

Article

Diverse Diets with Consistent Core Microbiome in Wild Bee Pollen Provisions

Rebecca M. Dew ¹, Quinn S. McFrederick ² and Sandra M. Rehan ^{1,*} ¹ Department of Biology, York University, Toronto, ON M3J 1P3, Canada; rebecca.dew11@gmail.com² Department of Entomology, University of California Riverside, Riverside, CA 92521, USA; quinnmc@ucr.edu

* Correspondence: sandra.rehan@gmail.com; Tel.: +1-416-736-2100

Received: 25 May 2020; Accepted: 29 July 2020; Published: 4 August 2020



Abstract: Bees collect pollen from flowers for their offspring, and by doing so contribute critical pollination services for our crops and ecosystems. Unlike many managed bee species, wild bees are thought to obtain much of their microbiome from the environment. However, we know surprisingly little about what plant species bees visit and the microbes associated with the collected pollen. Here, we addressed the hypothesis that the pollen and microbial components of bee diets would change across the range of the bee, by amplicon sequencing pollen provisions of a widespread small carpenter bee, *Ceratina calcarata*, across three populations. *Ceratina calcarata* was found to use a diversity of floral resources across its range, but the bacterial genera associated with pollen provisions were very consistent. *Acinetobacter*, *Erwinia*, *Lactobacillus*, *Sodalis*, *Sphingomonas* and *Wolbachia* were among the top ten bacterial genera across all sites. *Ceratina calcarata* uses both raspberry (*Rubus*) and sumac (*Rhus*) stems as nesting substrates, however nests within these plants showed no preference for host plant pollen. Significant correlations in plant and bacterial co-occurrence differed between sites, indicating that many of the most common bacterial genera have either regional or transitory floral associations. This range-wide study suggests microbes present in brood provisions are conserved within a bee species, rather than mediated by climate or pollen composition. Moving forward, this has important implications for how these core bacteria affect larval health and whether these functions vary across space and diet. These data increase our understanding of how pollinators interact with and adjust to their changing environment.

Keywords: pollination; microbiome; foraging; bees; *Ceratina*; bacteria

1. Introduction

The broader community of microbes within a host, the microbiome, can determine the health status of an individual. Many microbes provide beneficial functions for the host including metabolism and immunity [1,2]. In honey bees, certain *Lactobacillus* strains offered protection against a microsporidian and bacterial pathogen [3]. Similarly, in bumble bees, increased microbiome diversity was linked to reduced infection by the trypanosomatid parasite *Crithidia* [4]. In *Osmia ribofloris*, the pollen provision microbiome is crucial for larval development [5]. Therefore, it is important to characterize and understand the microbiome to understand bee health.

Our current knowledge on the microbiome of bees is predominantly based on honey bees (*Apis* spp.), and to a lesser degree bumble bees (*Bombus* spp.) [6–8]. Both of these are highly social and closely related members of the corbiculate apid bees, and as such they share a very similar core microbiome [7]. Outside of these genera, the bee microbiomes sequenced so far do not conform to the *Apis* and *Bombus* models. Even within the corbiculates, the stingless bees (Meliponini) and the orchid bees (Euglossini) lack some of the most common symbionts of *Apis* and *Bombus*, although several related symbionts are shared amongst the corbiculates [7]. Looking more broadly, bacteria that

were previously classified as *Lactobacillus* but have been recently split into the genera *Apilactobacillus*, *Bombilactobacillus* and *Lactobacillus sensu strictu* [9] are some of the few symbionts common to multiple bee taxa including *Apis*, *Bombus*, the small carpenter bee *Ceratina*, megachilid and halictid bees [10–15]. Microbe acquisition in *Apis* and *Bombus* occurs within the hive, facilitated by nestmate interactions or transfer from feces [6,12,16–18]. While honey and bumble bees live in large groups, this level of sociality is rare among bees, the vast majority of bee species being solitary [19]. Indeed, in the other bee species studied so far, much of their microbiota appears to be gained from the environment rather than through social transmission [11,13,14,20–22]. Therefore, differences in environmental and pollen-associated bacteria may have larger impacts on wild bee development than for the highly social corbiculate bee species.

Microbial acquisition from the environment may be influenced by the diet of the bee. As for bees, flowers harbor a variety of microbes, which can potentially be passed to foraging bees. *Crithidia* can be transmitted between foraging *Bombus* at flowers [23] and communities of pollinators have been found to share microbes [24,25]. For bees that use foliage to line their nests, both flower and foliar source affect their pollen provision microbiome [22]. However, there are also many more complex factors to consider such as flower morphology, volatiles and even the secondary compounds produced by the microbes themselves, that can alter floral bacterial communities and transmission to pollinators [26,27]. Therefore, diet may be an important factor to consider when looking at the wild bee microbiome, which is thought to be largely environmentally sourced.

To conserve wild bees, we need to understand their health, and their associated microbial symbionts. It seems likely that the microbes present in the environment, and therefore those gained environmentally by bees, will vary geographically with changes in climate, interacting insect species and floral communities.

Ceratina calcarata Robertson, 1900 is a small carpenter bee species that is a widespread and prominent pollinator across eastern North America [28–30]. This species nests in the dead stems of various plants, commonly raspberry and sumac (*Rubus* and *Rhus* species, respectively) [20,31]. The plants it nests in also produce flowers, potentially biasing pollen collection and thereby microbial acquisition. This bee constructs separate brood cells within the stem nest, each provisioned with a single pollen ball on which an egg is laid [31]. This brood provision is the only source of food given to the offspring until it reaches maturity. Study of these brood provisions from nests at the northern extent of its range found they contain multiple pollen species and a diversity of microbes dominated by *Lactobacillus*, *Wolbachia*, *Acinetobacter* and *Sodalis* [20]. However, *C. calcarata* is found across a broad geographic range in eastern North America and acquires at least part of its microbiome from the environment [20], therefore its microbiome may vary geographically with corresponding changes in climate and floral landscape.

The aim of this project was to investigate whether the microbiome of *C. calcarata* varies geographically by sequencing brood provisions spanning this species' range across the eastern United States. Specifically, we asked whether pollen or bacterial species vary in composition or diversity among sites and if there are identifiable plant/microbe associations. We also asked whether foraging was biased by the proximity of nest plant flowers.

2. Methods

2.1. Brood Provision Collection

Nests of *Ceratina calcarata* were collected from Athens, Georgia (33.95°19' N, 83.35°76' W), Lake Ozark, Missouri (38.19°86' N, 92.63°88' W) and Durham, New Hampshire, USA (43.13°39' N, 70.92°64' W). These sites cover the geographic and climatic range of this bee within the USA (Figure 1). Nests were collected in dead stems of *Rubus* spp. (raspberry) and *Rhus* spp. (sumac) from May to July 2016. The stems were split open and brood provisions from each nest were removed with

flame-sterilized forceps and placed into separate cryovials. These were stored at -80°C prior to DNA extraction.

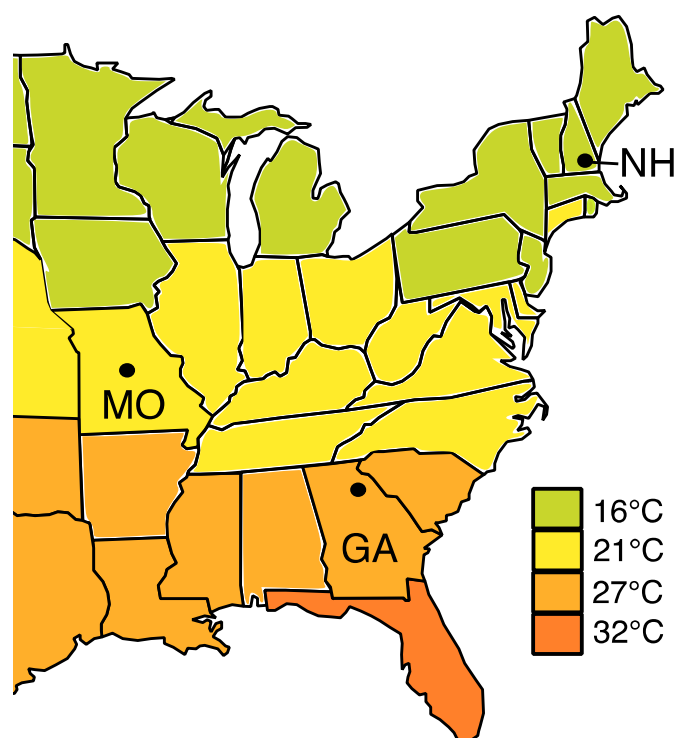


Figure 1. Map showing the average summer temperatures ($^{\circ}\text{C}$) across the eastern United States. Black dots indicate the sampled regions of GA: Georgia, MO: Missouri and NH: New Hampshire. Map was modified from the National Oceanic and Atmospheric Administration (<http://www.noaa.gov/climate>).

2.2. DNA Extraction and Sequencing

DNA extraction was performed with the DNeasy blood and tissue kit (Qiagen, Valencia, CA, USA). Each sample was homogenized with a sterile steel bead (5 mm) and 100 μL of sterile glass beads (0.1 mm) in 180 μL ATL buffer using a Qiagen tissue lyser. An aliquot of 20 μL proteinase K was added to the homogenized samples, which was then incubated overnight at 56°C . Three negative controls containing sterile water were simultaneously prepared.

PCR reactions targeted the 16S ribosomal RNA gene to sequence bacterial species, and the ribulose biphosphate carboxylase large subunit (*rbcl*) to sequence pollen species. Primers and PCR assays for both genes followed the protocols outlined by McFrederick and Rehan [20]. The products of the PCR were cleaned using the PureLink Pro Purification Kit (Invitrogen, Carlsbad, CA, USA). Then, a second PCR was performed with 1 μL of each of the cleaned PCR1 products as template to add the Illumina adapters. The PCR2 products were then normalized by running 18 μL of the product through SequelPrep normalization plates (Invitrogen, Carlsbad, CA, USA). The pooled normalized product was cleaned a second time via a speed bead cleaning method before the libraries were quality checked with a 2100 Bioanalyser (Agilent, Santa Clara, CA, USA). Sequencing was performed with the MiSeq Reagent Kit v3, using 2×300 cycles. The raw sequences can be accessed on the NCBI Sequence Read Archive (SRA) under accession number PRJNA454884.

3. Data Analysis

The data for both 16S and *rbcl* were demultiplexed in Qiime2 v.2017.10 [32]. Quality filtering and exact sequence variant (ESV) clustering were performed with the Qiime2 plug-in Dada2 [33]. The 16S rRNA gene reads were classified in Qiime2 using a naïve Bayes classifier trained on the Greengenes 99% database [34]. The resulting taxonomy was filtered to remove all mitochondrial and chloroplast

genes. The insect symbionts *Wolbachia* and *Sodalis* were recovered in reads. It is unlikely these were obtained from flowers, more likely coming from the mother or mites in the nest. We chose to keep these reads, as we do not know the source of any of the bacteria found in the pollen provision, or whether these are metabolically active, only that they are present. In this way, we give a complete snapshot of the community present.

There were some bacterial reads in one of the negative controls. Most of these were common laboratory contaminants and were removed from all reads [35]. However, *Sodalis* was found in 6.140% of the control reads. *Sodalis* has been previously identified in brood provisions of *C. calcarata* [20], and was present in 40 of our brood provision samples, with an average of 8.61% reads per sample. Given the high frequency of *Sodalis* in some samples, it is likely to represent a true component of the microbial community in the brood provisions. To compensate for this, we analysed the data in two separate ways, (i) removing all *Sodalis* reads from samples, and (ii) removing a portion of *Sodalis* from each sample, up to a maximum of 6.140% of the sample's reads, to correspond to the percentage of contamination present in the control. We did not use this adjusted ratio of *Sodalis* in analyses of abundance, as this is reliant on accurate read counts. Alpha and beta diversity analyses were performed on both of these datasets at a read depth of 139 reads, allowing inclusion of 87% and 89% of samples in each dataset, respectively. Read depth was shallow, as the quality and filtering steps removed a large proportion of reads, but the rarefaction curves plateaued at this depth, indicating that bacterial diversity was truly low in these samples (Figure 2).

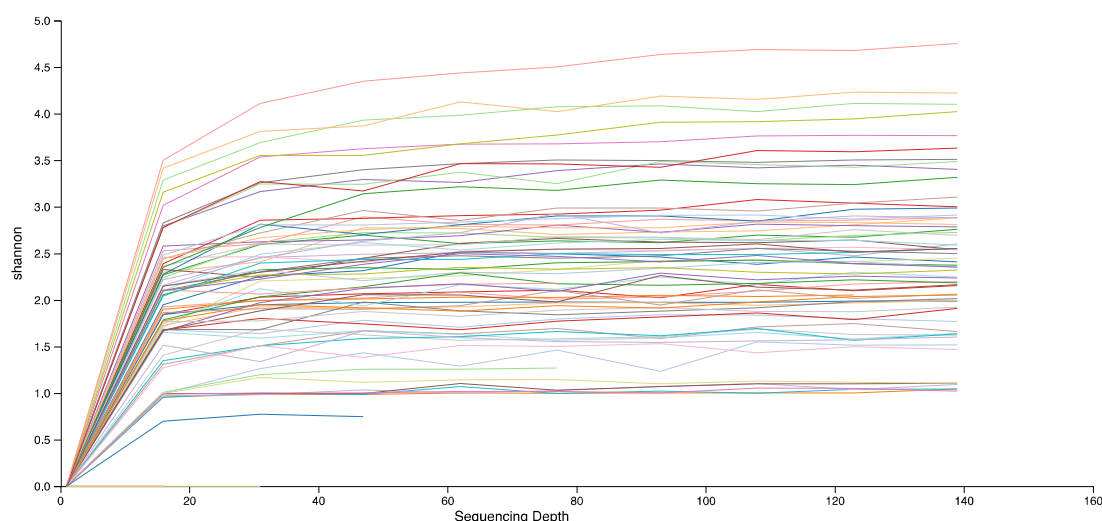


Figure 2. Rarefaction curve of Shannon diversity for the 16S reads for each sample to a read depth of 139 reads. Based on the dataset with the *Sodalis* reads removed.

For *rbcl*, read quality was low for all New Hampshire sequences, so these were not analysed further. However, *C. calcarata* pollen balls collected from New Hampshire in July 2014 were previously sequenced for *rbcl* by McFrederick and Rehan [20]. In that study, the sequences were analysed in Qiime1 using 97% OTU matching. This software has now been superseded by Qiime2, which employs exact sequence variant matching. As part of this change, the Qiime2 program also implements more stringent quality filtering. Here, we reanalyse the *rbcl* data from McFrederick and Rehan [20] in Qiime2, to allow for statistical comparison to the *rbcl* data generated in the current study for Missouri and Georgia. The New Hampshire reads went through demultiplexing, quality filtering and ESV binning separately, to account for sequencing run differences before amalgamation with the Georgia and Missouri *rbcl* reads.

The *rbcl* ESVs from Georgia, Missouri and New Hampshire that remained after quality filtering were assigned taxonomy based on local BLAST searches with 99% sequence identity [36]. All plant genera returned by BLAST searches were confirmed to be in Georgia and Missouri through searches

in the USDA PLANTS Database (<http://plants.usda.gov/checklist.html>) and the Missouri Botanical Gardens Plant Finder (<http://www.missouribotanicalgarden.org>). Diversity analyses were performed at a depth of 3791 reads, preserving 79% of samples.

We considered Faith's phylogenetic diversity as a measure of taxonomic richness; comparisons between groups were made with Kruskal–Wallis analysis (K–W). Abundance data were compared through PERMANOVA analysis of both Bray–Curtis dissimilarity (B–C) and weighted unifracs (W–U) indices as measures of beta-diversity and phylogenetic beta-diversity, respectively. Additional statistical comparisons were performed with Kruskal–Wallis and Wilcoxon tests in R v.1.0.136.

Co-associations between the pollen and bacterial components of the brood provisions were assessed using SparCC [37] and CoNet [38]. Both programs were used as SparCC and CoNet use correlation and dissimilarity methods, respectively, and the differences in the algorithms between these two approaches can lead to different results [39]. New Hampshire was not considered due to the exclusion of the *rbcl* data for this site, and these analyses have been previously reported for New Hampshire in [20]. Analyses were run on a combined dataset of all bacterial and plant ESVs identified to genus level from Georgia and Missouri. This was then repeated on separate datasets of reads from Georgia and Missouri separately to look for site differences in co-associations. Pseudo p-values in SparCC were calculated based on 100 bootstrap replicates. In CoNet, network edge scores were calculated with Pearson, Spearman, mutual information, Bray–Curtis and Kullback–Leibler. Bootstraps were calculated using Brown p-value merging and Benjamini–Hochberg multiple test correction. Positive correlations were only considered if recovered in both SparCC and CoNet with $p \leq 0.01$.

4. Results

After trimming, quality filtering and chimera removal, 350,472 sequences of 16S were obtained for 189 ESVs across 65 samples. This dropped to 250,810 reads in 62 samples when *Sodalis* was excluded due to its presence in the negative controls. There were 299,489 *rbcl* sequences for 249 ESVs across 79 samples from Missouri and Georgia, with a total of 44 samples with sequences for both gene regions. For the New Hampshire *rbcl* data reanalysed from [20], there were 356,354 *rbcl* sequences for 226 ESVs across 94 samples.

The Missouri and New Hampshire pollen *rbcl* samples were exclusively from nests formed in sumac plants, but the Georgia samples comprised 35 nests in raspberry and 22 nests in sumac (Table S1). The 62 samples successfully sequenced for 16S included 24 nests from Missouri, 31 from Georgia and 7 from New Hampshire. Nests from Missouri were exclusively from sumac, and New Hampshire nests were from raspberry ($n = 5$) and sumac ($n = 2$), as were the nests from Georgia (raspberry: $n = 17$; sumac: $n = 14$; Tables S2 and S3).

In total, 58 bacterial ESVs were classified through to genus, however the provisions were dominated by just a few of these genera. In each site, >96% of the sample reads were from the 10 most common genera. *Acinetobacter*, *Erwinia*, *Lactobacillus*, *Sodalis*, *Sphingomonas* and *Wolbachia* were found among those top 10 at all sites. For the analyses using an adjusted percentage of *Sodalis*, this was the most frequent genus in Missouri, covering 51% of the reads (Figure 3). The second most common genus at that site was *Wolbachia*. For Georgia, *Lactobacillus* was the top genus, having 32% of the reads (Figure 3). There were very few bacterial genera unique to each state: Georgia had nine unique genera, Missouri five and New Hampshire one. Consequently, phylogenetic microbial richness did not significantly differ among states (K–W: $N_{GA} = 31$, $N_{MO} = 27$, $N_{NH} = 7$, $H = 3.57$, $P = 0.167$) or with nesting substrate ($N_R = 22$, $N_S = 40$, $H = 0.966$, $P = 0.318$). The results were the same if *Sodalis* was excluded (states: $N_{GA} = 31$, $N_{MO} = 24$, $N_{NH} = 7$, $H = 0.867$, $P = 0.684$; nesting substrate: $N_R = 22$, $N_S = 40$, $H = 0.146$, $P = 0.702$). However, when abundance was considered in the analyses excluding *Sodalis*, the microbial composition of the brood provisions did significantly differ between Georgia nests in raspberry and the Missouri nests, all of which were in sumac (B–C: pseudo- $F_{41} = 4.63$, $q = 0.010$, W–U: pseudo- $F_{41} = 6.23$, $q = 0.040$). There was also a significant difference in phylogenetic diversity between Georgia nests in sumac and Missouri nests, though this was not significant for

non-phylogenetic diversity (W-U: pseudo- $F_{38} = 4.45$, $q = 0.045$; B-C: pseudo- $F_{38} = 2.90$, $q = 0.065$). The bacterial composition of brood provisions did not significantly differ between all other pairwise comparisons, including comparisons of Georgia and Missouri with New Hampshire.

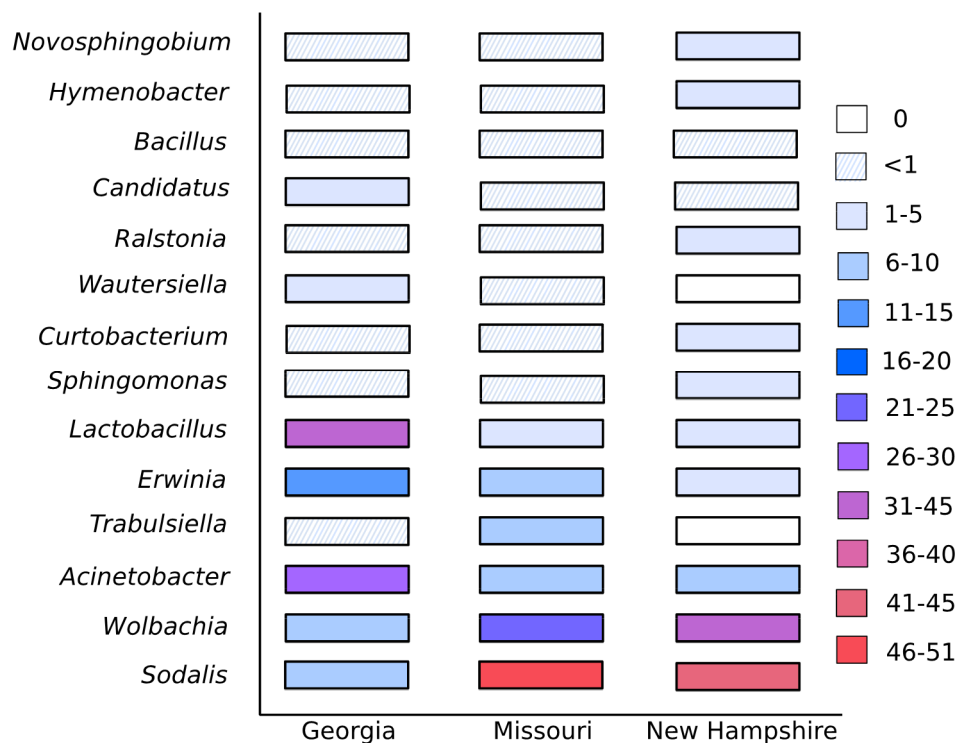


Figure 3. The percentage of 16S reads found across the top 14 bacterial genera in brood provisions from each site. Many bacteria were common to all sites but the relative abundance of the most common bacteria varied broadly.

There were in total 96 plant genera present in the brood provisions. Brood provisions in Georgia contained 52 plant genera, with *Liriodendron tulipifera* and *Rubus spp.* accounting for 31% and 28% of the reads, respectively (Figure 4). Missouri had 40 genera dominated by *Diospyros* (38%) and *Gleditsia* (15%; Figure 4). New Hampshire had 65 genera, predominantly consisting of *Rhamnus* (36%) and *Rhus* (32%). Phylogenetic richness was significantly higher in New Hampshire (mean = 0.44, s.e. = 0.23) when compared to Georgia (mean = 0.41, s.e. = 0.31; K-W: $N_{NH} = 94$, $N_{GA} = 57$, $H = 6.38$, $q = 0.035$). Neither Georgia nor New Hampshire significantly differed in richness from Missouri (mean = 0.40, s.e. = 0.29; K-W: $N_{NH} = 94$, $N_{MO} = 22$, $H = 1.64$, $q = 0.300$; $N_{GA} = 57$, $N_{MO} = 22$, $H = 0.62$, $P = 0.431$). There was also no significant difference in phylogenetic richness with nesting substrate (K-W: $N_{RUBUS} = 35$, $N_{RHUS} = 138$, $H = 0.91$, $q = 0.341$). However, the number of pollen genera in each individual brood provision was higher in Missouri (mean: 10.24 ± 1.3 s.e.) compared to Georgia (mean: 6.67 ± 0.5 s.e.; pairwise Wilcoxon, Bonferroni correction: $P = 0.024$), while the number of pollen genera in New Hampshire provisions did not significantly differ from either Missouri ($P = 0.305$) or Georgia ($P = 0.225$).

Brood provisions significantly differed in pollen beta-diversity between Missouri and New Hampshire (B-C: pseudo- $F_{116} = 21.53$, $q = 0.001$; W-U: pseudo- $F_{116} = 32.59$, $q = 0.001$), Missouri and Georgia (B-C: pseudo- $F_{79} = 12.35$, $q = 0.001$; W-U: pseudo- $F_{79} = 12.87$, $q = 0.001$) and New Hampshire and Georgia (B-C: pseudo- F_{151} , $q = 0.001$; W-U: pseudo- $F_{151} = 35.25$, $q = 0.001$). No significant differences in beta-diversity were found between *Rubus* and *Rhus* nests within Georgia (B-C: pseudo- $F_{57} = 0.82$, $q = 0.517$, W-U: pseudo- $F_{57} = 0.34$, $q = 0.742$). To test if pollen of nest plants is preferentially collected, we compared the number of *Rubus* reads between nests formed in *Rubus* or *Rhus* stems in Georgia, and we found no significant difference (Wilcoxon: $W = 421.5$, $p = 0.555$).

There was also no significant difference in the amount of *Rhus* pollen collected (Wilcoxon: $W = 365$, $p = 0.680$).

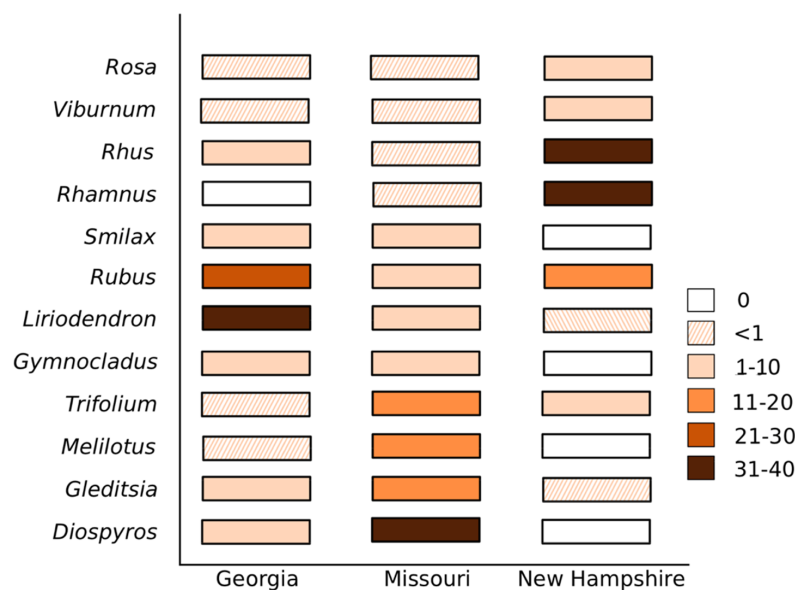


Figure 4. The percentage of rbcl reads found across the top 12 plant genera in brood provisions from each site. The most abundant plant genera varied between each state.

Significant positive correlations between plant and bacterial genera were not consistent between the analysis of the overall dataset and state-level analyses (Table 1). In the overall analysis of Georgia and Missouri combined, a significant correlation was found between the plant genus *Liriodendron* and *Lactobacillus*. Analysis of Georgia separately gave a significant correlation between *Liriodendron* and *Sphingomonas*. Within the Missouri data, a number of significant correlations were found, though none involving *Liriodendron*. For Missouri, the plant genera *Brunia*, *Camptotheca*, *Rhus* and *Smilax* were all correlated with *Wolbachia*. The plants *Gleditsia* and *Gymnocladus* were correlated with *Trabulsiella*, while *Trifolium* was correlated with *Wautersiella*. In New Hampshire, based on data presented in [20], correlations were found between *Gleditsia* and *Rosa* with *Lactobacillus* and *Rubus* was correlated with *Acinetobacter* and *Sodalis* (Table 1).

Table 1. Correlations coefficients between plant and bacteria genera as identified by SparCC. Only significant positive correlations are shown that were also found to be significant in CoNet ($p \leq 0.01$). Superscripts notate significant correlations found when analyses were restricted to Georgia (^{GA}), Missouri (^{MO}) or New Hampshire (^{NH}).

Bacteria	Plants									
	<i>Brunia</i>	<i>Camptotheca</i>	<i>Gleditsia</i>	<i>Gymnocladus</i>	<i>Liriodendron</i>	<i>Rhus</i>	<i>Rosa</i>	<i>Rubus</i>	<i>Smilax</i>	<i>Trifolium</i>
<i>Acinetobacter</i>								0.287 ^{NH}		
<i>Lactobacillus</i>			0.231 ^{NH}		0.393		0.225 ^{NH}			
<i>Sodalis</i>								0.249 ^{NH}		
<i>Sphingomonas</i>					0.420 ^{GA}					
<i>Trabulsiella</i>			0.579 ^{MO}	0.578 ^{MO}						
<i>Wautersiella</i>										0.654 ^{MO}
<i>Wolbachia</i>	0.600 ^{MO}	0.587 ^{MO}				0.574 ^{MO}			0.566 ^{MO}	

5. Discussion and Conclusions

The floral resources utilized by *C. calcarata* differed between regions, brood provisions being dominated by different pollen genera in each state, showing this generalist bee's local adaptation to regional floral communities (Figure 4). Foraging preference was not biased towards flowers present on the host nest plant, indicating that spatial assortment of floral resources alone does not determine foraging preferences. Despite changes in floral resources, the same core microbes dominated brood provisions across all states, although the relative abundance of these groups did vary between region (Figure 3). Our overall network analyses identified some correlations in plant and bacterial occurrences, however the broad changes in floral diet between states did not correspond to large changes in the bacterial community, suggesting that these floral–bacterial associations are transient or non-obligate (Table 1).

There were a number of core bacterial genera found across all sites but the relative abundance of these varied strongly (Figure 3). Comparison to McFrederick and Rehan's [20] study in New Hampshire suggests there may also be annual variation within sites. In 2016, *Lactobacillus* was the most common genus, while we recovered *Sodalis* and *Wolbachia* as the top two genera in the current study (Figure 3). Being obligate facultative endosymbionts in many insect species, the occurrence of *Wolbachia* and *Sodalis* is likely due to contamination from mites or other parasites, or transfer directly from the mother's crop rather than through any specific floral sources [40,41]. A major limitation of amplicon sequencing studies is that they can only determine the presence or absence of an organism's DNA, not whether that organism is metabolically active, and this holds true for all bacteria recovered in our samples. This aside, it is interesting that *Wolbachia* contamination is so prevalent, and future microscopic examination of pollen material for mite infestation and tests of possible vertical transmission via pollen are needed. *Lactobacillus* was only present in 2.2% of reads in the New Hampshire samples for this study. Within Georgia, the abundance of bacterial ESVs differed between nests in *Rubus* and *Rhus* but significant differences disappeared when the phylogenetic similarity of bacteria was considered. This suggests there could be some differences in the abundance of microbial species or strains between nesting substrates but overall the taxonomic distribution of bacteria in the brood provisions is similar. Nesting substrate, therefore, has a smaller or negligible effect on bacterial abundance compared to differences among states.

Ceratina calcarata foraged from a slightly greater phylogenetic richness of floral genera in New Hampshire than in Georgia. The phylogenetic richness of floral genera in Missouri did not significantly differ from either Georgia or New Hampshire, perhaps because of its mid-lying geographic position and climate. As expected, our reanalysis of the New Hampshire data with 99% ESV matching recovered fewer genera than in McFrederick and Rehan [20], who found 110 genera compared to this study with 65. The genera not recovered in our reanalysis were all present in less than 1% of reads in the original study. We identified the same five genera as being the most abundant (*Rhamnus*, *Rhus*, *Rubus*, *Viburnum*, *Trifolium*), these genera accounting for 92% of the reads (Figure 4). Despite this more conserved estimate of genera, the floral resources used in New Hampshire are still rich compared to those utilized in Georgia. We also found that foraging females in Missouri foraged from more plants to form a single pollen provision mass than those in Georgia (on average 10.2 genera for females in Missouri compared to 6.7 for females in Georgia). This suggests that suitable floral resources at the time of brood provisioning may not be as diverse in Georgia as more northern areas of *Ceratina*'s range, or that they were simply not locally abundant in the area around the collected nests.

Across its geographic range, *C. calcarata* encounters a broad variety of possible forage. Diets in Georgia, Missouri and New Hampshire were dominated by pollen from different plant genera (Figure 4). Out of the 96 floral genera found in provisions in this study, only *Rubus* was found in more than 1% of reads across all three states (Figure 4). All other genera, even if abundant in one or two states, were rare in provisions from the third. For example, sumac was a key floral resource in New Hampshire but made up less than 10% of the reads in Georgia and was hardly utilized at all in Missouri (<1% of reads; Figure 4). It is important to note that while read counts have been correlated with

microscopy pollen counts in many studies [27,42–44], factors such as pollen morphology can skew the abundance estimate obtained from DNA sequences [45]. Our study uses the marker *rbcl*, which has shown strong correlation with pollen counts, outperforming *trnL* and ITS2 [27]. With this in mind, comparison of relative abundance between sites shows state-wise differences in diet. Many of these plant genera are common to all three states, so perhaps these dietary variations are due to differences in bee and floral phenologies, as well as possible microhabitat distinctions in floral assemblages in proximity to the bee nest.

While we do not have data on floral distributions within each collecting site, our records of nest substrate allow us to determine that foraging was not skewed towards the host plants. *Rubus* was a common pollen source but even nests formed within *Rubus* plants did not show a bias in pollen collection. Different pollens vary in nutritional qualities, which may influence foraging decisions [46]. Pollen can also have toxic constituents, and some generalist foragers appear to actively utilize a broad range of floral resources to alleviate the effects these may have on brood development [47,48]. How these factors influence *C. calcarata* foraging is unknown but our results suggest that spatial orientation of floral resources alone does not determine foraging preferences.

The presence of a consistent core microbial community despite the variation in pollen sources suggests that many of the most common bacterial genera do not have specific floral associations. We identified a number of tentative bacteria–plant correlations, but these were not consistent among states (Table 1). In the overall analysis, the tupliptree genus *Liriodendron* was correlated with *Lactobacillus*, while the same plant genus was correlated with *Sphingomonas* in Georgia. In Missouri, *Wolbachia* was correlated with four plant genera: *Brunia*, *Camptotheca*, *Rhus* and *Smilax* but this bacterium was not correlated with plants in the other states or the overall analysis. The correlations found in Georgia and Missouri also differ to those previously identified in New Hampshire, following the same methodology [20]. These correlations broadly suggest that plants and bacteria are co-occurring but the variance in results between the overall dataset and the state-level analyses indicates these relationships are facultative or transient. Using read data to identify co-occurrence correlations is statistically challenging [39,49] and further experiments sampling pollen bacterial communities with and without pollinator visitation, such as the study by McFrederick et al. [13], are needed to directly test for plant–bacteria associations.

Whether plants harbor certain microbes over others or not, there are many factors altering microbial floral communities. Long-term artificial warming of grassland plots was found to alter the microbial communities of plant leaves, including microbial groups common to bees [50]. Aydoğan et al. [50] found *Acinetobacter* and *Wolbachia* increased in frequency, while *Sphingomonas* frequency decreased, these three bacterial genera being common to *C. calcarata* pollen provision and adult gut microbiomes [10,20]. These temperature based microbial changes could translate into changes in insect microbiomes, and indeed climate has been correlated with changes in microbiome composition in some species such as the red palm weevil [51], the chestnut weevil [52] and a spider mite [53]. Flower visitation by bees can transfer microbes to flowers, but herbivorous insects, other pollinators including thrips and wind are thought to contribute to microbe dispersal as well [13,54–56]. Similarly, the presence of potentially predatory or competitive species such as ants can reduce floral visitation and this in turn alters the microbes present on flowers [57]. Any and all of these could be important factors influencing the observed microbiome variation in *C. calcarata* and are important considerations when concerned with wild bee health generally.

Our study shows that the diet of *C. calcarata* varies widely with geography, with only *Rubus* found in more than 1% of reads at all three sites, indicating that this generalist bee species is able to utilize different resources as floral communities change. However, it seems that floral preference may not be simply determined by the proximity of the floral resource to the nest. The same six bacterial genera consistently dominated provisions in all sites but the relative abundance of these fluctuated widely. There are still many unknowns regarding how microbes are acquired, both in the pollen provisions and subsequently the bees themselves. Flowers appear to be general points of bacterial transmission,

but so far specific associations have not been identified. The current lack of knowledge on microbial associates is a major hindrance in our ability to maintain diverse wild bee populations.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2075-4450/11/8/499/s1>, Table S1: Rbcl reads for each sample binned by genus, Table S2: 16S reads with the number of *Sodalis* reads reduced by 6.140%, Table S3: 16S reads with all *Sodalis* reads removed.

Author Contributions: S.M.R. designed the study. Q.S.M. and R.M.D. performed laboratory work. All authors have read and agreed to the published version of the manuscript.

Funding: Funding from an NSERC Discovery Grant to S.M.R., Endeavour Research Fellowship R.M.D., the University of California Riverside to Q.S.M., US Department of Agriculture NIFA Hatch funds (CA-R-ENT-5109-H) to Q.S.M., and Foundation for Food and Agriculture Research Pollinator Health Fund 549038 to S.M.R. supported this work.

Acknowledgments: We thank Sean Lombard, Sarah Lawson and Wyatt Shell for assistance with nest collections, Kaleigh Russell and Jason Rothman for help with DNA extractions and library preparation.

Conflicts of Interest: The authors declare no conflict of interest.

Data Accessibility: NCBI Sequence Read Archive (SRA) accession number PRJNA454884.

References

- Alberoni, D.; Gaggia, F.; Baffoni, L.; Di Gioia, D. Beneficial microorganisms for honey bees: Problems and progresses. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 9469–9482. [\[CrossRef\]](#)
- Engel, P.; Moran, N.A. The gut microbiota of insects—Diversity in structure and function. *FEMS Microbiol. Rev.* **2013**, *37*, 699–735. [\[CrossRef\]](#)
- Arredondo, D.; Castelli, L.; Porrini, M.P.; Garrido, P.M.; Eguaras, M.J.; Zunino, P.; Antúnez, K. *Lactobacillus kunkei* strains decreased the infection by honey bee pathogens *Paenibacillus larvae* and *Nosema ceranae*. *Benef. Microbes* **2018**, *9*, 279–290. [\[CrossRef\]](#)
- Mockler, B.K.; Kwong, W.K.; Moran, N.A.; Koch, H. Microbiome structure influences infection by the parasite *Crithidia bombi* in bumble bees. *Appl. Environ. Microbiol.* **2018**, *84*, 1–11. [\[CrossRef\]](#)
- Dharampal, P.S.; Carlson, C.; Currie, C.R.; Steffan, S. Pollen-borne microbes shape bee fitness. *Proc. R. Soc. B Biol. Sci.* **2019**, *286*, 20182894–99. [\[CrossRef\]](#)
- Koch, H.; Schmid-Hempel, P. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 19288–19292. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kwong, W.K.; Medina, L.A.; Koch, H.; Sing, K.W.; Soh, E.J.Y.; Ascher, J.S.; Moran, N.A. Dynamic microbiome evolution in social bees. *Sci. Adv.* **2017**, *3*, 1–17. [\[CrossRef\]](#)
- Kwong, W.K.; Moran, N.A. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* **2016**, *14*, 374–384. [\[CrossRef\]](#)
- Zheng, J.; Wittouck, S.; Salvetti, E.; Franz, C.M.; Harris, H.M.; Mattarelli, P.; O’Toole, P.W.; Pot, B.; Vandamme, P.; Walter, J.; et al. A taxonomic note on the genus *Lactobacillus*: Description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. *Int. J. Syst. Evol. Microbiol.* **2020**, *70*, 2782–2858. [\[CrossRef\]](#)
- Graystock, P.; Rehan, S.M.; McFrederick, Q.S. Hunting for healthy microbiomes: Determining the core microbiomes of *Ceratina*, *Megalopta*, and *Apis* bees and how they associate with microbes in bee collected pollen. *Conserv. Genet.* **2017**, *18*, 701–711. [\[CrossRef\]](#)
- Lozo, J.; Berić, T.; Terzić-Vidojević, A.; Stanković, S.; Fira, D.; Stanisavljević, L. Microbiota associated with pollen, bee bread, larvae and adults of solitary bee *Osmia cornuta* (Hymenoptera: Megachilidae). *Bull. Entomol. Res.* **2015**, *105*, 470–476. [\[CrossRef\]](#)
- Martinson, V.G.; Danforth, B.N.; Minckley, R.L.; Rueppell, O.; Tingek, S.; Moran, N.A. A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol. Ecol.* **2011**, *20*, 619–628. [\[CrossRef\]](#)
- McFrederick, Q.S.; Thomas, J.M.; Neff, J.L.; Vuong, H.Q.; Russell, K.A.; Hale, A.R.; Mueller, U.G. Flowers and wild megachilid bees share microbes. *Microb. Ecol.* **2017**, *73*, 188–200. [\[CrossRef\]](#)
- McFrederick, Q.S.; Wcislo, W.T.; Taylor, D.R.; Ishak, H.D.; Dowd, S.E.; Mueller, U.G. Environment or kin: Whence do bees obtain acidophilic bacteria? *Mol. Ecol.* **2012**, *21*, 1754–1768. [\[CrossRef\]](#)
- McFrederick, Q.S.; Rehan, S.M. Wild bee pollen usage and microbial communities co-vary across landscapes. *Microb. Ecol.* **2019**, *77*, 513–522. [\[CrossRef\]](#)

16. Koch, H.; Abrol, D.P.; Li, J.; Schmid-Hempel, P. Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Mol. Ecol.* **2013**, *22*, 2028–2044. [\[CrossRef\]](#)
17. Martinson, V.G.; Moy, J.; Moran, N.A. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl. Environ. Microbiol.* **2012**, *78*, 2830–2840. [\[CrossRef\]](#)
18. Parmentier, A.; Meeus, I.; Van Nieuwerburgh, F.; Deforce, D.; Vandamme, P.; Smagghe, G. A different gut microbial community between larvae and adults of a wild bumblebee nest (*Bombus pascuorum*). *Insect Sci.* **2018**, *25*, 66–74. [\[CrossRef\]](#)
19. Michener, C.D. *The Social Behavior of the Bees*; Belknap Press of Harvard University Press: Cambridge, MA, USA, 1974.
20. McFrederick, Q.S.; Rehan, S.M. Characterization of pollen and bacterial community composition in brood provisions of a small carpenter bee. *Mol. Ecol.* **2016**, *25*, 2302–2311. [\[CrossRef\]](#)
21. McFrederick, Q.S.; Wcislo, W.T.; Hout, M.C.; Mueller, U.G. Host species and developmental stage, but not host social structure, affects bacterial community structure in socially polymorphic bees. *FEMS Microbiol. Ecol.* **2014**, *88*, 398–406. [\[CrossRef\]](#)
22. Rothman, J.A.; Andrikopoulos, C.; Cox-Foster, D.Q.S. McFrederick. Floral and foliar source affect the bee nest microbial community. *Microb. Ecol.* **2019**, *78*, 506–516. [\[CrossRef\]](#)
23. Durrer, S.; Schmid-hempel, P. Shared use of flowers leads to horizontal pathogen transmission. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **1994**, *258*, 299–302. [\[CrossRef\]](#)
24. Evison, S.E.F.; Roberts, K.E.; Laurenson, L.; Pietravalle, S.; Hui, J.; Biesmeijer, J.C.; Hughes, W.O.H. Pervasiveness of parasites in pollinators. *PLoS ONE* **2012**, *7*, e30641. [\[CrossRef\]](#)
25. Singh, R.; Levitt, A.L.; Rajotte, E.G.; Holmes, E.C.; Ostiguy, N.; VanEngelsdorp, D.; Cox-Foster, D.L. RNA viruses in hymenopteran pollinators: Evidence of inter-taxa virus transmission via pollen and potential impact on non-*Apis* hymenopteran species. *PLoS ONE* **2010**, *5*, e14357. [\[CrossRef\]](#)
26. McArt, S.H.; Koch, H.; Irwin, R.E.; Adler, L.S. Arranging the bouquet of disease: Floral traits and the transmission of plant and animal pathogens. *Ecol. Lett.* **2014**, *17*, 624–636. [\[CrossRef\]](#)
27. Richardson, R.T.; Lin, C.-H.; Quijia, J.O.; Riusech, N.S.; Goodell, K.; Johnson, R.M. Rank-based characterization of pollen assemblages collected by honey bees using a multi-locus metabarcoding approach. *Appl. Plant Sci.* **2015**, *3*, 1500043. [\[CrossRef\]](#)
28. Gardiner, M.A.; Tuell, J.K.; Isaacs, R.; Gibbs, J.; Ascher, J.S.; Landis, D.A. Implications of three biofuel crops for beneficial arthropods in agricultural landscapes. *Bioenergy Res.* **2010**, *3*, 6–19. [\[CrossRef\]](#)
29. Tucker, E.M.; Rehan, S.M. Farming for bees: Annual variation in pollinator populations across agricultural landscapes. *Agric. For. Entomol.* **2018**, *20*, 541–548. [\[CrossRef\]](#)
30. Tuell, J.K.; Ascher, J.S.; Isaacs, R. Wild bees (Hymenoptera: Apoidea: Anthophila) of the Michigan highbush blueberry agroecosystem. *Ann. Entomol. Soc. Am.* **2009**, *102*, 275–287. [\[CrossRef\]](#)
31. Rehan, S.M.; Richards, M.H. Nesting biology and subsociality in *Ceratina calcarata* (Hymenoptera: Apidae). *Can. Entomol.* **2010**, *142*, 65–74. [\[CrossRef\]](#)
32. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Knight, R. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [\[CrossRef\]](#)
33. Callahan, B.J.; McMurdie, P.J.; Rosen, M.J.; Han, A.W.; Johnson, A.J.A.; Holmes, S.P. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, *13*, 581–583. [\[CrossRef\]](#)
34. McDonald, D.; Price, M.N.; Goodrich, J.; Nawrocki, E.P.; DeSantis, T.Z.; Probst, A.; Hugenholtz, P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **2012**, *6*, 610–618. [\[CrossRef\]](#)
35. Salter, S.J.; Cox, M.J.; Turek, E.M.; Calus, S.T.; Cookson, W.O.; Moffatt, M.F.; Walker, A.W. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* **2014**, *12*, 1–12. [\[CrossRef\]](#)
36. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [\[CrossRef\]](#)
37. Friedman, J.; Alm, E.J. Inferring correlation networks from genomic survey data. *PLoS Comput. Biol.* **2012**, *8*, 1–11. [\[CrossRef\]](#)
38. Faust, K.; Sathirapongsasuti, J.F.; Izard, J.; Segata, N.; Gevers, D.; Raes, J.; Huttenhower, C. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput. Biol.* **2012**, *8*, e1002606. [\[CrossRef\]](#)

39. Layeghifard, M.; Hwang, D.M.; Guttman, D.S. Disentangling interactions in the microbiome: A network Perspective. *Trends Microbiol.* **2017**, *25*, 217–228. [\[CrossRef\]](#)
40. Dalet, C.; Maudlin, I. *Sodalis gen. nov.* and *Sodalis glossinidius sp. nov.* a microaerophilic secondary endosymbiont of the tsetse fly *Glossina morsitans morsitans*. *Int. J. Syst. Bacteriol.* **1999**, *49*, 267–275. [\[CrossRef\]](#)
41. Saeed, A.; White, J.A. Surveys for maternally-inherited endosymbionts reveal novel and variable infections within solitary bee species. *J. Invertebr. Pathol.* **2015**, *132*, 111–114. [\[CrossRef\]](#)
42. Keller, A.; Danner, N.; Grimmer, G.; Ankenbrand, M.; von der Ohe, K.; von der Ohe, W.; Steffan-Dewenter, I. Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biol.* **2015**, *17*, 558–566. [\[CrossRef\]](#)
43. Kraaijeveld, K.; de Weger, L.A.; Ventayol García, M.; Buermans, H.; Frank, J.; Hiemstra, P.S.; den Dunnen, J.T. Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol. Ecol. Resour.* **2015**, *15*, 8–16. [\[CrossRef\]](#)
44. Richardson, R.T.; Curtis, H.R.; Matcham, E.G.; Hua Lin, C.; Suresh, S.; Sponsler, D.B.; Johnson, R.M. Quantitative multi-locus metabarcoding and waggle dance interpretation reveal honey bee spring foraging patterns in Midwest agroecosystems. *Mol. Ecol.* **2019**, *28*, 686–697. [\[CrossRef\]](#)
45. Bell, K.L.; Burgess, K.S.; Botsch, J.C.; Dobbs, E.K.; Read, T.D.; Brosi, B.J. Quantitative and qualitative assessment of pollen DNA metabarcoding using constructed species mixtures. *Mol. Ecol.* **2018**, *28*, 431–455. [\[CrossRef\]](#)
46. Roulston, T.H.I.; Cane, J.H. Pollen nutritional content and digestibility for animals. *Plant Syst.* **2000**, *222*, 187–209. [\[CrossRef\]](#)
47. Eckhardt, M.; Haider, M.; Dorn, S.; Müller, A. Pollen mixing in pollen generalist solitary bees: A possible strategy to complement or mitigate unfavourable pollen properties? *J. Anim. Ecol.* **2014**, *83*, 588–597. [\[CrossRef\]](#)
48. Vanderplanck, M.; Decleves, S.; Roger, N.; Decroo, C.; Caulier, G.; Glauser, G.; Michez, D. Is non-host pollen suitable for generalist bumblebees? *Insect Sci.* **2018**, *25*, 259–272. [\[CrossRef\]](#)
49. Weiss, S.; Van Treuren, W.; Lozupone, C.; Faust, K.; Friedman, J.; Deng, Y.; Knight, R. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J.* **2016**, *10*, 1669–1681. [\[CrossRef\]](#)
50. Aydogan, E.L.; Moser, G.; Müller, C.; Kämpfer, P.; Glaeser, S.P. Long-term warming shifts the composition of bacterial communities in the phyllosphere of *Galium album* in a permanent grassland field-experiment. *Front. Microbiol.* **2018**, *9*, 1–17. [\[CrossRef\]](#)
51. Jia, S.; Zhang, X.; Zhang, G.; Yin, A.; Zhang, S.; Li, F.; Yu, J. Seasonally variable intestinal metagenomes of the red palm weevil (*Rhynchophorus ferrugineus*). *Environ. Microbiol.* **2013**, *15*, 3020–3029. [\[CrossRef\]](#)
52. Toju, H.; Fukatsu, T. Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: Relevance of local climate and host plants. *Mol. Ecol.* **2011**, *20*, 853–868. [\[CrossRef\]](#) [\[PubMed\]](#)
53. Zhu, Y.-X.; Song, Y.-L.; Zhang, Y.-K.; Hoffmann, A.A.; Zhou, J.-C.; Sun, J.-T.; Hong, X.-Y. Incidence of facultative bacterial endosymbionts in spider mites associated with local environments and host plants. *Appl. Environ. Microbiol.* **2018**, *84*, e02546-17. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Aizenberg-Gershtein, Y.; Izhaki, I.; Halpern, M. Do honeybees shape the bacterial community composition in floral nectar? *PLoS ONE* **2013**, *8*, e67556. [\[CrossRef\]](#)
55. Samuni-Blank, M.; Izhaki, I.; Laviad, S.; Bar-Massada, A.; Gerchman, Y.; Halpern, M. The role of abiotic environmental conditions and herbivory in shaping bacterial community composition in floral nectar. *PLoS ONE* **2014**, *9*, e99107. [\[CrossRef\]](#)
56. Vannette, R.L.; Fukami, T. Dispersal enhances beta diversity in nectar microbes. *Ecol. Lett.* **2017**, *20*, 901–910. [\[CrossRef\]](#)
57. Vannette, R.L.; Bichier, P.; Philpott, S.M. The presence of aggressive ants is associated with fewer insect visits to and altered microbe communities in coffee flowers. *Basic Appl. Ecol.* **2017**, *20*, 62–74. [\[CrossRef\]](#)

