found almost exclusively in FSR and CC, sites partly or entirely agricultural. Conversely, *Hedera*, a naturalized ornamental that has become widespread in urban greenspaces, was found almost exclusively in the urban sites BK and GH. Similarly, *Hydrangea* is a frequently cultivated ornamental, and was found mainly at sites BK and GH. These patterns highlight the joint importance of incidental and intentional effects of human land use on floral community composition.

Humans have mediated fundamental transitions in the distributions and abundances of both managed and non-managed species (Achard et al. 2002, Keeley 2002, Foley et al. 2005, Klimaszewski-Patterson and Mensing 2016). Such community compositional change continues at a rapid pace (Seebens et al. 2015, Tarolli 2016, Seebens et al. 2017) and is often driven broadly by landscape-scale changes (Tarolli et al. 2014, Otto et al. 2016). Understanding how alterations in landscape structure and composition affect the composition and functioning of communities is a central goal in applied ecology. Nutritional limitation is likely a bottom-up factor involved in shaping biological communities, especially with respect to pollinators (Samuelson et al. 2018, Baldock et al. 2019, Wood et al. 2019). Thus, a greater understanding of pollinator landscape and nutritional ecology is central to improving pollinator health. In applying improved molecular pollen identification methods to honey bee collected pollen, we show that local landscape composition alters forage diversity, composition and temporal dynamics along a

land use gradient. These findings provide direction for future integration of foraging ecology, landscape ecology and nutritional ecology on a broader range of taxa and environmental gradients.

Acknowledgements

The authors thank D. Denlinger and K. Goodell for laboratory space, C Kurkul for sample processing and N Douridas for assistance at the Molly Caren Agricultural Center. This work was funded by a Project Apis m.—Costco Honey Bee Biology Fellowship, an OARDC SEEDS grant and support provided by state and federal funds appropriated to The Ohio State University, Ohio Agricultural Research and Development Center (OHO01277).

References

- Achard, F., H. D. Eva, H.-J. Stibig, P. Mayaux, J. Gallego, T. Richards, and J.-P. Malingreau. (2002). Determination of deforestation rates of the world's humid tropical forests. *Science*, 297, 999-1002.
- Aronson, M. F. J., S. N. Handel, I. P. La Puma, and S. E. Clemants. (2015). Urbanization promotes non-native woody species and diverse plant assemblages in the New York metropolitan region. *Urban Ecosystems*, 18, 31-45.
- Aronson, M. F. J., F. A. L. Sorte, C. H. Nilon, M. Katti, M. A. Goddard, C. A. Lepczyk, ...M. Winter. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20133330.
- Baldock, K. C. R., M. A. Goddard, D. M. Hicks, W. E. Kunin, N. Mitschunas, H. Morse, ... J. Memmott. (2019). A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nature Ecology & Evolution*, 3, 363-373.
- Bell, K. L., K. S. Burgess, J. C. Botsch, E. K. Dobbs, T. D. Read, and B. J. Brosi. (2019). Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Molecular Ecology*, 28, 431-455.
- Bengtsson-Palme, J., M. Hartmann, K. M. Eriksson, C. Pal, K. Thorell, D. G. J. Larsson, and R. H. Nilsson. (2015). metaxa2: improved identification and taxonomic classification of small and large subunit rRNA in metagenomic data. *Molecular Ecology Resources*, 15, 1403-1414.
- Bengtsson-Palme, J., R. T. Richardson, M. Meola, C. Wurzbacher, É. D. Tremblay, K. Thorell, ... R. H. Nilsson. (2018). Metaxa2 Database Builder: Enabling taxonomic identification from metagenomic or metabarcoding data using any genetic marker. *Bioinformatics*, 34, 4027-4033.

- Chen, S., H. Yao, J. Han, C. Liu, J. Song, L. Shi, ... C. Leon. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLOS ONE*, 5, e8613.
- Couvillon, Margaret J., R. Schürch, and Francis L. W. Ratnieks. (2014a). Dancing bees communicate a foraging preference for rural lands in high-level agri-environment schemes. *Current Biology*, 24, 1212-1215.
- Couvillon, M. J., R. Schürch, and F. L. W. Ratnieks. (2014b). waggle dance distances as integrative indicators of seasonal foraging challenges. *PLOS ONE*, 9, e93495.
- Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science*, 14, 927-930.
- Döke, M. A., C. M. McGrady, M. Otieno, C. M. Grozinger, and M. Frazier. (2018). Colony size, rather than geographic origin of stocks, predicts overwintering success in honey bees (Hymenoptera: Apidae) in the Northeastern United States. *Journal of Economic Entomology*, 112, 525-533.
- Donaldson-Matasci, M. C., and A. Dornhaus. (2012). How habitat affects the benefits of communication in collectively foraging honey bees. *Behavioral Ecology and Sociobiology*, 66, 583-592.
- Eddy, S. R. (2011). Accelerated profile HMM searches. *PLOS Computational Biology*, 7, e1002195.
- Foley, J. A., R. DeFries, G. P. Asner, C. Barford, G. Bonan, S. R. Carpenter, ... P. K. Snyder. (2005). Global consequences of land use. *Science*, 309, 570-574.
- Garbuzov, M., R. Schürch, and F. L. W. Ratnieks. (2015). Eating locally: dance decoding demonstrates that urban honey bees in Brighton, UK, forage mainly in the surrounding urban area. *Urban Ecosystems*, 18, 411-418.
- Grüter, C., and W. M. Farina. (2009). The honeybee waggle dance: Can we follow the steps? *Trends in Ecology & Evolution*, 24, 242-247.
- Hallett, L. M., S. K. Jones, A. A. M. MacDonald, M. B. Jones, D. F. B. Flynn, J. Ripplinger, P. Slaughter, C. Gries, and S. L. Collins. (2016). codyn: An r package of community dynamics metrics. *Methods in Ecology and Evolution*, 7, 1146-1151.
- Härtel, S., and I. Steffan-Dewenter. (2014). Ecology: Honey bee foraging in human-modified landscapes. *Current Biology*, 24, R524-R526.
- Homer, C. G., J. Dewitz, L. Yang, S. Jin, P. Danielson, G. Z. Xian, J. Coulston, N. Herold, J. Wickham, and K. Megown. (2015). Completion of the 2011 National Land Cover Database for the conterminous United States – Representing a decade of land cover change information. *Photogrammetric Engineering and Remote Sensing*, 81, 345-354.
- Kamil Slowikowski, A. S., Sean Hughes, et al. (2019). ggrepel: Automatically position non-overlapping text labels with 'ggplot2'. R package version 0.8.1.
- Katoh, K., K. Misawa, K. i. Kuma, and T. Miyata. (2002). MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research*, 30, 3059-3066.

- Keeley, J. E. (2002). Native American impacts on fire regimes of the California coastal ranges. *Journal of Biogeography*, 29, 303-320.
- Keller, A., N. Danner, G. Grimmer, M. Ankenbrand, K. von der Ohe, W. von der Ohe, S. Rost, S. Härtel, and I. Steffan-Dewenter. (2015). Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biology*, 17, 558-566.
- Klimaszewski-Patterson, A., and S. A. Mensing. (2016). Multi-disciplinary approach to identifying Native American impacts on Late Holocene forest dynamics in the southern Sierra Nevada range, California, USA. *Anthropocene*, 15, 37-48.
- Kress, W. J., and D. L. Erickson. (2007). A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding *trnH-psbA* spacer region. *PLoS ONE*, 2, e508.
- Lau, P., V. Bryant, J. D. Ellis, Z. Y. Huang, J. Sullivan, D. R. Schmehl, A. R. Cabrera, and J. Rangel. (2019). Seasonal variation of pollen collected by honey bees (*Apis mellifera*) in developed areas across four regions in the United States. *PLoS ONE*, 14, e0217294.
- Lecocq, A., P. Kryger, F. Vejsnæs, and A. Bruun Jensen. (2015). Weight watching and the effect of landscape on honeybee colony productivity: Investigating the value of colony weight monitoring for the beekeeping industry. *PLoS ONE*, 10, e0132473.
- Lucek, K., A. Galli, S. Gurten, N. Hohmann, A. Maccagni, T. Patsiou, and Y. Willi. (2019). Metabarcoding of honey to assess differences in plant-pollinator interactions between urban and non-urban sites. *Apidologie*, 50, 317-329.
- McKinney, M. L. (2008). Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosystems*, 11, 161-176.
- Nürnberg, F., A. Keller, S. Härtel, and I. Steffan-Dewenter. (2019). Honey bee waggle dance communication increases diversity of pollen diets in intensively managed agricultural landscapes. *Molecular Ecology*, 28, 3602-3611.
- O'Rourke, M. K., and S. L. Buchmann. (1991). Standardized analytical techniques for bee-collected pollen. *Environmental Entomology*, 20, 507-513.
- Ohio Supercomputer Center (1987). *Citation*. Columbus, Ohio, USA. Retrieved from <http://osc.edu/ark:/19495/f5s1ph73>
- Otto, C. R. V., C. L. Roth, B. L. Carlson, and M. D. Smart. (2016). Land-use change reduces habitat suitability for supporting managed honey bee colonies in the Northern Great Plains. *Proceedings of the National Academy of Sciences*, 113, 10430-10435.
- Palmieri, L., E. Bozza, and L. Giongo. (2009). Soft fruit traceability in food matrices using real-time PCR. *Nutrients*, 1, 316-328.
- Richardson, R. T., H. R. Curtis, E. G. Matcham, C.-H. Lin, S. Suresh, D. B. Sponsler, L. E. Hearon, and R. M. Johnson. (2019). Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. *Molecular Ecology*, 28, 686-697.

- Richardson, R. T., C.-H. Lin, J. O. Quijia, N. S. Riusech, K. Goodell, and R. M. Johnson. (2015). Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Applications in Plant Sciences*, 3, 1500043.
- Richardson, R. T., D. B. Sponsler, H. McMinn-Sauder, and R. M. Johnson. (2019). MetaCurator: A hidden Markov model-based toolkit for extracting and curating sequences from taxonomically-informative genetic markers. *Methods in Ecology and Evolution*, <https://doi.org/10.1111/2041-210X.13314>
- Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584.
- Samuelson, A. E., R. J. Gill, M. J. F. Brown, and E. Leadbeater. (2018). Lower bumblebee colony reproductive success in agricultural compared with urban environments. *Proceedings of the Royal Society B: Biological Sciences*, 285, 20180807.
- Schürch, R., and C. Grüter. (2014). Dancing bees improve colony foraging success as long-term benefits outweigh short-term costs. *PLoS ONE*, 9, e104660.
- Seebens, H., T. M. Blackburn, E. E. Dyer, P. Genovesi, P. E. Hulme, J. M. Jeschke, ... F. Essl. (2017). No saturation in the accumulation of alien species worldwide. *Nature Communications*, 8, 14435.
- Seebens, H., F. Essl, W. Dawson, N. Fuentes, D. Moser, J. Pergl, ... B. Blasius. (2015). Global trade will accelerate plant invasions in emerging economies under climate change. *Global Change Biology*, 21, 4128-4140.
- Seeley, T. D. (1995). The wisdom of the hive: The social physiology of honey bee colonies.
- Sickel, W., M. J. Ankenbrand, G. Grimmer, A. Holzschuh, S. Härtel, J. Lanzen, I. Steffan-Dewenter, and A. Keller. (2015). Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecology*, 15, 20.
- Smart, M. D., R. S. Cornman, D. D. Iwanowicz, M. McDermott-Kubeczko, J. S. Pettis, M. S. Spivak, and C. R. V. Otto. (2017). A comparison of honey bee-collected pollen from working agricultural lands using light microscopy and ITS metabarcoding. *Environmental Entomology*, 46, 38-49.
- Smart, M. D., J. S. Pettis, N. Euliss, and M. S. Spivak. (2016). Land use in the Northern Great Plains region of the U.S. influences the survival and productivity of honey bee colonies. *Agriculture, Ecosystems & Environment*, 230, 139-149.
- Sponsler, D. B., and R. M. Johnson. (2015). Honey bee success predicted by landscape composition in Ohio, USA. *PeerJ* 3, e838.
- Sponsler, D. B., E. G. Matcham, C.-H. Lin, J. L. Lanterman, and R. M. Johnson. (2017). Spatial and taxonomic patterns of honey bee foraging: A choice test between urban and agricultural landscapes. *Journal of Urban Ecology*, 3.
- Steffan-Dewenter, I., and A. Kuhn. (2003). Honeybee foraging in differentially structured landscapes. *Proceedings of the Royal Society B: Biological Sciences*, 270, 569-575.
- Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. (1991). Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, 17, 1105-1109.

- Tarolli, P. (2016). Humans and the Earth's surface. *Earth Surface Processes and Landforms*, 41, 2301-2304.
- Tarolli, P., V. Vanacker, H. Middelkoop, and A. G. Brown. (2014). Landscapes in the Anthropocene: State of the art and future directions. *Anthropocene*, 6, 1-2.
- Team, Q. D. (2015). QGIS Geographic Information System. Open Source Geospatial Foundation Project.
- Team, R. C. (2014). R: A language and environment for statistical computing. Vienna, Austria.
- Visscher, P. K., and T. D. Seeley. (1982). Foraging strategy of honeybee colonies in a temperate deciduous forest. *Ecology*, 63, 1790-1801.
- Waddington, K. D., T. J. Herbert, P. K. Visscher, and M. R. Richter. (1994). Comparisons of forager distributions from matched honey bee colonies in suburban environments. *Behavioral Ecology and Sociobiology*, 35, 423-429.
- White, T. J., Bruns, T.D., Lee, S.B. and Taylor, J.W. (1990). Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. Pages 315-322 in M. A. Innis, Gelfand, D.H., Sninsky, J.J. and White, T.J., editor. PCR Protocols, a Guide to Methods and Applications. Academic Press, New York.
- Wood, T. J., J. Gibbs, K. K. Graham, and R. Isaacs. (2019). Narrow pollen diets are associated with declining Midwestern bumble bee species. *Ecology*, 100, e02697.
- Zhang, J., K. Kobert, T. Flouri, & A. Stamatakis. (2013). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30, 614-620.

Data Accessibility

All sequence data produced in this work can be downloaded from NCBI SRA, BioProject PRJNA574766. Descriptions of the bioinformatic analysis, R and python code, reference sequence databases, and supplemental excel files of the taxonomic inferences and the sequencing coverage per library can be found at <https://github.com/RTRichar/UrbanRuralPollen>.

Author Contributions

Conceptualization, DBS, RTR, RMJ; Investigation, GC, TE, CHL, RTR; Formal Analysis, RTR, DBS; Resources, RMJ, RTR; Writing – Original Draft, RTR, DBS; Writing – Review & Editing, RTR, DBS, RMJ, CHL, GC, TE; Visualization, RTR, DBS; Supervision, RMJ, RTR, DBS, CHL; Project Administration, RMJ; Funding Acquisition, RMJ, RTR, DBS.

Figures

Figure 1: A landcover map of Columbus, OH and surroundings. Black dots indicate apiary locations and circles are drawn at a 3 km radius surrounding each apiary.

Figure 2: Molecular proportions regressed against microscopic proportions for the nine samples analyzed by pollen morphology. Taxa are shown with unique colors and point types with the exception of minor taxa, those which never exceeded 2.5 percent relative abundance according to microscopy. The comparison was performed at the genus level with the exception of Asteraceae, for which genera are difficult to distinguish morphologically.

Figure 3: Proportional abundances of top 20 families detected, as measured by mean frequency of occurrence across sites for 2013 and 2014 separately. Families are ordered by decreasing frequency of detection for visualization purposes. Panels are arranged from the most agricultural site (FSR) to the most urban site (GH).

Figure 4: Proportional abundances of the top 35 genera detected, as measured by mean frequency of occurrence across sites for 2013 and 2014 separately. Genera are grouped by family for visualization purposes. Panels are arranged from the most agricultural site (FSR) to the most urban site (GH).

Figure 5: Time series and boxplots of Shannon diversity, mean rank shift and qualitative turnover (appearances plus disappearances between consecutive samplings) for 2013 and 2014. Open triangles in boxplots represent distribution averages. Boxplots are arranged from most urban (GH) to most agricultural (FSR).

Figure 6: An NMDS biplot ($k = 2$, stress = 0.06) of forage compositional similarity between sites and years. Only genera with a maximum proportional abundance of at least 2.5 percent are plotted. Genus scores are plotted using weighted averages.