A Meta-Analysis of the 16S-rRNA Gut Microbiome Data in Honeybees (*Apis Mellifera*)

Authors: Alexis G. Gkantiragas¹, Jacopo Gabrielli¹

¹ = University College London (UCL). Department of Structural and Molecular Biology

1. Abstract

Honeybees (*Apis Mellifera*) perform an essential role in the ecosystem and economy through pollination of insect-pollinated plants, but their population is declining. Many causes of honeybees' decline are likely to be influenced by the microbiome which is thought to play an important role in bees and is particularly susceptible to infection and pesticides. However, there has been no systematic review or meta-analysis on honeybee microbiome data. Therefore, we conducted the first systematic meta-analysis of 16S-rRNA data to address this gap in the literature. Four studies were in a usable format – accounting for 336 honeybee's worth of data – the largest such dataset to the best of our knowledge. We analysed these datasets in QIIME2 and visualised the results in R-studio. For the first time, we conducted a multi-study evaluation of the core and rare bee microbiome and confirmed previous compositional microbiome data. We established that *Snodgrassella*, *Lactobacillus*, *Bifidobacterium*, *Fructobacillus* and *Saccaribacter* form part of the core microbiome and identify 251 rare bacterial genera. Additional components of the core microbiome were likely obscured by incomplete classification. Future studies should refine and add to our existing dataset to produce a more conclusive and high-resolution portrait of the honeybee microbiome. Furthermore, we emphasise the need for an actively curated dataset and enforcement of data sharing standards.

Figure 1 - Graphical abstract. Made by the author in Biorender.com.

2. Introduction

Honeybees (*Apis Mellifera*) play an essential role in the global ecosystem and economy, pollinating most flowering plants and numerous food crops, such as apples and cucumbers¹. Recent evidence also implies that insect pollinators (honeybees included) perform a role in the pollination of wind-blown pollinated crops, including rice and corn². Thus, the plight of honeybee populations in recent decades³ bears grim practical implications.

Aside from loss of flora, the decline in honeybee populations has been attributed to numerous factors — notably, pesticide use and pathogens⁴. These will, to some degree, interact with the microbiome. For instance, the pesticide chlorothalonil was found to lower commensal bacterial sugar metabolism⁵. Similarly, bees live in densely populated colonies with $\geq 50,000$ individuals⁶ and must contest with greater risk from infectious agents. For example, lethal foulbrood infections are caused by the bacterium *Paenibacillus larvae*⁷. In other organisms, symbiotic microbes protect against infection through a combination of outcompeting and direct killing⁸.

A reason honeybees are appealing for microbiome studies is that workers and queens are genetically identical yet display profound morphological and behavioural differences — implying that these are environmentally determined. Individuals become queens when they are fed royal jelly in the early stages of larval development⁹. While the mechanism whereby these developmental changes manifest is still largely unknown, evidence suggests that bacterial colonisation from workers' glands contributes to the queen-rearing process¹⁰. Equally, it might also be that the different nutrient composition of royal jelly nurtures different bacterial populations.

While microbiome studies often struggle to demonstrate causative relevance, there is evidence that changes in the microbiome reflect or even dictate phenotypic and behavioural differences in honeybees¹¹. Nonetheless, it is worth noting that the hitherto assumed unbiased nature of metagenomics studies have recently been questioned due to substantial bias introduced in DNA extraction and PCR steps¹².

Bees are generally thought to have a relatively conserved resident population of 8 or so bacterial species¹³. However, these studies principally focused on worker bees, and the microbiome of larvae, drones and queen bees is still largely uncharacterised as it is thought to be more dependent on environmental variation^{10,14}. There has not, to our knowledge, been any large-scale validation of these findings to date. The current evidence on the bee microbiome's role is summarised in an excellent review by Engel *et al.*¹⁵.

Moreover, most of the individual microbiome members that have been identified have not been linked to specific symbiotic functions within the honey bee gut¹⁶. Overall, there is general agreement that the bee microbiome is consistent and stable within workers in a colony which can be attributed both the social nature of bees, the large colony size (up to 50,000 bees per colony), and the use of propolis which works as an antiseptic and reduces bacterial diversity in the hive^{16,17}.

Despite the likely importance of the microbiome in bees, relatively few studies have been conducted studying the honeybee microbiome, and there currently exists no meta-analysis or systematic review collating the available bee microbiome data. While the Bee-Biome Consortium has received funding to construct a data portal, this has yet to be completed

<https://wp.unil.ch/beebiome/>. Consequently, this meta-analysis sought to aggregate available datasets on the honeybee microbiome and conduct, for the first time, a multi-study characterisation of the honeybee microbiome. We hope that this aggregated dataset will prove helpful for those conducting future studies.

Although the ‘core’ bee microbiome is referred to in the literature¹⁸, and can usually be cultured in the lab¹³, this has primarily been summarised in non-systematic reviews and studies. For example, one paper claiming to establish a core bee microbiome contained only 41 honeybees¹⁹. This leaves such results open to unconscious bias and risks excluding or overrepresenting certain groups of bacteria – particularly as there is no fixed definition of ‘core microbiome’²⁰.

We focused upon amplicon sequencing data – a PCR derived approach to perform metagenomic sequencing in which a small sequence or ‘amplicon’ is amplified via PCR and sequenced. Such approaches are essential when studying mixed microbiological samples since most microbes cannot be cultured in laboratory settings²¹. Of equal relevance, amplicon sequencing permits vastly higher throughput experiments.

While more sophisticated methods such as whole-genome next-generation sequencing (NGS) have become more widespread in microbiome studies, including in bees²², the examination of such methods was beyond this study's scope. Though the cost is ever decreasing, these methods are more expensive and demand a more complex workflow. Nonetheless, future studies may wish to consider incorporating such data into a meta-analysis.

3. Methods

3.1. Systematic Search

Firstly, we performed a systematic review of the literature on PubMed using the search terms: (*microbiome AND "16S rRNA" AND "Apis mellifera"*)

This search was performed on 22/10/2020 yielding 26 results. Next, searches for data availability were performed independently by two researchers. A total of 14 studies were excluded on the basis of lack of available data or data available in an unusable form (e.g. in PDFs). This left a total of 12 studies to be used in the meta-analysis. The annotated paper database can be found in our Github repository (see data availability statement).

Of the 12 studies where data was accessible, most were either multiplexed and with no barcode file to allow demultiplexing, had imprecise structure and/or metadata or had considerably low-quality sequences, making them hard to extract and analyse. Finally, only four studies had a clear data structure and were fully accessible and reproducible on QIIME2. Accordingly, these studies were used subsequently and are detailed below.

We did not intentionally include studies on bumblebees or other related species as, while valuable, they were beyond the scope of this study. Additionally, papers that did not have experimental data available publicly were excluded due to time constraints on this study. We also excluded studies investigating *nosema* infestation as we reasoned this would likely perturb normal microbiome function. Nonetheless, these would not be picked up since bee parasites such as *nosema* are eukaryotes, and *nosema* infestation may be asymptomatic²³. Studies investigating the specific effect of *nosema* infestation on microbiome function and corresponding meta-analysis would be valuable additions to the literature.

We focused upon 16S-rRNA data since this is a well-established bacterial marker used routinely in amplicon metagenomics sequencing. We used the GreenGene database for the identification of 16S-rRNA <https://greengenes.secondgenome.com/>. Pipeline Files in .sra format were downloaded from SRA, where they were converted to FASTA format, readable by QIIME2 and aggregated. Various data types were obtained; thus, adjustments and labelling had to be carried out.

3.1.1.A – Saelao *et al.*, 2020¹⁷

The authors studied 62 bees from 32 colonies. They used an equal number of bees from propolis-rich and propolis-poor colonies and concluded that the former was associated with a ‘stabilised’ microbiome. However, they did not link these to any colony outcomes and neglected to describe how propolis richness was determined.

3.1.2.B – Horton *et al.*, 2015²⁴

This paper claimed to define and confirm earlier experimental findings of the ‘core’ bee microbiome. Unfortunately, they did not clarify the number of samples used in their experiments – only 35 were present in the dataset.

3.1.3.C – Erban *et al.*, 2017²⁵

This study used 49 bees covering 27 colonies. Of note, some of the colonies in this sample were affected by foulbrood, likely accounting for this genera's presence within the rare bee microbiome, as examined later.

3.1.4.D – Powell *et al.*, 2016²⁶

This study contained 221 samples from an unspecified number of colonies. Unfortunately, we failed to exclude some bumblebee samples from this group, possibly accounting for the large number of outliers in D, as discussed later.

3.2. Diversity Metrics

Measures of alpha diversity were performed to ascertain each study's microbial diversity and the variation in diversity. Shannon diversity (Equation 1)(first described in 1948²⁷) was used as a non-phylogenetic measure of alpha diversity, while faiths-PD (Equation 2)(first described 1992²⁸) was employed as a phylogenetic measure of alpha diversity. Statistical significance was determined using Turkey's test²⁹.

$$H = - \sum_{i=1}^S p_i \ln(p_i)$$

Equation 1 - Shannon diversity index. Wherein:

H = Shannon diversity index

S = Number of species

P_i = Proportion of abundance represented by ith species

$$\text{Faiths PD} = \text{Branch length 1} + \text{Branch length 2} + \dots$$

Equation 2 - Faiths PD. Calculated via the sum of branch lengths from a phylogenetic tree.

Beta diversity was assessed in a PCoA plot via a Jaccard index (first developed 1912³⁰) as shown in Equation 3 and Bray-Curtis dissimilarity matrix (first developed 1957³¹), shown in Equation 4. PERMANOVA³² was used for comparisons of beta diversity.

$$J(A, B) = \frac{|A \cap B|}{|A \cup B|}$$

Equation 3 – Jaccard similarity matrix.

$$BC_d = \frac{\sum |x_i - x_j|}{\sum |x_i + x_j|}$$

Equation 4 - Bray-Curtis dissimilarity matrix. Wherein:

BC_d = Bray-curtis dissimilarity value

x_i = abundance of species in sample i

x_j = abundance of species in sample j

3.3. Data Visualisation and Analysis

Analysis was performed in QIIME2, and visualisation of data was performed in RStudio due to its superior processing times compared to similar pipelines³³. A QIIME2 pipeline (Figure 2) was utilised as described previously³⁴. We utilised both DADA2³⁵ and the QIIME2 deblur function for the denoising step due to difficulties in aggregating multiple datasets. We used 97% OTU clustering.

Figure 2 – Overall microbiome study- and QIIME2 pipeline. We conducted only the microbiome pipeline's final step – that is, analysis due to lack of access to a laboratory space. Made by the author in Biorender.com.

We used several libraries to handle and visualise the data. These are described in our Github repository https://github.com/JacopoGabrielli/Bee_Microbiome?fbclid=IwAR0AvpfowNnWHNDIFXM5z24a6ws9f5zKmF9o1n8DRXWH9SLqJ-YPHQo4qVg.

4. Results

4.1. Diversity Measures Imply a Consistent Microbiome

We began by measuring alpha diversities of each study to confirm previous findings that the honeybee microbiome is relatively consistent. Samples were rarefied to 1300 reads to provide sufficient coverage.

All studies except A and B had statistically significant differences in non-phylogenetic alpha diversity as measured by Shannon diversity (Figure 3A). Even A/B have such large error bars that the result might be obscured by the sizeable signal-to-noise ratio. This suggests that the microbial diversity varied across studies. Likewise, there appears to be a great deal of intra-study variation between bees, with data points in group A showing almost 10-fold differences in diversity. This implies that individual bees possess vastly dissimilar microbial diversity. Overall, it is worth noting that these are somewhat low alpha diversity values.

Figure 3 - Differences in diversity between bee microbiome studies. A) Shannon diversity and comparison between studies (shown below). B) Faiths-PD values for each study and comparisons between them. Error bars show extreme upper and lower bounds while boxes represent upper and lower quartiles with the centre line's median. Each data point shows a single

sample (Bee). 95% family wide confidence levels were calculated using Turkey's test. For A-B the box plot shows study letter/code on the x-axis and the diversity index on the y-axis.

Once adjusting for phylogeneticity using Faiths-PD, this picture changes (Figure 3B). Most noticeably, the intra-study variation appears significantly diminished, with comparatively tiny error bars for each study. However, this might be a result of the altered scale. While D contains many outliers at approximately 10-fold the mean, this is presumably a product of the large sample size D used. It is interesting that these outliers only become apparent once phylogeneticity is accounted for. These results suggest that their variation in the bee microbiome primarily occur between related microbial species.

Limited information can be divined from a PCoA plot of the studies. While C and D showed overlapped, all other studies showed no similarity to each other. This applies both when using a Jaccard- (Figure 4A) and Bray-Curtis similarity matrix (Figure 4B).

Figure 4 - PCoA plot using A) Jaccard similarity matrix. B) Bray-Curtis dissimilarity matrix. For A) and B), the x-axis represents PC1 which accounts for the greatest variance in the data while the y-axis represents PC2 which accounts for the second greatest variance in the data. Each data point represents an individual sample within its respective study.

Overall, we can see that the bacterial bee microbiome appears to be dominated by commensal, gram-positive species whose supplementation promotes health. Since honeybees exist in huge colonies, it is unsurprising that they would have conserved commensal microbiota, which protects against the higher risk of infection. It also may be a way to help ensure stability and consistency across individuals. While the main genera are all shared across studies, the incomplete classification of many taxonomic groups makes the results challenging to interpret.

The presence of *Bombiscardovia* in C is likely a result of contamination with bumblebees since this genus is exclusively found in bumblebees⁴⁸. Indeed, after revisiting the dataset, it appears that we failed to exclude bumblebee samples. It is also worth noting that many of D's species (which were also the largest group) are either unidentified or unassigned.

Other studies reported *Bartonella*⁴⁹ and *Gillamella* genera which we did not observe. However, this is doubtless due to poor sequencing information as *Bartonella* can likely be accounted for by the “Rhizobiales...NA...NA” (for it falls within this order). Interestingly, in the case of *Gillamella* – past studies utilising QIIME had noted that it mislabelled *Gillamella* and *Frischella* as “Pasteurellales”⁵⁰. This is somewhat unusual given that *Gillamella* does not belong to the order Pasteurellales.

4.3. The Core and Rare Microbiome

We attempted to characterise the core (Figure 6A) and rare (Figure 6B) honeybee microbiome. Most obvious is that ‘unknown’ comprise the most common group in the data. This implies that the bee microbiome is poorly characterised. The difference between Figure 6 and Figure 5 can largely be accounted for by the large number of unknowns (labelled NA in the previous graph) and the data's heterogeneity. Furthermore, group D is far more heavily weighted here owing to the larger sample size. Nevertheless, the same genera are represented here, such as *Snodgrassella*, *Lactobacillus* and *Bifidobacterium*.

Figure 6 - Heatmaps of the honeybee microbiome. This represents a weighted and aggregated form of the four studies. A) The core bee microbiome. These were sorted in order of prevalence. B) The core microbiome at order level. The y-axis indicates the genus (A) and order (B) while the x-axis denotes the detection threshold in A-B.

Interestingly, *Fructobacilla* and *Saccharibacter* are both recently discovered gram-negative genera⁵¹. *Fructobacilla* are, as the name suggests, a form of lactic acid bacteria which thrive in fructose-rich environments and have been well described in bee guts⁵². *Saccharibacter* are a genus with presently only one species (*Saccharibacter floricol*) which has been described in honeybees¹⁹ and was first isolated from flower pollen⁵³ – from which one might speculate the guts were colonised. They both play a role in sugar metabolism – suggesting a probable commensal function as microbes often aid in digestion⁵⁴.

In considering the core bee microbiome orders (Figure 6B), we see that various unusual bacteria are present, including *Rhizobium*, most well known as a nitrogen-fixing bacteria found in the root nodules of legume plants⁵⁵. These are likely *Gilliamella* which belong to the order *Rhizobium*.

A total of 251 rare genera were identified. These were filtered by setting a maximum prevalence of 3%. Interestingly, this included 8 genera of archaea which were primarily involved in methane processing. This may have been due to improper storage or preservation of the bee samples, which may have led to decay in one or two instances. This occurs rapidly for a small organism like a honeybee. Nonetheless, archaea have been reported (albeit at very low abundances) in Africanized honeybees⁵⁶.

Finally, two *Paenibacillus* genera were detected in the rare microbiome, a known honeybee pathogen (American foulbrood)⁷. This is likely the result of foulbrood samples within C – though recent studies reporting asymptomatic foulbrood⁵⁷, *nosema*²³, *crithidia*⁵⁸ and chronic paralysis virus⁵⁸ presence do raise the question of whether *Paenibacillus* might have been present in other studies. A more complete bee microbiome dataset might clarify this issue – particularly given that this would have significant impacts for both wild and captive bee management.

Differential abundances for classes of bacteria were conducted to evaluate these groups' abundance between studies. Interestingly, study A showed the greatest variation in samples (bees) in abundance (Figure 7). Proteobacteria in study A had outliers approximately 4-fold higher than those of any other groups. These are gram-negative bacteria which include *Snodgrassella* and *Gilliamella*. We see that there is a greater median abundance in A and B. Additionally, A and B also show far greater variation. For visualisation purposes, a single outlier in B was excluded.

Figure 7 - Differential abundance for the classes *proteobacteria*, *actinobacteria* and *firmicutes*. Error bars show extreme upper and lower bounds; boxes show upper and lower quartiles with the median represented by the centre line. Individual data points represent individual bees. The y-axis denotes abundance (though scales vary) and the x-axis shows the study letter/code.

Again, *Actinobacteria* were more abundant in study A (Figure 7). These are a gram-positive phylum of bacteria frequently found in bee guts and have even been found to produce compounds that inhibited *Paenibacillus larvae*⁵⁹. Although this study was conducted in *Melipona scutellaris* (Stingless bees), the results are likely translatable to honeybees. These correspond to orders such as *Streptomyces* noted in the core bee microbiome (Figure 5B). Oddly, these are generally thought to have higher abundance than we obtained⁶⁰.

Firmicutes' differential abundances appear to be relatively homogenous across all samples except for D, which has a very low median (Figure 7). This is quite strange and might be due to impurities or sequencing issues since this group had a large amount of NA data. There was also greater variation in abundance between samples in A with several outliers showing around 10-fold the abundance versus the median. These correspond to genera such as *Lactobacillus*.

5. Discussion

5.1. Agreement with Prior Studies

Overall, our results confirmed previous studies' findings, confirming the 'core' bee microbiome in the first meta-analysis in this area. However, we also highlighted the need for further classification of the bee microbiome and failed to replicate some of previous studies' findings, such as the prevalence of *Gillamella*.

5.2. Limitations

Our investigations had several constraints. Unfortunately, we were unable to access data from most available studies. Similarly, data that required laborious mining techniques could not be used due to time constraints. Searches on Web of Science and other databases could have been searched but were not due to material and time constraints on the study. The authors will conduct these at a later date.

Due to investigating only 16S-rRNA, we were unable to study viruses, prions or fungi – which nonetheless play a role in bee health^{61,62}. However, these should be studied separately since these would require sequencing different regions and consequently, datasets which contained all types of microbes would likely be too small. Similarly, we did not look into specific subset populations like the bees' surface and the specific composition of each body part e.g. malpighian tubules vs hindgut and honey crop (Figure 8). This is because many studies use whole bee samples – thereby prohibiting organ-based differentiation. While subsequent studies may wish to take such a differential approach, this would require skilled and time-consuming dissections.

Figure 8 - Honeybee worker digestive anatomy. Shown enlarged are the major components of the bee digestive system, which would be well suited to closer study. Made by the author in Biorender.com.

Many different measures of microbial diversity could have been used. In particular, studying phylogenetic beta diversity, such as via UniFrac analysis⁶³, would be a worthwhile addition since we only used non-phylogenetic measures. This might increase clustering between groups, given that variation in alpha diversity appeared reduced upon using Faith's-PD.

Moreover, due to the nature of meta-analyses, we have no way of knowing whether contamination in the samples may have occurred. We must simply hope that these do not directionally bias the results of our study.

Finally, since this is an active research area, several studies (such as Ptaszyńska *et al.*, 2021⁶⁴) were added throughout our analysis that would have fulfilled inclusion criteria. This underlines the importance of a curated and active database as proposed by Engel *et al.* We intend to include all studies published since the initial search in a later iteration of this study.

5.3. Future Directions

Subsequent studies should consider contacting authors of papers with registered but uninterpretable datasets, checking for papers published since the initial literature search and finally conducting similar meta-analyses for different bee types (e.g. bumblebees). This approach should be easily translatable to similar insect types. In the future, we aim to improve our dataset by adding recently published papers and constructing a web app to make the data more accessible.

Given the ever-increasing affordability of genetic sequencing, a well-characterised bee microbiome may have eventual commercial relevance in smart bee monitoring. This could provide an additional dimension to be considered alongside hive temperature, humidity and honey production⁶⁵.

Similarly, while various isolated incidents of asymptomatic infections (described earlier) have been reported, such findings can only truly be evaluated by large scale meta-analysis to ensure that the results were not simply anomalous. If chronic infections are prevalent in bees, this could shift emphasis of bee management towards general maintenance of hive health as opposed to control of infection sources. Moreover, if, as one study of over 1000 bees suggested⁵⁸ chronic asymptomatic infections are quite literally ubiquitous in bees, it would be more proper to consider them a resident component of the bee microbiome and regard them as opportunistic pathogens. Thus, future studies should avoid excluding infected bees.

While incredibly ambitious, one might consider constructing a time-course-based meta-analysis, matching bees by age, thus assembling a picture of the bee microbiome over the life course. Although huge datasets would be required for such subgrouping, it might provide enhanced resolution of the bee microbiome and provide novel insights into bee physiology.

Finally, a potentially exciting area for study would be creating a separate 'wild' honeybee core dataset. Wild bees have different foraging habits⁶⁶ and have overall smaller, less densely packed colonies — particularly as captive bees are kept in apiaries with many hives close together.

6. Conclusions

While significant evidence supports the importance of the microbiome in bees, its true relevance cannot be determined until the standard bee microbiome has been elucidated. Once established, deviations from this core microbiome, and indeed all insights in bees' microbiomes will gain greater relevance. Unfortunately, there is presently no standard core or rare microbiome in the honeybee. We have established, for the first time, a core honeybee microbiome though further characterisation is needed. To facilitate future research in the field, we have made this publicly accessible to the community as a reference. Our meta-analysis demonstrates that there is a need for standardisation and enforcement of data sharing policies. Specialist journals in the field would be well advised to require accessible data sharing.

7. Data Availability and Conflict of Interest Statement

All data is freely and publicly available on Github (https://github.com/JacopoGabrielli/Bee_Microbiome?fbclid=IwAR0AvpfowNnWHNDIFXM5z24a6ws9f5zKmF9o1n8DRXWH9SLqJ-YPHQo4qVg). Both authors read and approved the final manuscript.

Alexis Gkantiragas is a voting member of the BeeXML standardisation Apimondia working group, though does not receive financial compensation for this involvement.

8. Acknowledgments

We would like to thank Dr Toryn Poolman his assistance with this study. Furthermore, we would like to thank the colleagues who helped review this manuscript.

9. References

1. Spivak, M., Mader, E., Vaughan, M. & Euliss, N. H. The plight of the bees. *Environ. Sci. Technol.* (2011) doi:10.1021/es101468w.
2. Saunders, M. E. Insect pollinators collect pollen from wind-pollinated plants: implications for pollination ecology and sustainable agriculture. *Insect Conservation and Diversity* (2018) doi:10.1111/icad.12243.
3. Potts, S. G. *et al.* Declines of managed honey bees and beekeepers in Europe. *J. Apic. Res.* (2010) doi:10.3896/IBRA.1.49.1.02.
4. Goulson, D., Nicholls, E., Botías, C. & Rotheray, E. L. Bee declines driven by combined Stress from parasites, pesticides, and lack of flowers. *Science* (2015) doi:10.1126/science.1255957.
5. Kakumanu, M. L., Reeves, A. M., Anderson, T. D., Rodrigues, R. R. & Williams, M. A. Honey bee gut microbiome is altered by in-hive pesticide exposures. *Front. Microbiol.* (2016) doi:10.3389/fmicb.2016.01255.
6. Cutler, G. C., Scott-Dupree, C. D. & Drexler, D. M. Honey bees, neonicotinoids and bee incident reports: The Canadian situation. *Pest Manag. Sci.* (2014) doi:10.1002/ps.3613.
7. Genersch, E. American Foulbrood in honeybees and its causative agent, *Paenibacillus larvae*. *J. Invertebr. Pathol.* (2010) doi:10.1016/j.jip.2009.06.015.
8. Abt, M. C. & Pamer, E. G. Commensal bacteria mediated defenses against pathogens. *Current Opinion in Immunology* (2014) doi:10.1016/j.coi.2014.03.003.
9. Townsend, G. F. & Lucas, C. C. The chemical nature of royal jelly¹. *Biochem. J.* (1940) doi:10.1042/bj0341155.
10. Tarpy, D. R., Mattila, H. R. & Newton, I. L. G. Development of the honey bee gut microbiome throughout the queen-rearing process. *Appl. Environ. Microbiol.* (2015) doi:10.1128/AEM.00307-15.
11. Vernier, C. L. *et al.* The gut microbiome defines social group membership in honey bee colonies. *Sci. Adv.* (2020) doi:10.1126/sciadv.abd3431.
12. Brooks, J. P. *et al.* The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol.* (2015).
13. Moran, N. A. Genomics of the honey bee microbiome. *Current Opinion in Insect Science* (2015) doi:10.1016/j.cois.2015.04.003.
14. Kapheim, K. M. *et al.* Caste-Specific Differences in Hindgut Microbial Communities of Honey Bees (*Apis mellifera*). *PLoS One* **10**, e0123911 (2015).
15. Engel, P. *et al.* The bee microbiome: Impact on bee health and model for evolution and ecology of host-microbe interactions. *MBio* (2016) doi:10.1128/mBio.02164-15.
16. Kwong, W. K. & Moran, N. A. Gut microbial communities of social bees. *Nat. Rev. Microbiol.* **14**, 374–384 (2016).
17. Saelao, P., Borba, R. S., Ricigliano, V., Spivak, M. & Simone-Finstrom, M. Honeybee microbiome is stabilized in the presence of propolis. *Biol. Lett.* (2020) doi:10.1098/rsbl.2020.0003.
18. Raymann, K., Bobay, L. M. & Moran, N. A. Antibiotics reduce genetic diversity of core species in the honeybee gut microbiome. *Mol. Ecol.* (2018) doi:10.1111/mec.14434.
19. 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) doi:10.1007/s10592-017-0937-7.
20. Risely, A. Applying the core microbiome to understand host–microbe systems.

- Journal of Animal Ecology* (2020) doi:10.1111/1365-2656.13229.
21. Streit, W. R., Daniel, R. & Jaeger, K. E. Prospecting for biocatalysts and drugs in the genomes of non-cultured microorganisms. *Current Opinion in Biotechnology* (2004) doi:10.1016/j.copbio.2004.05.006.
22. Haddad, N. *et al.* Next generation sequencing of *Apis mellifera syriaca* identifies genes for Varroa resistance and beneficial bee keeping traits. *Insect Sci.* (2016) doi:10.1111/1744-7917.12205.
23. Fernández, J. M. *et al.* Asymptomatic presence of *Nosema* spp. in Spanish commercial apiaries. *J. Invertebr. Pathol.* (2012) doi:10.1016/j.jip.2012.06.008.
24. Horton, M. A., Oliver, R. & Newton, I. L. No apparent correlation between honey bee forager gut microbiota and honey production. *PeerJ* (2015) doi:10.7717/peerj.1329.
25. Erban, T. *et al.* Bacterial community associated with worker honeybees (*Apis mellifera*) affected by European foulbrood. *PeerJ* (2017) doi:10.7717/peerj.3816.
26. Powell, E., Ratnayake, N. & Moran, N. A. Strain diversity and host specificity in a specialized gut symbiont of honeybees and bumblebees. *Mol. Ecol.* (2016) doi:10.1111/mec.13787.
27. Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* (1948) doi:10.1002/j.1538-7305.1948.tb01338.x.
28. Faith, D. P. Conservation evaluation and phylogenetic diversity. *Biol. Conserv.* (1992) doi:10.1016/0006-3207(92)91201-3.
29. Tukey, J. W. Comparing Individual Means in the Analysis of Variance. *Biometrics* (1949) doi:10.2307/3001913.
30. Jaccard, P. THE DISTRIBUTION OF THE FLORA IN THE ALPINE ZONE. *New Phytol.* (1912) doi:10.1111/j.1469-8137.1912.tb05611.x.
31. Bray, J. R. & Curtis, J. T. An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecol. Monogr.* (1957) doi:10.2307/1942268.
32. Anderson, M. J. Permutational Multivariate Analysis of Variance (PERMANOVA). in *Wiley StatsRef: Statistics Reference Online* (2017). doi:10.1002/9781118445112.stat07841.
33. Marizzoni, M. *et al.* Comparison of Bioinformatics Pipelines and Operating Systems for the Analyses of 16S rRNA Gene Amplicon Sequences in Human Fecal Samples. *Front. Microbiol.* (2020) doi:10.3389/fmicb.2020.01262.
34. Estaki, M. *et al.* QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. *Curr. Protoc. Bioinforma.* (2020) doi:10.1002/cpbi.100.
35. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* (2016) doi:10.1038/nmeth.3869.
36. Picard, C. *et al.* Review article: Bifidobacteria as probiotic agents - Physiological effects and clinical benefits. *Alimentary Pharmacology and Therapeutics* (2005) doi:10.1111/j.1365-2036.2005.02615.x.
37. Praet, J. *et al.* Bifidobacterium commune sp. nov. isolated from the bumble bee gut. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* (2015) doi:10.1007/s10482-015-0425-3.
38. Anderson, K. E. *et al.* Draft genome sequences of two Bifidobacterium sp. from the honey bee (*Apis mellifera*). *Gut Pathog.* (2013) doi:10.1186/1757-4749-5-42.
39. Alberoni, D. *et al.* Bifidobacterium xylocopae sp. nov. and Bifidobacterium aemilianum sp. nov., from the carpenter bee (*Xylocopa violacea*) digestive tract. *Syst. Appl. Microbiol.* (2019) doi:10.1016/j.syapm.2018.11.005.
40. Baffoni, L. *et al.* Effect of dietary supplementation of Bifidobacterium and Lactobacillus strains in *Apis mellifera* L. against *Nosema ceranae*. *Benef. Microbes*

- (2016) doi:10.3920/BM2015.0085.
41. Naito, E. *et al.* Beneficial effect of oral administration of *Lactobacillus casei* strain Shirota on insulin resistance in diet-induced obesity mice. *J. Appl. Microbiol.* (2011) doi:10.1111/j.1365-2672.2010.04922.x.
42. Daisley, B. A. *et al.* *Lactobacillus* spp. attenuate antibiotic-induced immune and microbiota dysregulation in honey bees. *Commun. Biol.* (2020) doi:10.1038/s42003-020-01259-8.
43. Audisio, M. C. & Benítez-Ahrendts, M. R. *Lactobacillus johnsonii* CRL1647, isolated from *Apis mellifera* L. bee-gut, exhibited a beneficial effect on honeybee colonies. *Benef. Microbes* (2011) doi:10.3920/BM2010.0024.
44. Kwong, W. K. & Moran, N. A. Cultivation and characterization of the gut symbionts of honey bees and bumble bees: Description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family Neisseriaceae of the betaproteobacteria, and *Gilliamella apicola* gen. nov., sp. nov., a memb. *Int. J. Syst. Evol. Microbiol.* (2013) doi:10.1099/ijs.0.044875-0.
45. Moller, K. & Kilian, M. V factor-dependent members of the family Pasteurellaceae in the porcine upper respiratory tract. *J. Clin. Microbiol.* (1990) doi:10.1128/jcm.28.12.2711-2716.1990.
46. Bisgaard, M. Ecology and significance of pasteurellaceae in animals. *Zentralblatt fur Bakteriologie* (1993) doi:10.1016/S0934-8840(11)80487-1.
47. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. U. S. A.* (2014) doi:10.1073/pnas.1405838111.
48. Killer, J. *et al.* *Bombiscardovia coagulans* gen. nov., sp. nov., a new member of the family Bifidobacteriaceae isolated from the digestive tract of bumblebees. *Syst. Appl. Microbiol.* (2010) doi:10.1016/j.syapm.2010.08.002.
49. Kešnerová, L., Moritz, R. & Engel, P. *Bartonella apis* sp. nov., a honey bee gut symbiont of the class Alphaproteobacteria. *Int. J. Syst. Evol. Microbiol.* (2016) doi:10.1099/ijs.0.000736.
50. Ludvigsen, J. *et al.* Shifts in the midgut/pyloric microbiota composition within a honey bee apiary throughout a season. *Microbes Environ.* (2015) doi:10.1264/jsme2.ME15019.
51. Endo, A. & Okada, S. Reclassification of the genus *Leuconostoc* and proposals of *Fructobacillus fructosus* gen. nov., comb. nov., *Fructobacillus durionis* comb. nov., *Fructobacillus ficulneus* comb. nov. and *Fructobacillus pseudoficulneus* comb. nov. *Int. J. Syst. Evol. Microbiol.* (2008) doi:10.1099/ijs.0.65609-0.
52. Khan, K. A. *et al.* Investigation of gut microbial communities associated with indigenous honey bee (*Apis mellifera jemenitica*) from two different eco-regions of Saudi Arabia. *Saudi J. Biol. Sci.* (2017) doi:10.1016/j.sjbs.2017.01.055.
53. Jojima, Y. *et al.* *Saccharibacter floricola* gen. nov., sp. nov., a novel osmophilic acetic acid bacterium isolated from pollen. *Int. J. Syst. Evol. Microbiol.* (2004) doi:10.1099/ijs.0.02911-0.
54. Krajmalnik-Brown, R., Ilhan, Z. E., Kang, D. W. & DiBaise, J. K. Effects of gut microbes on nutrient absorption and energy regulation. *Nutrition in Clinical Practice* (2012) doi:10.1177/0884533611436116.
55. Fisher, R. F. & Long, S. R. Rhizobium-plant signal exchange. *Nature* (1992) doi:10.1038/357655a0.
56. Saraiva, M. A. *et al.* Relationship between honeybee nutrition and their microbial communities. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* (2015) doi:10.1007/s10482-015-0384-8.

57. Khezri, M. *et al.* Prevalence of American foulbrood in asymptomatic apiaries of Kurdistan, Iran. *Vet. World* (2018) doi:10.14202/vetworld.2018.281-285.
58. D'Alvise, P., Seeburger, V., Gihring, K., Kieboom, M. & Hasselmann, M. Seasonal dynamics and co-occurrence patterns of honey bee pathogens revealed by high-throughput RT-qPCR analysis. *Ecol. Evol.* (2019) doi:10.1002/ece3.5544.
59. Rodríguez-Hernández, D. *et al.* Actinobacteria associated with stingless bees biosynthesize bioactive polyketides against bacterial pathogens. *New J. Chem.* (2019) doi:10.1039/c9nj01619h.
60. Grubbs, K. J. *et al.* Pollen *Streptomyces* Produce Antibiotic That Inhibits the Honey Bee Pathogen *Paenibacillus* larvae. *Front. Microbiol.* (2021) doi:10.3389/fmicb.2021.632637.
61. Ullah, A. *et al.* Viral impacts on honey bee populations: A review. *Saudi Journal of Biological Sciences* (2021) doi:10.1016/j.sjbs.2020.10.037.
62. González, G., Hinojo, M. J., Mateo, R., Medina, A. & Jiménez, M. Occurrence of mycotoxin producing fungi in bee pollen. *Int. J. Food Microbiol.* (2005) doi:10.1016/j.ijfoodmicro.2005.05.001.
63. Lozupone, C. & Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* (2005) doi:10.1128/AEM.71.12.8228-8235.2005.
64. Ptasińska, A. A. *et al.* Amplicon Sequencing of Variable 16S rRNA from Bacteria and ITS2 Regions from Fungi and Plants, Reveals Honeybee Susceptibility to Diseases Results from Their Forage Availability under Anthropogenic Landscapes. *Pathogens* (2021) doi:10.3390/pathogens10030381.
65. Cecchi, S., Spinsante, S., Terenzi, A. & Orcioni, S. A smart sensor-based measurement system for advanced bee hive monitoring. *Sensors (Switzerland)* (2020) doi:10.3390/s20092726.
66. Tuell, J. K., Fiedler, A. K., Landis, D. & Isaacs, R. Visitation by wild and managed bees (Hymenoptera: Apoidea) to eastern U.S. native plants for use in conservation programs. *Environ. Entomol.* (2008) doi:10.1603/0046-225X(2008)37[707:VBWAMB]2.0.CO;2.