

Relationship between honeybee nutrition and their microbial communities

Miriane Acosta Saraiva · Ana Paula Pegoraro Zemolin · Jeferson Luis Franco ·
Juliano Tomazzoni Boldo · Valdir Marcos Stefenon · Eric W. Triplett ·
Flávio Anastácio de Oliveira Camargo · Luiz Fernando Wurdig Roesch

Received: 30 October 2014 / Accepted: 12 January 2015
© Springer International Publishing Switzerland 2015

Abstract The microbiota and the functional genes actively involved in the process of breakdown and utilization of pollen grains in beebread and bee guts are not yet understood. The aim of this work was to assess the diversity and community structure of bacteria and archaea in Africanized honeybee guts and beebread as well as to predict the genes involved in the microbial bioprocessing of pollen

using state of the art ‘post-light’ based sequencing technology. A total of 11 bacterial phyla were found within bee guts and 10 bacterial phyla were found within beebread. Although the phylum level comparison shows most phyla in common, a deeper phylogenetic analysis showed greater variation of taxonomic composition. The families Enterobacteriaceae, Rickettsiaceae, Spiroplasmataceae and Bacillaceae, were the main groups responsible for the specificity of the bee gut while the main families responsible for the specificity of the beebread were Neisseriaceae, Flavobacteriaceae, Acetobacteraceae and Lactobacillaceae. In terms of microbial community structure, the analysis showed that the communities from the two environments were quite different from each other with only 7 % of species-level taxa shared between bee gut and beebread. The results indicated the presence of a highly specialized and well-adapted microbiota within each bee gut and beebread. The beebread community included a greater relative abundance of genes related to amino acid, carbohydrate, and lipid metabolism, suggesting that pollen biodegradation predominantly occurs in the beebread. These results suggests a complex and important relationship between honeybee nutrition and their microbial communities.

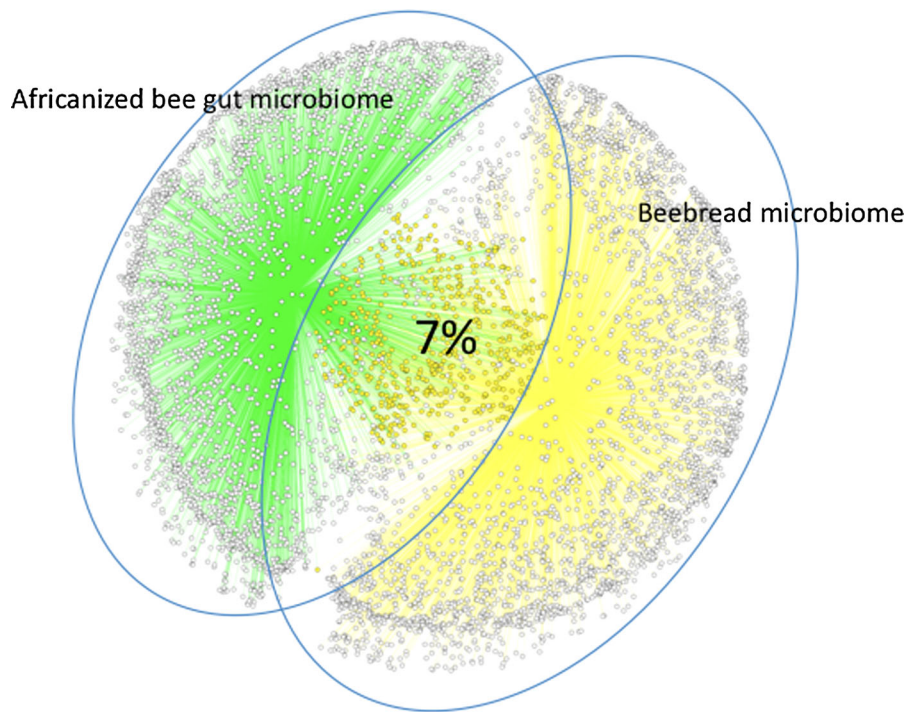
Electronic supplementary material The online version of this article (doi:[10.1007/s10482-015-0384-8](https://doi.org/10.1007/s10482-015-0384-8)) contains supplementary material, which is available to authorized users.

M. A. Saraiva · A. P. P. Zemolin · J. L. Franco ·
J. T. Boldo · V. M. Stefenon · L. F. W. Roesch (✉)
Universidade Federal do PAMPA, Campus São Gabriel,
Avenida Antônio Trilha, 1847, São Gabriel,
Rio Grande do Sul 97300-000, Brazil
e-mail: luizroesch@unipampa.edu.br

E. W. Triplett
Microbiology and Cell Science Department, Institute of
Food and Agricultural Sciences, University of Florida,
Gainesville, FL 32611-0700, USA

F. A. de Oliveira Camargo
Departamento de Solos, Universidade Federal do Rio
Grande do Sul, Av. Bento Gonçalves, 7712, Porto Alegre,
Rio Grande do Sul 91540-000, Brazil

Graphical abstract



Keywords Post-Light™ ion semiconductor sequencing · Microbial functional prediction · Microbial phylogeny · Africanized honeybee · *Apis mellifera*, 16S rRNA

Introduction

Honeybees are considered the most important group of plant pollinators in many ecosystems (Bawa et al. 1985; Bawa 1990; Shipp et al. 1994; Heard 1999). They contribute significantly to agricultural productivity and profitability.

The microbial gut community of *Apis mellifera* might be involved in nutritional processes, such as breakdown and utilization of pollen grains as well as degradation of toxic compounds found in the environment (Engel et al. 2012). These gut bacteria may also provide defense responses against pathogens by producing inhibitory compounds or by monopolizing nutrients within the gut (Forsgren et al. 2010; Koch and Schmid-Hempel 2011; Engel et al. 2012). Although many functions of the bee microbiota have

been discovered, adult honeybees and bumblebees harbor a specialized but surprisingly species-poor community of bacteria in their guts (Koch and Schmid-Hempel 2011; Martinson et al. 2011).

The best food sources for bee colonies are nectar and pollen, and various insects and vertebrates are able to use pollen as food source (Roulston and Cane 2000). However, pollen is difficult to digest and honeybees are unable to consume fresh unprocessed pollen. The potential microbial involvement in pollen breakdown is especially important for supplemental feeding of honeybee colonies during non-foraging periods. Many artificial high-protein diets containing no pollen but rather protein from soybean are currently commercialized. Although the soybean contains all essential amino acids required by the bee's diet, the lack of a microbial pre-fermentation might limit the total use of this protein source. This observation might also explain why fermented diets are consumed more readily than unfermented ones (Ellis et al. 2006). Although pollen is nutrient-rich, the carbohydrate exine, which is refractory to most digestive systems, protects pollen from digestion (Roulston and Cane 2000). Pollen stored in the honeycomb receives the

addition of digestive enzymes through the workers' saliva (Gilliam 1979). These enzymes' actions are capable of changing the pH in the honeycomb, making the environment selective to some microbial taxa. While changes in the nutritional composition of pollen appear to be guided by microbial activity, the microbiota and the functional genes actively involved in this process in beebread and bee guts are not yet understood. Pioneer studies characterized the microbiota of pollen and beebread by culture-dependent approaches (Gilliam 1979). Those studies were very important to access the cultivable microorganisms; however, many microorganisms defy standard cultivation approaches. As a result, molecular methods to sequence 16S rRNA genes by high-throughput sequencing and the phylogenetic investigation of communities by reconstruction of unobserved states should enhance our understanding of the bee microbiome.

Most of the studies involving associations with bees and microorganisms have been conducted with European bees or bumblebees. However, studies of Africanized bees, which are more resistant to diseases and more productive than European bees, are scarce. The intent of this work was to enhance our knowledge about microbial communities inhabiting both bee guts and beebread by assessing their similarities and/or differences as well as predict the genes involved in the microbial bioprocess of pollen using state of the art 'post-light'-based sequencing technology.

Materials and methods

Ethics Statement

All colonies were sampled in an apiary of the Cooperativa Apícola do Pampa Gaúcho located in São Gabriel, Rio Grande do Sul State, Brazil. Our field collections did not involve endangered or protected species and no specific permissions were required since the responsible beekeeper granted access to the apiary. The authors declare no conflict of interest.

Experiment design and DNA extraction

An apiary containing 10 honeybee colonies (located 5 m away from each other) was established during springtime (September–December, 2012) from queens that had been inseminated naturally. According to

Whitfield et al. (2006), honeybee colonies in Southern Brazil are product of a genetic mixture of 25 % European and 75 % African populations as a result of the introduction of *A. mellifera scutellata* in 1956. The apiary was located at the boundary between a pristine forest segment and a 7-year old *Eucalyptus* sp. forest. A general scheme of the experimental design and analysis is shown on Fig. 1. Bee gut samples were obtained from 10 worker nurse bees, removed from each one of the colonies ($n = 100$) after a 30-day period after the establishment of the colony. Worker bees were placed in a 15 mL tube (1 tube per colony), kept on ice, brought to the lab, and immediately dissected. Gut organs (crop, hindgut, midgut, ileum, and rectum) were humidified with sterile saline solution (0.9 % w/v NaCl) and removed using sterile tweezers and scissors. After dissection, gut organs were stored at -20°C for a maximum of 24 h. For beebread samples, a frame containing combs filled with beebread was removed from each colony, also after a 30-day period from colonies' establishment. These combs were placed inside a sterile plastic bag, kept on ice, and brought to the lab. Three samples of beebread were taken from the combs using sterile tweezers and needles and kept at -20°C until further use. Worker bees (=bee gut) and beebread were sampled in the same day. Microbial DNA extraction from gut and beebread samples was carried out using the QIAamp DNA Stool kit (QIAGEN, USA), according to the manufacturer's instructions. For each sample type per colony, individual DNA extractions were pooled. The nucleic acid concentration and purity of each sample was evaluated by the NanoVue spectrophotometer (GE Life Sciences, USA). Extracted DNA samples were kept at -20°C until further use.

For the overall analysis of the microbial diversity and structure comparisons from each bee gut or beebread, the 16S rRNA gene from the V3 region was amplified using the primer fusion technique and sequenced with the Ion Torrent Personal Genome Machine (Life Technologies, USA). Finally, to obtain high sequence coverage and to detect the rare microbes inhabiting the samples, a second sequencing run was performed using one single sample randomly chosen from each bee gut and beebread. The 16S rRNA gene from the V6 region was amplified using the adaptor ligation technique and sequenced with the Ion Torrent Personal Genome Machine (Life Technologies, USA).

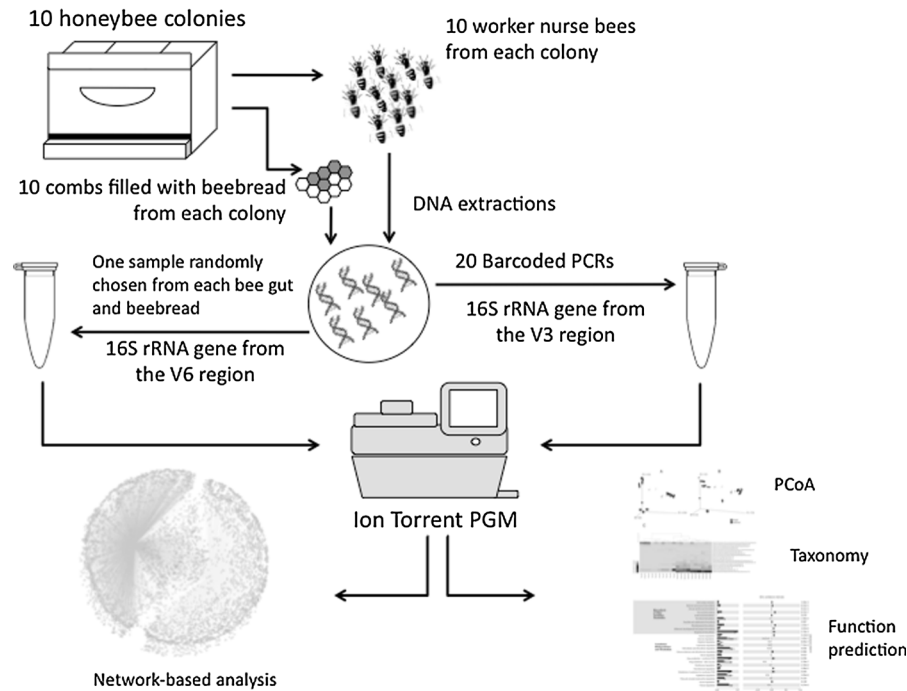


Fig. 1 General scheme of the experimental design and workflow conducted for the analysis of the microbial communities in the bee guts and bee bread

Amplicon libraries preparation using primer fusion

Three independent PCR reactions were performed for each of the 20 samples with the 515F and 806R primers (Caporaso et al. 2012). The forward primer was modified adding the P1 adaptor sequence (5'-CCTCTCTATGGGCAGTCGGTGAT-3') at the 5'-end of the primer. To reduce any effect the composite primer might have on PCR efficiency, a two-base linker (GT) was inserted between the primer and the P1 adaptor. Barcoded primers were used to multiplex the amplicon pools so they could be sequenced together and computationally separated afterwards. To do this, 12-base barcodes were added at the 5'-end of the reverse primers using the self-correcting barcode method of Hamady et al. (2008). A two-base linker (CC) was inserted between the primer and the barcode. The A1 adapter (5'-CCATCTCATCCCTGCGTGCTCCGATCTCAG-3') was then added to the 5'-end of the barcoded primer.

PCR was performed with the High Fidelity PCR Enzyme Mix (Thermo Scientific, USA). The mixtures contained 5 µl of 10X high fidelity PCR buffer with 15 mM MgCl₂, 0.2 mM of each dNTP, 100 mM of

each primer, 2.5 U of high fidelity PCR enzyme mix, and approximately 100 ng of DNA template in a final volume of 50 µl. The PCR conditions were 94 °C for 2 min, 25 cycles of 94 °C for 45 s; 55 °C for 30 s; and 72 °C for 30 s extension, followed by 72 °C for 4 min. The PCR products for each of the 20 samples were purified and combined in equimolar ratios using the quantitative DNA binding method (SequalPrep Kit, Invitrogen, Carlsbad, CA, USA) to create a DNA pool for template amplification (EmPCR) and sequencing.

Short amplicon library preparation using adaptor ligation

Three independent PCR reactions were performed for each sample with the primers 917F (Keijser et al. 2008) and 1046R (Sogin et al. 2006) primers, for the amplification of approximately 130 base pairs of the V6 region of the 16S rRNA gene. PCR was performed with the High Fidelity PCR Enzyme Mix (Thermo Scientific, USA). The mixtures contained 5 µl of 10X high fidelity PCR buffer with 15 mM MgCl₂, 0.2 mM of each dNTPs, 100 mM of each primer, 2.5 U of high fidelity PCR enzyme mix, and approximately 100 ng

of DNA template in a final volume of 50 µl. The PCR conditions were 94 °C for 2 min, 25 cycles of 94 °C for 45 s; 56 °C for 30 s; and 72 °C for 30 s extension, followed by 72 °C for 4 min.

Prior to Ion Torrent PGM sequencing, the short amplicon libraries were processed in order to add the barcoded adaptors, A and P1, necessary for sequencing. The adaptors were added to the amplicons using the Ion Plus Fragment Library Kit and the Ion Xpress™ Barcode Adapters (Life Technologies, USA). The reactions were performed based on the user bulletin MAN0006846 revision 3.0 available at <http://ioncommunity.lifetechnologies.com>, with minor modifications during the amplicon purification step as follows: the bead suspension with the DNA was incubated with the Agencourt® AMPure® XP Reagent (Beckman Coulter, USA) (2× sample volume) at room temperature for 10 min and all washing steps were performed with 500 µl of freshly prepared 80 % ethanol during 30 s. All other steps for preparing short amplicon libraries (end-repair, barcoded adaptors ligation and nick-repair) were performed according to the user bulletin mentioned above. The following barcodes TTGGAGTGTC and TCTAGAGGTC were added to the short amplicons from bee gut and beebread samples, respectively.

All barcoded amplicons were quantified by quantitative real-time PCR using Ion Library Quantitation Kit (following the user guide Ion Library Quantitation Kit TaqMan® assay quantitation of Ion Torrent libraries Publication Part Number 4468986 Rev. A) and the Applied Biosystems® 7500 Fast Real-Time PCR System according to the manufacturer's instructions. The samples were adjusted to 15×10^6 molecules per microliter and mixed in equal amounts to obtain an equimolar pool of amplicons that was used for template amplification onto Ion Sphere™ Particles (ISPs).

Template amplification and sequencing

The template-positive ISPs containing clonally amplified 16S rRNA genes (either fusion primer technique or adaptors ligation technique) was prepared with the Ion OneTouch™ System using the Ion OneTouch™ 200 Template Kit v2 following the user guide Publication Number 4478372 Revision B (available at <http://ioncommunity.lifetechnologies.com>). The resulting ISPs were sequenced on Ion 316™

microchip using the Ion Torrent Personal Genome Machine (Life Technologies, USA) and the Ion PGM™ 200 Sequencing Kit following the workflow suggested in the user guide Part Number 4474246. After sequencing, the sequence reads were filtered within the PGM software that removed low quality and polyclonal sequences. All PGM filtered data were exported as Fastq file that was used for the subsequent bioinformatics analysis.

16S rRNA reads processing for downstream analyses

A fundamental problem with the use of next-generation sequencing for single read analysis is the number of artifacts that might exacerbate biases via the presence of chimeric sequences and sequence errors (Schloss et al. 2011). In attempt to reduce the sequencing error, a stringent quality-filtering method was applied to improve our downstream analyses. The Fastq file exported from the Ion PGM™ System was processed using mothur v.1.30.2 (Schloss et al. 2009). The multiplexed reads were first filtered for quality and assigned to the starting samples. The filtering criteria removed any sequence that: (1) contained a homopolymer greater than 8 bases, (2) contained any ambiguous base call, (3) had more than one mismatch to the barcode sequence, and (4) was shorter than 100 bases in length. Also, the sequences were quality screened using a moving window that was 50 bases long. Within that window, any read was removed with an average quality score (inferred as Phred score) below 25. Finally the chimeric sequences were checked using the chimera.slayer command and any sequences that were of mitochondrial or chloroplast origin were removed from the dataset.

Library comparisons

For the overall comparison of significant differences among bacterial communities from bee guts and beebread, Principal Coordinates Analysis (PCoA) was performed with the amplicon libraries. PCoA is an ordination method based on multivariate statistical analysis that maps the samples in different dimensions and reflects the similarity of the biological communities. The calculations were performed within the QIIME pipeline (Caporaso et al. 2010). A matrix using the UniFrac metric (weighted and unweighted) for

each pair of environments was calculated. The distances were converted into points in space with the number of dimensions one less than the number of samples. The first three principal dimensions were used to plot a three-dimensional graph that was visualized using KING (Chen et al. 2009). To test whether the results were robust, a sequence-jackknifing technique was used in which the PCoA clusters were regenerated using a subset corresponding to about 70 % of the total number of sequences obtained in the sample with the smaller number of sequences randomly selected from each sample for 100 replicate trials. In addition, to see which taxa were more prevalent in different areas of the PCoA plot, the ten most abundant taxa were added to the PCoA plots. Ellipses were drawn around the mean values to represent the interquartile ranges (IQRs), which measures the statistical dispersion between the upper and lower quartiles. If the ellipses are small, the same result would likely to be achieved with a different set of sequences from the same environment. Alternatively, if the ellipses are large, a different result might be expected. In addition, to determine which taxa were more prevalent in different areas of the PCoA plot, the nine most abundant family-level taxa were combined with the jackknifed principal coordinates.

To identify the organisms whose abundances differ between the bee gut and beebread samples, each sequence was assigned to an OTU at 95 % similarity by using uclust (Edgar 2010). The number of sequences in each OTU found in each sample was used to construct a table with OTUs (lines) and samples (columns). This table was used to generate a color-coded Clustered Image Map (heat map) with CIMminer (Weinstein et al. 1997). Finally, representative sequences of each OTU were subjected to the RDP naïve Bayesian rRNA Classifier (Wang et al. 2007), which attaches complete taxonomic information from domain to species for each sequence in the database with 80 % taxonomy confidence and an *e*-value of 0.001. That means some OTUs might only be classified at certain coarse taxonomic level because a deeper level (e.g. genus) would not present the required degree of confidence.

To examine the co-occurrence of OTUs between the samples a network-based analysis was applied. The network was performed using the dataset generated with the short amplicon libraries because these

libraries had greater sequencing coverage (Lemos et al. 2011). The network allows for the visualization of the OTUs that are either unique or shared by specific groups of bee gut and beebread samples. Connections were drawn between samples and OTUs, with edge weights defined as the number of sequences from each OTU that occurred in each sample. To cluster the OTUs from the bee gut or beebread samples in the network, the spring embedded algorithm layout was used. In this algorithm, the nodes act like steel rings that exert mechanical forces creating an attractive force between nodes that are far apart, and a repulsive force between nodes that are close together (Shannon et al. 2003). In the resultant graph, samples that share more OTUs cluster closer together. The network diagram was visualized with Cytoscape (Shannon et al. 2003) with two kinds of nodes; OTU-nodes, grouped using 97 % similarity cutoff, and bee gut and beebread sample nodes.

Metagenome prediction

The 16S rRNA database was used to predict the microbial functions within the bee gut and the beebread using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt). The software uses an extended ancestral-state reconstruction algorithm to predict which gene families are present and then combines genes families to estimate the composite metagenome (Langille et al. 2013). Briefly, the demultiplexed reads were subjected to a closed-reference OTU picking protocol (Caporaso et al. 2010) where the sequences were searched against a greengenes reference database at 97 % identity. Any read that did not hit the reference collection was discarded. The outputted OTU table was normalized according to the 16S rRNA gene copy number per genome and a final metagenome functional prediction was created.

To determine statistical differences between the bee gut and the beebread metagenomes, the statistical analysis of metagenomic profiles (STAMP) software package was used (Parks and Beiko 2010). The statistical hypothesis tests were performed using the Welch's *t* test (Bluman 2007) while confidence intervals were calculated using the Welch's inverted method, and correction was made using Bonferroni multiple test correction (Adbi 2007).

Data accessibility

Raw sequences were submitted to the NCBI Sequence Read Archive under the study number SRP043294, experiment number SRX603452.

Results

Using the library prepared with the primer fusion technique as described in the “[Amplicon libraries preparation using primer fusion](#)” section, a total of 505,606 reads were obtained. After filtering the reads by base quality and removing reads shorter than 100 bases, a total of 207,628 high-quality 16S rRNA gene sequences were obtained in this study (76,628 and 131,000 sequences for bee gut and beebread samples, respectively). After taxonomic assignment and chimera removal, a total of 11 phyla (Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Firmicutes, Gemmatimonadetes, Planctomycetes, Proteobacteria, TM7, Tenericutes and Verrucomicrobia) were found within bee guts and 10 phyla (Acidobacteria, Actinobacteria, Bacteroidetes, Firmicutes, Gemmatimonadetes, OD1, Planctomycetes, Proteobacteria, Tenericutes and Verrucomicrobia) were found within beebread. Firmicutes and Proteobacteria were the most abundant phyla within both beebread (29 % of Firmicutes and 62 % of Proteobacteria) and bee guts (16 % of Firmicutes and 19 % of Proteobacteria). No archaeal sequences were found in the beebread samples and only one sequence of the phylum Crenarchaeota and five sequences of the phylum Euryarchaeota were found in the bee gut samples.

In terms of microbial community structure, the analysis showed that the communities from the two environments were quite different from each other (Fig. 2a, b). The first three axes of the Jackknifed unweighted PCoA explained 42 % of the total variation found within microbial communities while the first three axes of the weighted Jackknifed PCoA accounted for 89.3 % of the variation. The results indicated that the overall differences between the clusters were related to both presence/absence and the abundance of specific OTUs. The group made up by the microbial communities from the beebread is illustrated as a dense cluster indicating that the microbial communities from this environment were closely related in terms of membership and structure. On the other hand, the group

represented by the microbial communities from the bee guts is shown as a dispersed cluster indicating a greater variation of microbial membership and structure when compared with the beebread. The families Enterobacteriaceae, Rickettsiaceae, Spiroplasmataceae and Bacillaceae, were the main groups responsible for the specificity of the bee gut with the first two families being the most abundant (Fig. 2c). The main families responsible for the specificity of the beebread were Neisseriaceae, Flavobacteriaceae, Acetobacteraceae and Lactobacillaceae.

To examine the co-occurrence of OTUs between the samples, a network-based analysis was applied using the dataset generated with the short amplicon libraries as described in the “[Short amplicon library preparation using adaptor ligation](#)” section. From this library, a total of 448,117 and 360,797 high-quality sequences from the V6 region were obtained from the bee guts and beebread, respectively. These data provided coverage greater than 99 % as calculated by the Good’s coverage (Good 1953). The network consisted of 6,208 OTUs (called nodes), in which 2,646 OTUs were found exclusively in the bee guts and 3,032 were present only in the beebread (Fig. 3). The network analysis revealed that only 7 % (530 OTUs) were shared between bee gut and beebread, confirming the very distinctive microbiota of bee guts and beebread.

On the basis of the PICRUSt analysis, 6,909 enzymes involved in metabolism processes were predicted (please refer to “[Discussion](#)” section for more information about the limitations and interpretation of these results) within the bee guts and beebread. These enzymes were grouped in 114 functional profiles related to 12 main metabolic functions (Figs. 4, 5). Those metabolic functions were found in both bee gut and beebread. However, it was possible to detect statistically significantly higher abundance of sequences matching the following functional profiles/pathways in the beebread: (a) amino acid metabolism—11 pathways; (b) carbohydrate metabolism—10 pathways; (c) glycan biosynthesis and metabolism—7 pathways; (d) lipid metabolism—11 pathways; (e) metabolism of cofactors and vitamins—7 pathways; (f) metabolism of other amino acids—7 pathways; (g) nucleotide metabolism—2 pathways; (h) xenobiotics biodegradation and metabolism—10 pathways. In contrast, the following functional profiles/pathways were predicted to

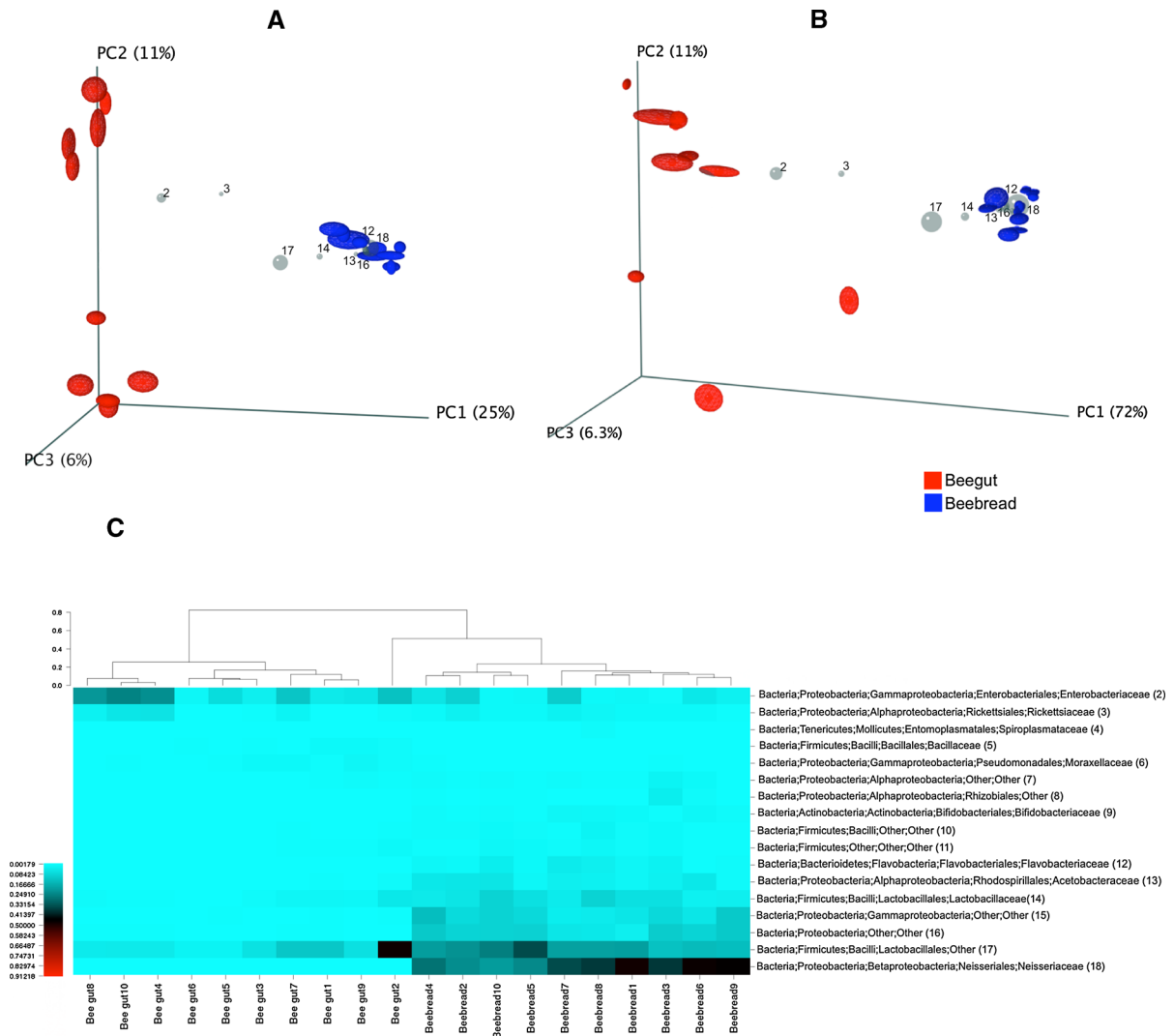


Fig. 2 Overall comparisons of microbial communities based on their composition. **a, b** 3D Jackknifed Principal Coordinates (PCoA) biplots depicting the clusters of bacterial communities within 20 samples each from the bee guts (red) and beebread (blue). **a** Unweighted UniFrac distance metrics. **b** Weighted UniFrac distance metrics. The positions of the points are the average for the jackknife replicates and ellipses around points represent the interquartile range (IQR) for the 1,000 jackknife replicates. Each point represents a microbial community. Points closer to each other represent similar microbial communities

while points far from each other represent dissimilar microbial communities. The gray circles (sized according to the bacterial relative abundance) represent the bacterial taxa, plotted in the same PCoA space, contributing to the PCoA clusters. The numbers next the gray circles correspond to bacterial families as depicted in the heatmap. **c** Heat map generated from 16S rRNA data reflecting the major differences in taxa abundance across bee gut and beebread samples. Note that only the most representative taxa are shown in the heat map. A complete heat map with all taxa grouped by family can be found in the Fig. S1

be statistically significantly more abundant in the bee guts: (a) biosynthesis of other secondary metabolites—8 pathways (mainly related to resistance and biosynthesis of antibiotics); (b) energy metabolism—5 pathways; (c) enzyme families—2 pathways (peptidases and protein kinases).

Discussion

In this study, we focused on the analysis and comparison of *Apis mellifera* gut and beebread microbiome by using state of the art ‘post-light’ based sequencing technology (Rothberg et al. 2011). We

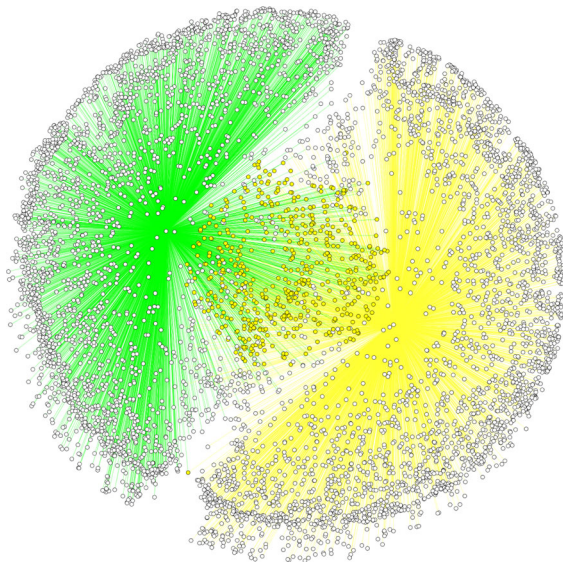


Fig. 3 Network based analysis of microbial communities. White circles are representative sequences from each OTU clustered based on the relatedness of the sequences (97 % similarity). Yellow circles are representative sequences shared between bee gut and beebread. Each bee gut or beebread samples are connected with the OTUs through edges color-coded according to the origin of the sequence. Green bee gut and yellow beebread

investigated the presence/absence of microbial taxa, the microbial community structure and its possible involvement with nutritional processes such as the breakdown and utilization of pollen grains.

The relative abundance of archaea in these samples was very low within the guts of *Apis mellifera*. Only 6 of 207,628 high-quality 16S rRNA gene sequences were classified as Archaea. In fact, Archaea do not appear to be important members of these communities. However, it should be mentioned that biases at the steps of DNA extraction, PCR amplification, primer choice and sequencing, might present some degree of interference in the results obtained. Nevertheless, as Probe Match (<http://rdp.cme.msu.edu/probematch/search.jsp>) predicted that the primers used in this study would amplify 16S rRNA from Archaea, these biases were unlikely to have missed many archaeal taxa.

Large variability of the bee microbiome within different colonies was observed through the PCoA analysis (Fig. 2). Several exogenous as well as endogenous factors could affect the intestinal microbiota in these bees. The environment of the sampled

insects was the same, thus minimizing its contributions to microbiome variability. Another important factor in determining the microbial diversity and community structure inside the bee guts is the age of the insects and their respective function in the colony. Younger workers may have maternally inherited microbiomes while older workers, which leave the hive to forage, might present microbiomes acquired from the environment or via flowers (Martinson et al. 2011; McFrederick et al. 2012). Finally, the microbiome acquisition might be determined by the bees' health status (Cox-Foster et al. 2007).

The microbial communities of the bee guts and the beebread were easily distinguished (Figs. 2, 3). Such large differences were expected since both guts and beebread are from environments that differ in temperature, pH and osmotic potential. Thus, only seven percent of the total OTUs were found in common between guts and beebread. Both environments presented high diversity of taxa and those taxa present different functions involving nutrient processing. The exclusive conditions of each environment supported a set of distinct microbial functions.

The Lactobacillaceae dominated beebread samples (Fig. 2c). The presence of representatives of Lactobacillaceae has already been reported in honeybee workers (Vojvodic et al. 2013). In fact, Lactobacillales were also found in association with the bee guts (see Fig. 2c). Recent evidence point to the presence of a microbial core community within the European bee guts and the bacteria belonging to the Lactobacillales order appear to be present in this core community (Moran et al. 2012). Sampling two localities in USA and five bees from four colonies, Moran et al. (2012), found *Gilliamella* (Gamma 1) *Snodgrassella* (Beta), Firm4 and Firm5 present in all European bees and Bifido, Gamma2 and Alpha2 present in most European bees (for detailed phylogenetic identification of those groups see Martinson et al. 2011). In our work, Africanized bees were collected from only one locality in Brazil but from ten colonies. At the phylum level, the same representatives of the microbial core were found in Africanized bees, indicating that this bacterial core might also be present in Africanized populations. Within the beebread, Lactobacillaceae might be involved in fermentative processes or in the production of secondary metabolites responsible for changes in the pH of the beebread. Studying the origin of *Lactobacillus* in the *Apis* genus, McFrederick et al.

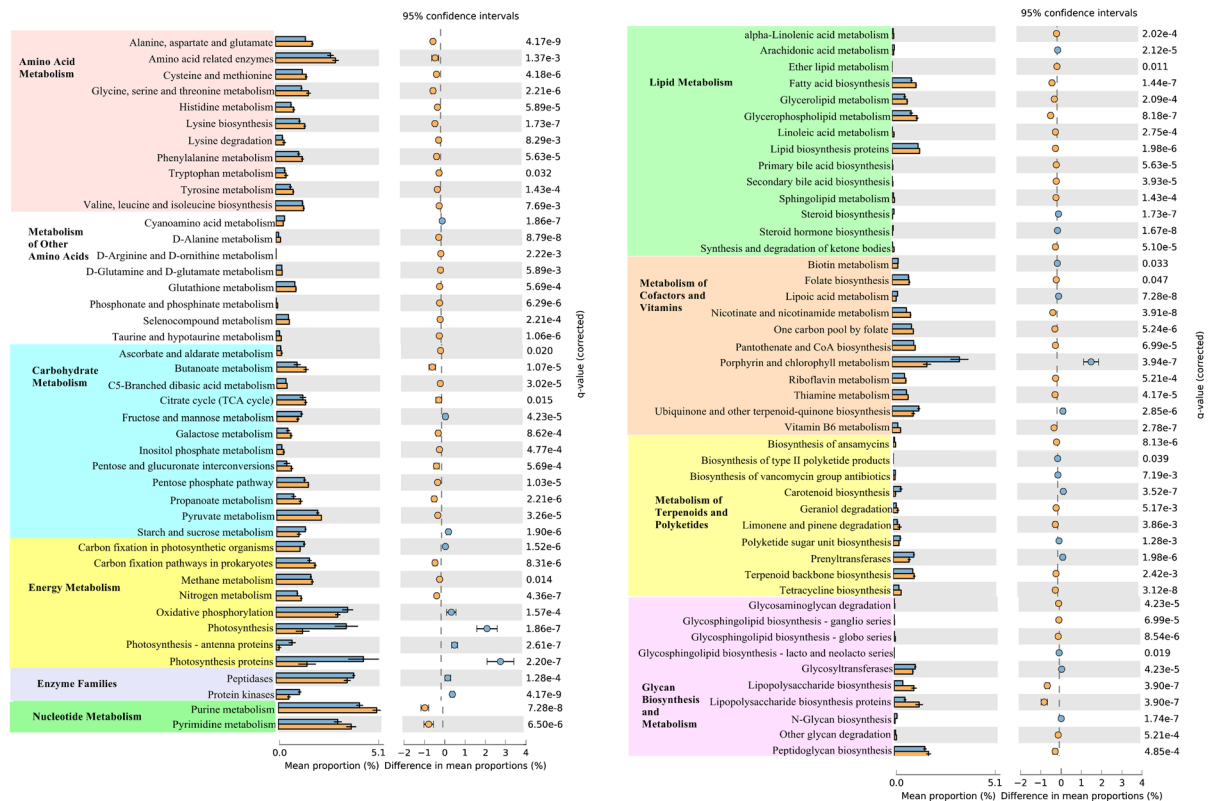


Fig. 4 Relative abundance of microbial metabolic functions from 10 composite bee guts (blue) and 10 composite beebread (orange) samples. The error bars show calculated standard variation. The colored circles represent the 95 % confidence

intervals calculated using the Welch's inverted method. Corrected q-values were calculated using Bonferroni Multiple test correction (q-value < 0.05)

(2012) reported the presence of close relatives of several bacteria that have been isolated from flowers. In addition to Lactobacillaceae, Enterobacteriaceae were abundant within bee guts, while Neisseriaceae were common within beebread. Working with European bees captured in two different seasons (spring and fall) in an apiary at Arizona, USA, Corby-Harris et al. (2014) also found *Lactobacillus* and sporadically abundant Enterobacteriaceae in crops of foragers' bees. Such observations are in line with the hypothesis that at least Lactobacillaceae is frequently found in bee guts irrespective of the geographic localization, environmental conditions (Moran et al. 2012) or bee genetics (European bees from USA or Africanized bees from Brazil as found in our study). Anderson et al. (2013) sampled beebread and bee gut of honey bees from apiaries located at the USDA Carl Hayden Bee Research Center in Tucson AZ, USA. Using

cultivation techniques and rRNA and rDNA sequencing, the researchers compared bacterial diversity from floral nectar, different segments of the honey bee alimentary tract, and beebread. Similar to our findings, within the beebread they found Firmicutes (*Lactobacillus*), Actinobacteria, Proteobacteria (Acetobacteriaceae) and Bacteroidetes.

An important challenge faced by microbial ecologists is to discover the metabolic differences that are affected by the differences in microbial taxa. This can be assessed with metagenomic or metatranscriptomic approaches. However, in this work we took advantage of a bioinformatics tool developed by Langille et al. (2013), capable of predicting the metagenome based on phylogenetic information. According to Langille et al. (2013), with only 105 assigned 16S rRNA gene reads, PICRUSt can provide the same accuracy of 15,000 annotated metagenomic sequences. PICRUSt

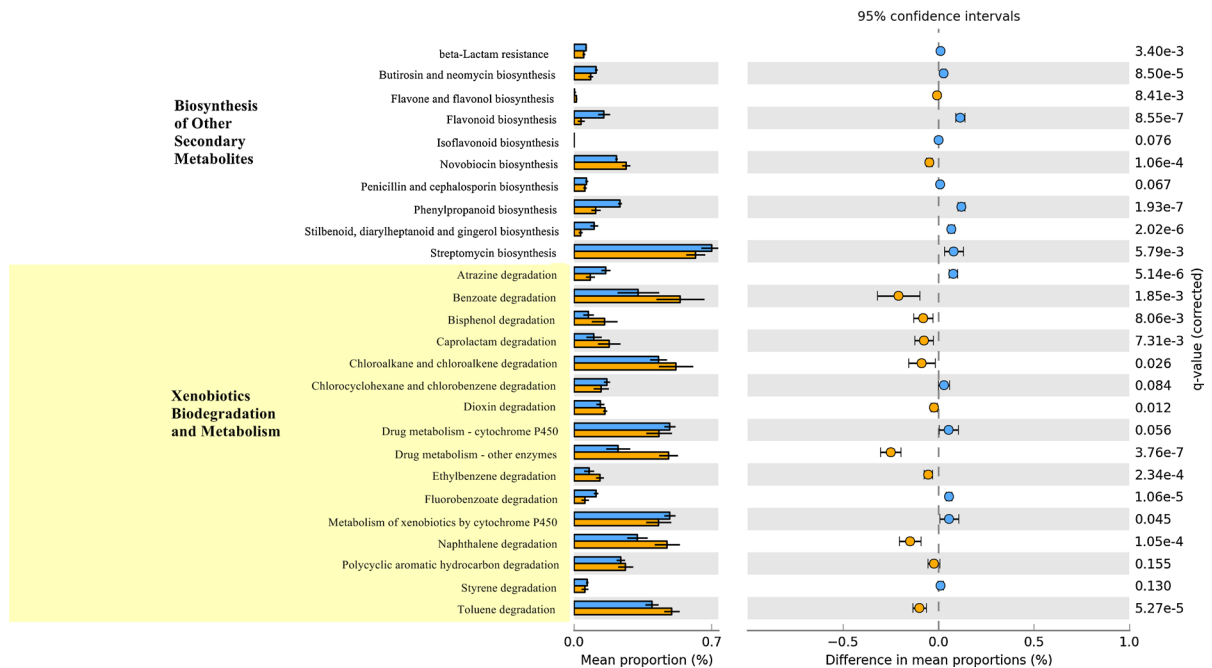


Fig. 5 Relative abundance of microbial genes related to biosynthesis of secondary metabolites and xenobiotics biodegradation and metabolism from 10 composite bee guts (blue) and 10 composite beebread (orange) samples. The error bars show

calculated standard variation. The colored circles represent the 95 % confidence intervals calculated using the Welch's inverted method. Corrected q-values were calculated using Bonferroni Multiple test correction (q-value < 0.05)

does not provide direct evidence of a community's functional capabilities but might provide useful insights about the functions the microbial community might be able to perform (Langille et al. 2013).

Genes involved in the biosynthesis of streptomycin were found in greater proportion than the genes related to the biosynthesis of other antibiotics. Streptomycin is an antibiotic that belongs to the aminoglycoside category and acts by inhibiting the protein synthesis commonly against gram-negative bacteria (Singh and Mitchison 1954). The presence of such genes raises the hypothesis about the possible role of commensal or normal microbiota in protecting *Apis mellifera* against pathogenic bacteria and in maintaining the healthy status of the hive.

In addition to the possibility of protection against pathogens, the microbiota from the beebread possessed many genes related to xenobiotics biodegradation, including phase I detoxification systems such as cytochrome P450 (CYP). The superfamily of CYP proteins is responsible for the catabolism of 90 % of all drugs (Liu et al. 2013). Thus, these genes may

provide important defense mechanisms against exogenous chemicals (Guengerich 2007). The presence of xenobiotic substances in a given organism is usually associated with environmental exposures to a range of compounds including drugs, natural compounds, as well as synthetic substances such as pesticides. Pollinators, such as honeybees, are directly exposed to chemical agents used in agriculture. Our metagenome prediction suggested the presence of many genes that can degrade the most commonly used pesticides in agriculture such as atrazine and fluorobenzoate. Such compounds are routinely used for weed control.

This work represents the evidence for the presence of microbial genes related to the pollen degradation in the beebread highlighting the importance of the microbiota associated with the beebread and its potential role in pollen breakdown. In addition, the results highlight the possible involvement of the microbiota associated with the bee guts in building up a defense system based on the biosynthesis of antibiotics and xenobiotics biodegradation. Xenobiotic degradation may be important in the

degradation of herbicides or insecticides carried with pollen and deposited in the combs.

In summary, the microbial communities of the bee guts and the beebread were very distinctive in terms of community structure and composition indicating the presence of a highly specialized and well-adapted microbiota within each bee gut and beebread. In fact, only a small proportion of the microbial taxa were found in common between bee guts and beebread. These differences might be related with the role of these communities within each environment. For instance, the microbial community from the beebread works as a bioreactor for pollen digestion, and contributes to the detoxification processes that provide bees with an increased ability to eliminate the toxic compounds carried with pollen.

Acknowledgments The National Council for Scientific and Technological Development (CNPq—Brazil), the Coordination for the Improvement of Higher Education Personnel (CAPES—Brazil), the Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS—Brazil) and Programa de Apoio aos Polos Tecnológicos; Process 474-2500/13-9 supported this work. LFW Roesch, JL Franco and FAO Camargo received research fellowships from the CNPq. MA Saraiva received research fellowships from Fundação de Amparo a Pesquisa do Estado do Rio Grande do Sul (FAPERGS—Brazil). Special thanks are due to the Cooperativa Apícola do Pampa Gaúcho and Mr. Aldo Machado dos Santos for granting access to their apiaries.

References

- Adbi H (2007) Encyclopedia of measurement and statistics. Sage, Thousand Oaks
- Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, Schwan MR, Walton BM, Jones BM, Corby-Harris V (2013) Microbial ecology of the hive and pollination landscape: bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). PLoS One 8:e83125
- Bawa KS (1990) Plant–pollinator interactions in tropical rain forests. Annu Rev Ecol Syst 21:399–422
- Bawa KS, Bullock SH, Perry DR, Coville RE, Grayum MH (1985) Reproductive biology of tropical lowland Rain Forest Trees. II. Pollination systems. Am J Bot 72:346–356
- Bluman AG (2007) Elementary statistics: a step by step approach, 6th edn. McGraw Hill Higher Education, New York
- Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Knight R (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7:335–336
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624
- Chen VB, Davis IW, Richardson DC (2009) KING (Kinemage, Next Generation): a versatile interactive molecular and scientific visualization program. Prot Sci 18:2403–2409
- Corby-Harris V, Maes P, Anderson KE (2014) The bacterial communities associated with honey bee (*Apis mellifera*) foragers. PLoS One 9:e95056
- Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, Moran NA, Quan PL, Briesse T, Hornig M, Geiser DM, Martinson V, van Engelsdorp D, Kalkstein AL, Drysdale A, Hui J, Zhai J, Cui L, Hutchison SK, Simons JF, Egholm M, Pettis JS, Lipkin WI (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. Science 318:283–287
- Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461
- Ellis JS, Knight ME, Darvill B, Goulson D (2006) Extremely low effective population sizes, genetic structuring and reduced genetic diversity in a threatened bumblebee species, *Bombus sylvarum* (Hymenoptera: Apidae). Mol Ecol 15:4375–4386
- Engel P, Martinson VG, Moran NA (2012) Functional diversity within the simple gut microbiota of the honey bee. Proc Natl Acad Sci USA 109:11002–11007
- Forsgren E, Olofsson TC, Vásquez A, Fries I (2010) Novel lactic acid bacteria inhibiting *Paenibacillus larvae* in honey bee larvae. Apidologie 41:99–108
- Gilliam M (1979) Microbiology of pollen and beebread: the yeasts. Apidologie 10:43–53
- Good IJ (1953) The population frequencies of species and the estimation of population parameters. Biometrika 40:237–264
- Guengerich FP (2007) Cytochrome P450 and chemical toxicology. Chem Res Toxicol 21:70–83
- Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. Nat Methods 5:235–237
- Heard TA (1999) The role of stingless bees in crop pollination. Annu Rev Entomol 44:183–206
- Keijser BJ, Zaura E, Huse SM, Van der Vossen JM, Schuren FH, Montijn RC, Crielaard W (2008) Pyrosequencing analysis of the oral microflora of healthy adults. J Dent Res 87:1016–1020
- Koch H, Schmid-Hempel P (2011) Bacterial communities in central European bumblebees: low diversity and high specificity. Microb Ecol 62:121–133
- Langille MG, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, Huttenhower C (2013) Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. Nat Biotechnol 31:814–821
- Lemos LN, Fulthorpe RR, Triplett EW, Roesch LFW (2011) Rethinking microbial diversity analysis in the high-throughput sequencing era. J Microbiol Methods 86:42–51
- Liu J, Tawa GJ, Wallqvist A (2013) Identifying cytochrome p450 functional networks and their allosteric regulatory elements. PLoS One 8:e81980
- Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, Moran NA (2011) A simple and distinctive microbiota associated with honey bees and bumble bees. Mol Ecol 20:619–628

- McFrederick QS, Wcislo WT, Taylor DR, Ishak HD, Dowd SE, Mueller UG (2012) Environment or kin: whence do bees obtain acidophilic bacteria? *Mol Ecol* 21:1754–1768
- Moran NA, Hansen AK, Powell JE, Sabree ZL (2012) Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. *PLoS One* 7:e36393
- Parks DH, Beiko RG (2010) Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* 26:715–721
- Rothberg JM, Hinz W, Rearick TM, Schultz J, Mileski W, Davey M, Bustillo J (2011) An integrated semiconductor device enabling non-optical genome sequencing. *Nature* 475:348–352
- Roulston TH, Cane JH (2000) Pollen nutritional content and digestibility for animals. *Plant Syst Evol* 222:187–209
- Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Weber CF (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75:7537–7541
- Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One* 6:e27310
- Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13:2498–2504
- Shipp JL, Whitfield GH, Papadopoulos AP (1994) Effectiveness of the bumble bee, *Bombus impatiens* Cr. (Hymenoptera: Apidae), as a pollinator of greenhouse sweet pepper. *Sci Hort* 57:29–39
- Singh B, Mitchison DA (1954) Bactericidal activity of streptomycin and isoniazid against tubercle bacilli. *Br Med J* 1:130–132
- Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM, Neal PR, Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored “rare biosphere”. *Proc Natl Acad Sci USA* 103:12115–12120
- Vojvodic S, Rehan SM, Anderson KE (2013) Microbial gut diversity of Africanized and European honey bee larval instars. *PLoS One* 8:e72106
- Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73:5261–5267
- Weinstein JN, Myers TG, O’Connor PM, Friend SH, Fornace AJ, Kohn KW, Paull KD (1997) An information-intensive approach to the molecular pharmacology of cancer. *Science* 275:343–349
- Whitfield CW, Behura SK, Berlocher SH, Clark AG, Johnston JS, Sheppard WS, Tsutsui ND (2006) Thrice out of Africa: ancient and recent expansions of the honey bee, *Apis mellifera*. *Science* 314:642–645