

# Genome Sequences of *Apibacter* spp., Gut Symbionts of Asian Honey Bees

Waldan K. Kwong<sup>1,\*</sup>, Margaret I. Steele, and Nancy A. Moran

Department of Integrative Biology, University of Texas at Austin

<sup>1</sup>Present address: Department of Botany, University of British Columbia, Vancouver, BC, Canada

\*Corresponding author: E-mail: waldankwong@gmail.com.

Accepted: April 3, 2018

**Data deposition:** Genome sequences have been deposited in GenBank under accession nos. PSZL00000000 (strain wkB309), PSZN00000000 (strain wkB180), and PSZM00000000 (strain wkB301).

## Abstract

Honey bees have distinct gut microbiomes consisting almost entirely of several host-specific bacterial species. We present the genomes of three strains of *Apibacter* spp., bacteria of the Bacteroidetes phylum that are endemic to Asian honey bee species (*Apis dorsata* and *Apis cerana*). The *Apibacter* strains have similar metabolic abilities to each other and to *Apibacter mensalis*, a species isolated from a bumble bee. They use microaerobic respiration and fermentation to catabolize a limited set of monosaccharides and dicarboxylic acids. All strains are capable of gliding motility and encode a type IX secretion system. Two strains and *A. mensalis* have type VI secretion systems, and all strains encode Rhs or VgrG proteins used in intercellular interactions. The characteristics of *Apibacter* spp. are consistent with adaptations to life in a gut environment; however, the factors responsible for host-specificity and mutualistic interactions remain to be uncovered.

**Key words:** gut microbiome, honey bee, bumble bee, strain variation, host specificity.

## Introduction

Honey bees are critical agricultural pollinators worldwide and have suffered from high rates of colony failure in recent years (Gallai et al. 2009). Worker honey bees harbor distinctive gut bacterial communities that typically consist of fewer than 10 member genera (Kwong and Moran 2016a). These bacteria are important for maintaining proper immune function (Emery et al. 2017; Kwong et al. 2017), gut physiology (Zheng et al. 2017), and nutrient processing (Lee et al. 2015; Kešnerová et al. 2017; Zheng et al. 2017) within bee hosts. The microbiome of the Western honey bee (*Apis mellifera*) has been the attention of many recent studies; however, little is known about the microbiota of its Asian relatives. Two species, the Eastern honey bee (*Apis cerana*) and the giant honey bee (*Apis dorsata*), are widespread throughout South and East Asia. These species have high local economic value as pollinators and as producers of honey and other hive products (Oldroyd and Wongsiri 2006). In particular, *Apis cerana* is kept domestically and accounts for a substantial proportion of the apiculture industry in some nations, including India and China.

Surveys of the gut microbiota of these bees show the presence of bacteria from the phylum Bacteroidetes as resident members (Ahn et al. 2012; Kwong et al. 2017); in contrast, bacteria of this lineage are rarely found in *Apis mellifera*. Strains have been isolated and classified as a novel genus, *Apibacter* (Kwong and Moran 2016b; Praet et al. 2016). In *Apis cerana* and *Apis dorsata*, >80% of adult worker bees are colonized with *Apibacter* (Kwong et al. 2017). *Apibacter* is also found in bumble bees, although with more sporadic occurrence than in the Asian honey bees (Koch and Schmid-Hempel 2011, 2012; Lim et al. 2015; Praet et al. 2016; Kwong et al. 2017). Different bee species appear to harbor different strains of *Apibacter*, suggesting that these bacteria are adapted to specific hosts (Kwong and Moran 2016b; Kwong et al. 2017). In bumble bees, *Apibacter* may be a beneficial symbiont, as it is associated with decreased infection by *Crithidia bombi*, a trypanosomatid gut parasite (Mockler et al. 2017).

Here, we present the genomes of three *Apibacter* strains from the honey bees *Apis cerana* and *Apis dorsata*. Together with a publically available *Apibacter mensalis* strain from a

**Table 1**Genome Statistics of *Apibacter* spp.

	<i>Apibacter mensalis</i> R-53146	<i>Apibacter</i> sp. wkb309	<i>Apibacter advertoris</i> wkb180	<i>Apibacter advertoris</i> wkb301
<b>Host</b>	Red-tailed bumble bee ( <i>Bombus lapidarius</i> )	Eastern honey bee ( <i>Apis cerana</i> )	Giant honey bee ( <i>Apis dorsata</i> )	Giant honey bee ( <i>Apis dorsata</i> )
<b>Collection site</b>	Ghent, Belgium	Genting, Malaysia	Singapore	Kuala Lumpur, Malaysia
<b>Genome size (bp)</b>	2,331,098	2,289,083	2,625,614	2,756,245
<b>G+C mol.</b>	30.5%	30.6%	28.9%	29.1%
<b>Plasmids (size)</b>	–	1 (25.4 kb)	–	–
<b>Proteins coding genes</b>	2,097	1,975	2,323	2,550
<b>tRNAs</b>	38	36	36	41
<b>Assembled contigs</b>	48	22	25	47
<b>Contig N50 (bp)</b>	158,456	340,669	165,551	244,991
<b>Coverage</b>	302×	130×	75×	100×
<b>Completeness<sup>a</sup></b>	97.95%	97.95%	99.49%	99.49%
<b>GenBank accession no.</b>	LIVM000000000	PSZL000000000	PSZN000000000	PSZM000000000

<sup>a</sup>Completeness estimates based on presence of 286 marker genes common to 419 strains of the phylum Bacteroidetes.

bumble bee, we conduct preliminary genomic analyses to reveal their encoded functions and potential role in the bee gut community.

## Materials and Methods

Cultivation and genomic sequencing of *Apibacter* was conducted as previously described (Kwong and Moran 2016b). Briefly, strains were grown on heart infusion agar (Hardy Diagnostics) supplemented with 5% defibrinated sheep blood, with incubation at 35°C in 5% CO<sub>2</sub>. DNA was extracted for paired-end sequencing on the Illumina MiSeq platform at the Genome Sequencing and Analysis Facility at the University of Texas at Austin (Kwong and Moran 2016b). In total, 2.6 million 300-bp Illumina MiSeq reads were acquired for strains wkb180, wkb301, and wkb309. Reads for strains wkb180 and wkb301 were assembled using MaSuRCA 3.2.2 (Zimin et al. 2013). Assembly of strain wkb309 was performed with Velvet 1.2.10 (Zerbino and Birney 2008) and CLC Genomics Workbench 5.5 (QIAGEN), and improved by mapping reads back onto assembled contigs using Bowtie 2 (Langmead and Salzberg 2012) and manual inspection. This multistep assembly for wkb309 was done to achieve the best possible genome assembly with the available short-read (Illumina) data. The wkb309 genome was selected as it had the highest quality assembly following initial assembly with MaSuRCA 3.2.2 and Velvet 1.2.10 (N50 of 260 kb); the laborious nature of this method precluded its use for all the *Apibacter* genomes. The genome of *Apibacter mensalis* R-53146 was retrieved from GenBank (accession no. LIVM000000000.1).

Genome completeness was estimated with CheckM v1.0.11 (Parks et al. 2015), using the taxonomic-specific workflow with the Bacteroidetes phylum taxon set. It should

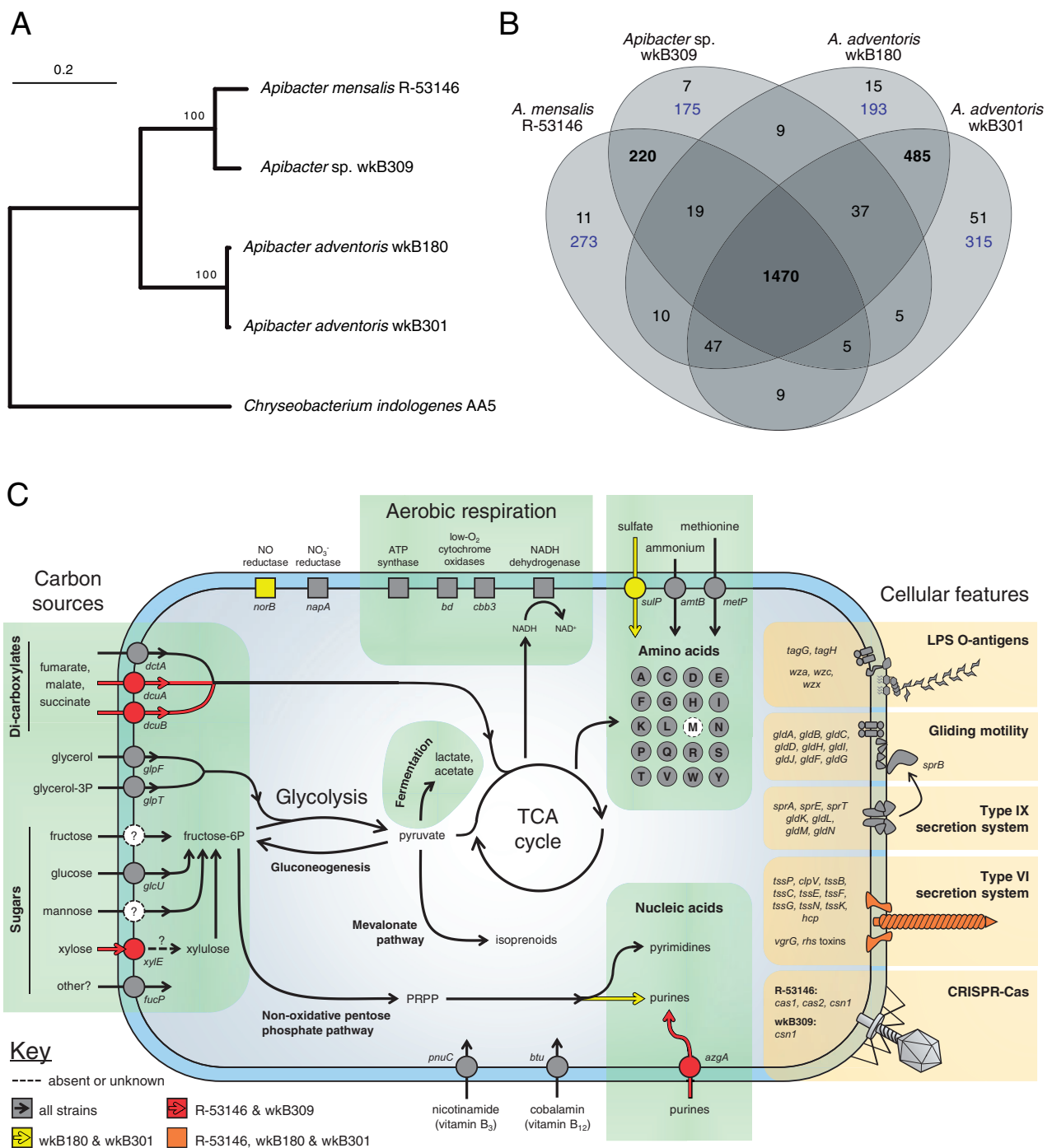
be noted that a lack of completeness by this metric may also be due to lineage-specific losses of marker genes within *Apibacter*, and not necessarily poor assembly. Gene content prediction and annotation for all four *Apibacter* genomes was done on the Rapid Annotation using Subsystem Technology (RAST) 2.0 platform (Overbeek et al. 2014). Orthologous gene clusters were identified and compared using OrthoVenn (Wang et al. 2015). Single-copy orthologs (1,132 proteins; 401,782 positions) were retrieved using OrthoFinder (Emms and Kelly 2015), aligned with MUSCLE (Edgar 2004), and concatenated to build a phylogenetic tree using the LG amino acid substitution model in RAXML v8 (Stamatakis 2014). Metabolic pathways were predicted with Pathway Tools 21.0 (Karp et al. 2002) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) mapper function on RAST 2.0. Pathway holes were manually inspected to verify completeness. CRISPR arrays were predicted with CRISPRFinder (Grissa et al. 2007). The NCBI Batch CD-Search tool (Marchler-Bauer and Bryant 2004) was used to identify conserved domains in putative secreted proteins, and RAXML v8 was used to build the TssB phylogeny.

Genome sequences have been deposited in GenBank under accession nos. PSZL000000000 (strain wkb309), PSZN000000000 (strain wkb180), and PSZM000000000 (strain wkb301).

## Results and Discussion

### Genome Content

*Apibacter* spp. genomes average 2.5 Mb in size, and have ~30% G + C content (table 1). The two strains from *Apis dorsata* (wkb180 and wkb301) have larger genomes and lower G + C content than the strains from *Apis cerana* (wkb309) or the bumble bee (R-53146). These differences



**FIG. 1.**—(A) Phylogenetic relationships of sequenced *Apibacter* strains, based on 1,132 single-copy orthologous proteins. Tree was built using the maximum-likelihood algorithm. Bootstrap support (1,000 replicates) indicated at nodes; bar, substitutions per site. *Chryseobacterium indologenes* was used as the outgroup. (B) Number of gene clusters (orthologs or paralogs) shared between *Apibacter* strains. Each cluster contains at least 2 genes. Blue text indicates singleton genes unique to a single genome and thus not belonging to orthologous or paralogous clusters. (C) Predicted metabolism and cellular features of *Apibacter* spp. Pathways and components colored according to presence across the sequenced strains. Abbreviations: PRPP, 5-phospho- $\alpha$ -D-ribose 1-pyrophosphate; LPS, lipopolysaccharide; CRISPR, clustered regularly interspaced short palindromic repeats.

in genome structure are reflected in their evolutionary relationships, as shown by a whole genome phylogeny (fig. 1A). Strains wkB180 and wkB301 are closely related (98.4%

ortholog amino acid identity). Strains R-53146 and wkB309 group together, but are more dissimilar (88.3% identity). A large majority of genes are shared by all strains (fig. 1B). Three

rRNA gene operons are predicted in strain wkB309, the best assembled genome, based on read coverage of contigs. Strain wkB309 carries a single 25.4-kb plasmid, as inferred from the presence of plasmid partitioning and mobilization genes, matching contig end sequences, and higher read coverage for that contig.

### Metabolism

Based on their genomic content, all strains possess the same core metabolic functionalities (fig. 1C), which include the Embden–Meyerhof–Parnas glycolysis pathway, the non-oxidative branch of the pentose phosphate pathway, the ability for gluconeogenesis, and a complete tricarboxylic acid (TCA) cycle. All strains likely use oxidative phosphorylation for energy production, as they encode NADH dehydrogenase and two cytochrome oxidases (*bd* and *cbb<sub>3</sub>*). Genes are present for the de novo synthesis of all proteinogenic amino acids except methionine, for which there is an encoded transporter. Based on their genomic content, pyrimidines can be synthesized de novo; however, only strains wkB301 and wkB180 can produce purines. The genomes also encode pathways for de novo production of several vitamins (pantothenate, tetrahydrofolate, vitamin B6, and likely riboflavin). However, other vitamins or their precursors, including biotin, cobalamin, nicotinamide, and thiamine, need to be imported, as inferred from the absence of genes encoding the necessary biosynthetic enzymes.

*Apibacter* spp. likely subsist on a mix of simple sugars and dicarboxylic acids, based on their gene repertoires. Pathways for glucose, fructose, and mannose degradation are present, which is consistent with in vitro substrate utilization assays (Kwong and Moran 2016b; Praet et al. 2016). Phosphotransferase systems for carbohydrate import are lacking, which is also the case in some other Bacteroidetes species (Barabote and Saier 2005). Instead, saccharides are likely imported by alternative, or unidentified transporters (fig. 1C). The bee gut is a low oxygen environment (Zheng et al. 2017), and this is reflected in how *Apibacter* is predicted to metabolize carbon sources. While the presence of the TCA cycle and NADH dehydrogenase implies a preference for aerobic respiration (Morris and Schmidt 2013), enzymes with activity in low-O<sub>2</sub> (cytochrome *bd*, cytochrome *cbb<sub>3</sub>*) or anaerobic (*dcuA*, *dcuB*, *glpT*) conditions are also present. If not oxidized in the TCA cycle, substrates broken down by glycolysis are predicted to be fermented to lactate or acetate. Dicarboxylic acids are likely utilized directly by the TCA cycle, in a manner similar to that in *Snodgrassella alvi*, another bee gut symbiont (Kwong et al. 2014).

### Cellular Features

*Apibacter* spp. lack genes encoding flagellum or pilus apparatuses for motility. However, they have a full complement of

genes for gliding motility that is typical of Flavobacteriaceae (McBride and Zhu 2013). Strain wkB309 has been observed to glide, and it is likely that all other strains also have this capability (Kwong and Moran 2016b). The type IX secretion system, which is also characteristic of many Bacteroidetes and is necessary for gliding motility (McBride and Nakane 2015; Lasica et al. 2017), is present in all strains.

Bacteria within the bee gut may engage in competitive interactions with each other, using type VI secretion systems (T6SS) to inject toxins into neighboring cells (Steele et al. 2017). Strains R-53146, wkB180, and wkB301 each encode a single T6SS that is related to other Bacteroidetes T6SSs (supplementary fig. S1, Supplementary Material online). All strains encode potential T6SS-secreted proteins: R-53146, wkB180, wkB301, and wkB309 encode 7, 9, 10, and 0 *vgrG*-like genes, respectively, as well as 7, 9, 41, and 4 Rhs-domain containing genes. *S. alvi* and *Gilliamella apicola*, two other Gram-negative members of the bee gut microbial community, also harbor large numbers of Rhs toxin genes (Steele et al. 2017). Further analysis of the genomic organization and predicted functionalities of *Apibacter* T6SSs and putative effectors is presented in supplementary results, Supplementary Material online.

Only *A. mensalis* R-53146 carries a potentially complete CRISPR-Cas system, a bacterial immune mechanism against viruses and foreign DNA (Barrangou et al. 2007). Seven CRISPR spacers were predicted in R-53146, three spacers were in strain wkB309 (which has only the *csn1* gene), while wkB180 and wkB301 each had only a single predicted spacer. Another mechanism to limit the transfer and integration of foreign DNA is through restriction modification (R–M) systems (Wilson and Murray 1991). Each strain has at least one type I R–M system. R–M systems might be shared by members of a coevolving community (Furuta and Kobayashi 2011; Kwong et al. 2014). One R–M system in strain R-53146 shares 97% protein sequence identity with an R–M system in *G. apicola*, suggesting recent transfer of genes between coresident bee gut bacterial species.

### Conclusions

*Apibacter* spp. are microaerobic members of the bee gut community that subsist on a limited set of monosaccharides, dicarboxylates, and glycerol. They are unlikely to participate in the digestion of complex polysaccharides, which stands in contrast to some other bee gut symbionts (e.g., *Gilliamella* and *Bifidobacterium*) as well as gut Bacteroidetes of other animals (e.g., *Bacteroides* and *Prevotella*) (Flint et al. 2008). Based on its physiology, we predict that *Apibacter* colonizes the gut wall, where O<sub>2</sub> concentration is the highest (Zheng et al. 2017). The encoded type IX secretion system and gliding motility apparatus may help it form biofilms (Kita et al. 2016), perhaps in a similar fashion to *S. alvi*, which is also aerobic and produces biofilms along the gut wall (Martinson et al. 2012;

Powell et al. 2016). *Apibacter* co-occurs with *S. alvi* in the Eastern honey bee *Apis cerana*, while the giant honey bee *Apis dorsata* generally lacks *S. alvi* (Kwong et al. 2017). Although the data are sparse, there is possibly an inverse relationship between *Apibacter* and *S. alvi* abundance in bumble bees, perhaps indicating competitive exclusion within a limited ecological niche (Koch and Schmid-Hempel 2012; Lim et al. 2015; Mockler et al. 2017).

The bee gut microbiome is a dynamic community with multiple interacting members. This is reflected in differences among *Apibacter* strains in their complements of accessory elements, including T6SSs, Rhs toxins, CRISPR-Cas systems, and R-M systems. Unlike the core metabolic pathways, these elements probably evolve quickly in response to changes in microbiome composition (such as the presence of bacterial competitors or viruses). Most other noncore genes (i.e., those not shared by all strains, fig. 1B) have no known function; those annotated as “hypothetical” constitute 30% of total predicted genes. Closer scrutiny of these strain-specific genes is needed to help uncover the factors responsible for *Apibacter*'s ability to colonize and interact with specific bee hosts.

## Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

## Authors' Contributions

W.K.K. and N.A.M. designed the study. W.K.K. and M.I.S. analyzed and interpreted the data. W.K.K. drafted the manuscript. All authors read, edited, and approved the final manuscript.

## Acknowledgments

We thank Hao Zheng for helpful discussions and suggestions. This work was supported by the National Institutes of Health award 1R01-GM108477-01 to N.A.M.

## Literature Cited

- Ahn JH, et al. 2012. Pyrosequencing analysis of the bacterial communities in the guts of honey bees *Apis cerana* and *Apis mellifera* in Korea. *J Microbiol.* 50(5):735–745.
- Barabote RD, Saier MH Jr. 2005. Comparative genomic analyses of the bacterial phosphotransferase system. *Microbiol Mol Biol Rev.* 69(4):608–634.
- Barrangou R, et al. 2007. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315(5819):1709–1712.
- Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32(5):1792–1797.
- Emery O, Schmidt K, Engel P. 2017. Stimulation of the immune system by the gut symbiont *Frischella perrara* in the honey bee (*Apis mellifera*). *Mol Ecol.* 26:2576–2590.
- Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* 16:157.
- Flint HJ, Bayer EA, Rincon MT, Lamed R, White BA. 2008. Polysaccharide utilization by gut bacteria: potential for new insights from genomic analysis. *Nat Rev Microbiol.* 6(2):121–131.
- Furuta Y, Kobayashi I. 2011. Restriction-modification systems as mobile epigenetic elements. Austin (TX): Landes Bioscience, 2000–2013.
- Gallai N, Salles JM, Settele J, Vaissiere BE. 2009. Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecol Econ.* 68(3):810–821.
- Grissa I, Vergnaud G, Pourcel C. 2007. CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35(Web Server issue):W52–W57.
- Karp PD, Paley S, Romero P. 2002. The Pathway Tools software. *Bioinformatics* 18(Suppl 1):S225–S232.
- Kešnerová L, et al. 2017. Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol.* 15(12):e2003467.
- Kita D, et al. 2016. Involvement of the type IX secretion system in *Capnocytophaga ochracea* gliding motility and biofilm formation. *Appl Environ Microbiol.* 82(6):1756–1766.
- Koch H, Schmid-Hempel P. 2011. Bacterial communities in central European bumblebees: low diversity and high specificity. *Microb Ecol.* 62(1):121–133.
- Koch H, Schmid-Hempel P. 2012. Gut microbiota instead of host genotype drive the specificity in the interaction of a natural host-parasite system. *Ecol Lett.* 15(10):1095–1103.
- Kwong WK, Engel P, Koch H, Moran NA. 2014. Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc Natl Acad Sci U S A.* 111(31):11509–11514.
- Kwong WK, et al. 2017. Dynamic microbiome evolution in social bees. *Sci Adv.* 3(3):e1600513.
- Kwong WK, Moran NA. 2016a. Gut microbial communities of social bees. *Nat Rev Microbiol.* 14(6):374–384.
- Kwong WK, Moran NA. 2016b. *Apibacter adventoris* gen. nov., sp. nov., a member of phylum *Bacteroidetes* isolated from honey bee guts. *Int J Syst Evol Microbiol.* 66(3):1323–1329.
- Langmead B, Salzberg S. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–359.
- Lasica AM, Ksiazek M, Madej M, Potempa J. 2017. The type IX secretion system (T9SS): highlights and recent insights into its structure and function. *Front Cell Infect Microbiol.* 7:215.
- Lee FJ, Rusch DB, Stewart FJ, Mattila HR, Newton IL. 2015. Saccharide breakdown and fermentation by the honey bee gut microbiome. *Environ Microbiol.* 17(3):796–815.
- Lim HC, Chu C-C, Seufferheld MJ, Cameron SA. 2015. Deep sequencing and ecological characterization of gut microbial communities of diverse bumble bee species. *PLoS One* 10(3):e0118566.
- Marchler-Bauer A, Bryant SH. 2004. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32(Web Server issue):W327–W331.
- Martinson VG, Moy J, Moran NA. 2012. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl Environ Microbiol.* 78(8):2830–2840.
- McBride MJ, Nakane D. 2015. *Flavobacterium* gliding motility and the type IX secretion system. *Curr Opin Microbiol.* 28:72–77.
- McBride MJ, Zhu Y. 2013. Gliding motility and Por secretion system genes are widespread among members of the phylum bacteroidetes. *J Bacteriol.* 195(2):270–278.
- Mockler BK, Kwong WK, Moran NA, Koch H. 2017. Microbiome structure influences infection by the parasite *Crithidia bombi* in bumble bees. *Appl Environ Microbiol.* 84(7):e02335–17.

- Morris RL, Schmidt TM. 2013. Shallow breathing: bacterial life at low O<sub>2</sub>. *Nat Rev Microbiol.* 11(3):205–212.
- Oldroyd BP, Wongsiri S. 2006. Asian honey bees: biology, conservation and human interactions. Cambridge: Harvard University Press.
- Overbeek R, et al. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42(Database issue):D206–D214.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25(7):1043–1055.
- Powell JE, Leonard SP, Kwong WK, Engel P, Moran NA. 2016. Genome-wide screen identifies host colonization determinants in a bacterial gut symbiont. *Proc Natl Acad Sci U S A.* 113(48):13887–13892.
- Praet J, et al. 2016. *Apibacter mensalis* sp. nov.: a rare member of the bumblebee gut microbiota. *Int J Syst Evol Microbiol.* 66(4): 1645–1651.
- Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
- Steele MI, Kwong WK, Whiteley M, Moran NA. 2017. Diversification of type VI secretion system toxins reveals ancient antagonism among bee gut microbes. *MBio* 8(6):e01630-17.
- Wang Y, Coleman-Derr D, Chen G, Gu YQ. 2015. OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 43(W1):W78–W84.
- Wilson GG, Murray NE. 1991. Restriction and modification systems. *Annu Rev Genet.* 25:585–627.
- Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18(5):821–829.
- Zheng H, Powell JE, Steele MI, Dietrich C, Moran NA. 2017. Honeybee gut microbiota promotes host weight gain via bacterial metabolism and hormonal signaling. *Proc Natl Acad Sci U S A.* 114(18):4775–4780.
- Zimin AV, et al. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29(21):2669–2677.

Associate editor: Tal Dagan