

# Reclassification of seven honey bee symbiont strains as *Bombella apis*

Eric A. Smith<sup>1†</sup>, Kirk E. Anderson<sup>2</sup>, Vanessa Corby-Harris<sup>2</sup>, Quinn S. McFrederick<sup>3</sup>, Audrey J. Parish<sup>1</sup>, Danny W. Rice<sup>1</sup> and Irene L. G. Newton<sup>1\*</sup>

## Abstract

Honey bees are important pollinators of many major crops and add billions of dollars annually to the US economy through their services. Recent declines in the health of the honey bee have startled researchers and lay people alike as honey bees are agriculture's most important pollinator. One factor that may influence colony health is the microbial community. Although honey bee worker guts have a characteristic community of bee-specific microbes, the honey bee queen digestive tracts are colonized predominantly by a single acetic acid bacterium tentatively named '*Parasaccharibacter apium*'. This bacterium is related to flower-associated microbes such as *Saccharibacter floricola*, and initial phylogenetic analyses placed it as sister to these environmental bacteria. We used a combination of phylogenetic and sequence identity methods to better resolve evolutionary relationships among '*P. apium*', strains in the genus *Saccharibacter*, and strains in the closely related genus *Bombella*. Interestingly, measures of genome-wide average nucleotide identity and aligned fraction, coupled with phylogenetic placement, indicate that many strains labelled as '*P. apium*' and *Saccharibacter* species are all the same species as *Bombella apis*. We propose reclassifying these strains as *Bombella apis* and outline the data supporting that classification below.

## INTRODUCTION

The honey bee (*Apis mellifera*) is extremely important economically because of the pollination services it provides to numerous agricultural crops. As a result, there is increasing interest in determining how the microbiome supports and influences bee function. While a honey bee colony is made up of individuals with distinct roles, or castes, the majority of studies on bee microbiomes have focused on workers. The microbial community of worker bees consists of eight to ten core bacterial phylotypes [1, 2]. The characterization of these groups led to speculation about their role in honey bee health and whether or not they provision nutrients [1] or assist in the breakdown of plant-derived carbohydrates [3, 4], as is the case in other insect-microbe interactions [5, 6]. There has also been speculation as to the role of the microbiome in resistance to pathogens, as microbial communities have been shown to protect the bumble bee (*Bombus terrestris*) from the parasite *Crithidia bombi* [7].

Honey bee-associated microbes interact with each other in diverse ways both *in vitro* and *in vivo*, suggesting that they may interact syntrophically within workers [8, 9]. While these studies focused on honey bee workers are intriguing, it is surprising that only recently was the microbiome of queen bees throughout development characterized [10].

Interestingly, the microbial community associated with queen bees is vastly different than that of workers and comprises a large percentage of acetic acid bacteria, a group of bacteria present only at very small percentages in workers. One of the primary bacteria that differentiate queens from workers was tentatively named '*Parasaccharibacter apium*' [11]. This bacterium is in the family *Acetobacteraceae* and occupies defined niches within the hive, including: queen guts, nurse hypopharyngeal glands, nurse crops and royal jelly, and is only rarely found in high abundance outside of these areas [12, 13]. Evidence suggests that it might play a role in protecting developing larvae and worker bees from

**Author affiliations:** <sup>1</sup>Department of Biology, Indiana University, Bloomington, Indiana, USA; <sup>2</sup>Carl Hayden Bee Research Center, USDA-ARS, Tucson, Arizona, USA; <sup>3</sup>Department of Entomology, University of California, Riverside, California, USA.

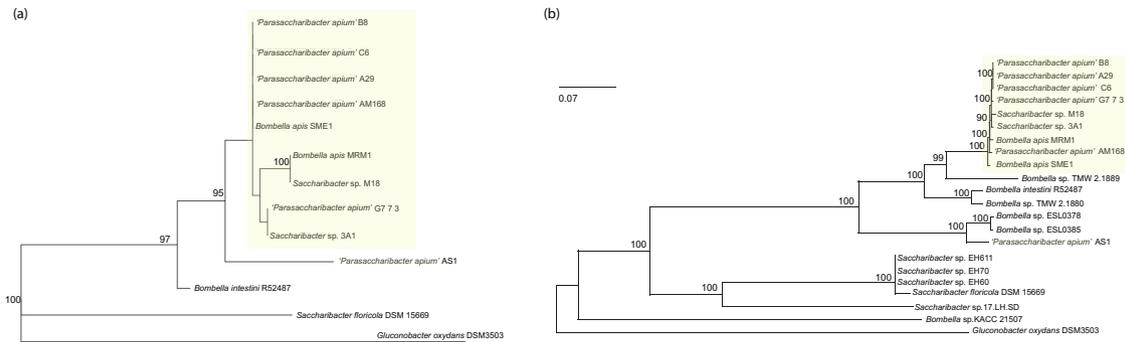
**\*Correspondence:** Irene L. G. Newton, irnewton@indiana.edu

**Keywords:** average nucleotide identity; bacterial classification; evolution; honey bee; phylogenetics.

**†Present address:** Diversigen, INC, New Brighton, MN, USA.

This Whole Genome Shotgun project has been deposited at DDBJ/ENA/GenBank under the accession JAGJTM000000000. The version described in this paper is version JAGJTM010000000.

Three supplementary tables are available with the online version of this article.



**Fig. 1.** Maximum-likelihood 16S rRNA gene phylogeny (a) and core ortholog phylogeny (b). Numbers at nodes represent bootstrap support from 1000 bootstrap pseudoreplicates; only support greater than 80 shown. The 'clade of interest' is highlighted in yellow.

fungal pathogens, such as *Nosema* [11, 14, 15]. Given that this bacterium makes up a large proportion of the queen gut microbiome, it is possible that it plays an important role in queen nutrition, protection from pathogens, and possibly modulating queen fertility, fecundity, and longevity [16].

We sought to determine the evolutionary relationships between strains of '*P. apium*', strains of the closely related genus *Saccharibacter*, and characterized sequences from strains of the recently named genus *Bombella* [17], using previously published data (Table S1, available in the online version of this article, for information on all genomes used in this study). Using a combination of phylogenetic and sequence similarity methods, we found that many genomes labelled as '*P. apium*' or *Saccharibacter* species are actually the same species as strains of the previously described species *Bombella apis*. We reclassify these strains and outline the data supporting this reclassification below.

## METHODS

### 16S rRNA phylogeny

To determine relatedness of the strains classified as '*P. apium*' and *Saccharibacter* species compared to existing *Bombella* and *Saccharibacter* species, we compared 16S rRNA gene sequences of these genomes to one another. We first downloaded all 16S rRNA gene sequences from the NCBI that were labelled as *Bombella*, '*Parasaccharibacter*' or *Saccharibacter* (Table S1). We also included *Gluconobacter oxydans* DSM3503 for use as an outgroup. All sequences were aligned using the SINA aligner [18]; parameters used were set using the --auto option. A maximum-likelihood phylogeny was reconstructed using RAxML with the GTRGAMMA substitution model and 1000 bootstrap replicates (version 8.2.11 [19]). The final tree was visualized using ggtree in R (version 3.12; <https://bioconductor.org/packages/release/bioc/html/ggtree.html>).

### Sequencing and assembly of *B. apis* MRM1

*B. apis* MRM1 was requested of the Riken JCM repository (31623<sup>T</sup>) and the microbe streaked to isolation on de

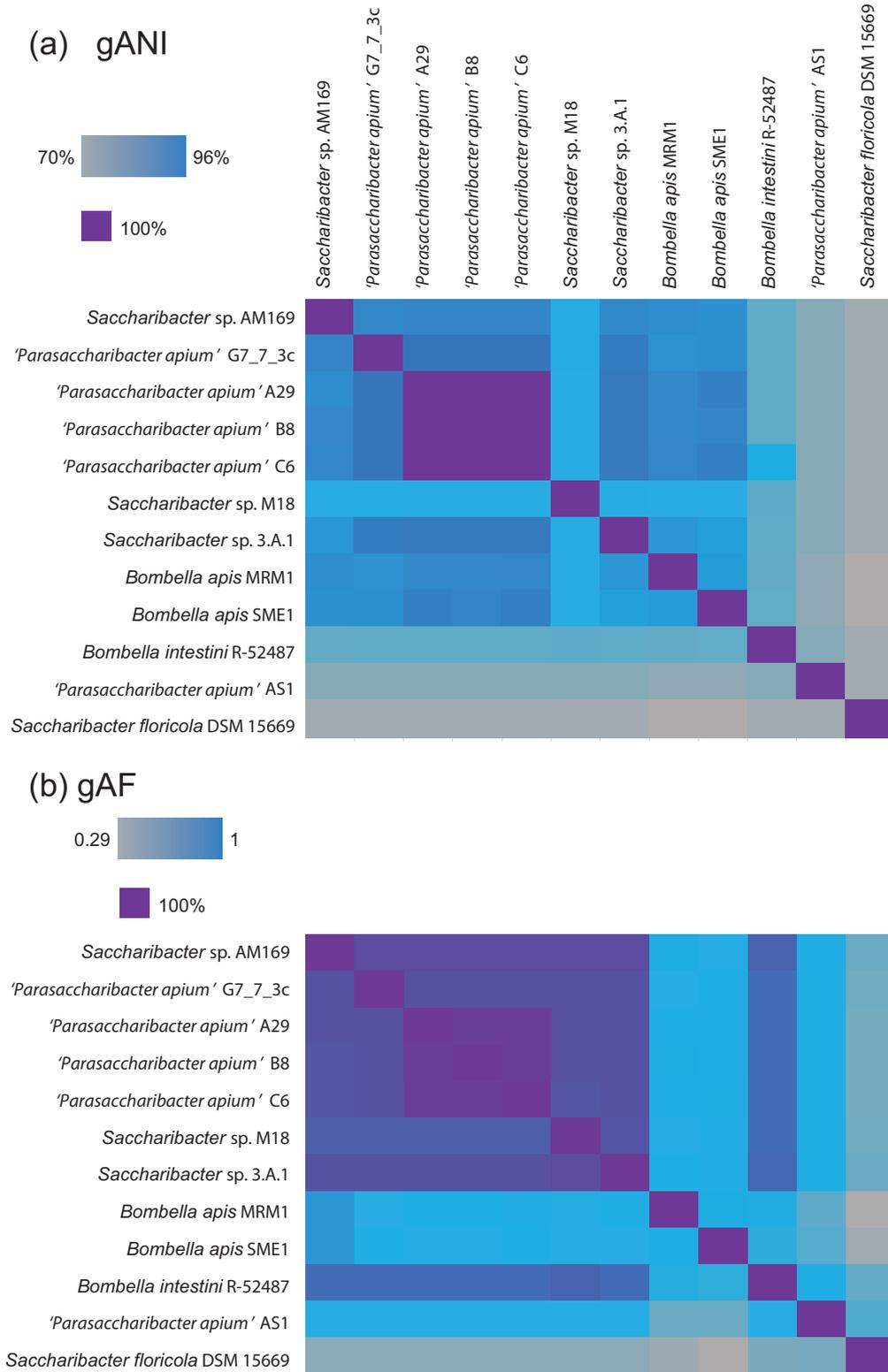
Man–Rogosa–Sharpe (MRS) agar at 34 °C. A single isolated colony was used to inoculate MRS broth and grown overnight at 34 °C and 250 r.p.m. before harvesting cells for DNA extraction. DNA was isolated using a standard beat beating method followed by phenol–chloroform extraction. Libraries were generated using an NEBNext Ultra II DNA sequencing library kit and sequenced on a NextSeq instrument using 300 cycles and a paired-end sequencing strategy. Reads were binned by their barcode and Masurca was used to assemble the reads using the guided assembly approach with the *B. apis* A29 genome as a reference. The resulting assembly was quality controlled using QUAST and CheckM and comprises six contigs, with an N50 of 41 1263 and a total size of 2048899 bp.

### Core ortholog phylogeny

To determine genome-wide phylogenetic relationships between strains, we clustered genes from all genomes in Table S1 into orthologous gene clusters using reciprocal blast. Whole genome and amino acid sequences were downloaded from the NCBI (see accessions in Table S1) and clustering was performed using default parameters and all versus all BLASTp with default options. An orthologous group of genes was defined by complete linkage such that all members of the group had to be the reciprocal best hit of all other members of the group. Thus, a particular strain could have at most one gene per ortholog group. We reconstructed a phylogeny using concatenated amino acid alignments of all single-copy orthologs. The amino acid sequences were aligned using the MAFFT L-INS-I algorithm (version 7.310, [20]) and alignments were then concatenated and used to reconstruct a maximum-likelihood phylogeny using RAxML with the GTRGAMMA substitution model and 1000 bootstrap replicates (version 8.2.11, [19]). The final tree was visualized using ggtree in R (version 3.12; <https://bioconductor.org/packages/release/bioc/html/ggtree.html>).

### Calculation of genomic similarity

To determine relatedness and proper species assignment, we calculated genome-wide average nucleotide identity (gANI)



**Fig. 2.** Pairwise genome-wide average nucleotide identity (A) and aligned fraction (B) of all strains used in this study. Accession numbers are given with current names in parentheses. Colors in each cell scale with the metric. Genome-wide average nucleotide identity scales from 0-100% while aligned fraction scales from 0-1. Both metrics can vary depending on which genome is considered the reference for alignment, and we made the calculations in both directions. Genomes listed on the horizontal axis were considered as the reference for these calculations.

**Table 1.** GenBank accession number, current name, isolation source, genome size, G+C content, and proposed new names of the strains used in this study

GenBank accession	Current name	Isolation source	References	Genome size (Mbp)	G+C content (mol%)	Proposed new name
GCA_002079945.1	' <i>Parasaccharibacter apium</i> ' G7_7_3 c	<i>Apis mellifera</i> hindgut	Corby-Harris and Anderson 2018 [23]	2.01	59.42	<i>Bombella apis</i> G7_7_3 c
GCA_002917995.1	' <i>Parasaccharibacter apium</i> ' A29	<i>Apis mellifera</i> larva	Corby-Harris and Anderson 2018 [23]	2.01	59.39	<i>Bombella apis</i> A29
GCA_002917945.1	' <i>Parasaccharibacter apium</i> ' B8	<i>Apis mellifera</i> larva	Corby-Harris and Anderson 2018 [23]	2.01	59.38	<i>Bombella apis</i> B8
GCA_002917985.1	' <i>Parasaccharibacter apium</i> ' C6	<i>Apis mellifera</i> larva	Corby-Harris and Anderson 2018 [23]	2.01	59.38	<i>Bombella apis</i> C6
GCA_000723565.1	' <i>Parasaccharibacter apium</i> ' AM168	<i>Apis mellifera</i> stomach	Chouaia et al. 2014 [24]	1.98	59.32	<i>Bombella apis</i> AM168
GCA_002150105.1	<i>Saccharibacter</i> sp. M18	<i>Apis mellifera</i> stomach	Veress et al. 2017 [25]	2.01	59.35	<i>Bombella apis</i> M18
GCA_002150125.1	<i>Saccharibacter</i> sp. 3.A.1	Honey	Veress et al. 2017 [25]	2.01	59.41	<i>Bombella apis</i> 3.A.1
GCA_009362775.1	<i>Bombella apis</i> SME1	Honey	Smith et al. 2019 [26]	2.09	59.56	NA
PRJNA719376	<i>Bombella apis</i> MRM1	<i>Apis mellifera</i> midgut	Yun et al., 2017 [22]	2.01	59.39	NA

and aligned fraction (AF) for each pairwise comparison using ANIcalculator [21]. Predicted transcript sequences for each pairwise comparison were passed to the software, which output gANI and AF in each direction for the pairwise comparison. As gANI and AF can vary depending on the direction of comparison due to differences in genome lengths, we report results for comparisons in both directions.

## RESULTS

### 16S rRNA gene phylogeny

To determine phylogenetic relationships of all genomes labelled '*Parasaccharibacter*', *Saccharibacter* and *Bombella* on the NCBI, we reconstructed a maximum-likelihood phylogeny using 16S rRNA sequences (Fig. 1a). This phylogeny, based on the 16S rRNA gene sequences from the following genomes, suggests that the strains are all very closely related and represent a single, distinct clade: *Saccharibacter* sp. AM169, '*P. apium*' G7\_7\_3 c, '*P. apium*' A29, '*P. apium*' B8, '*P. apium*' C6, *Saccharibacter* sp. M18, *Saccharibacter* sp. 3.A.1, *B. apis* MRM1, *B. apis* SME1. For simplicity, this clade, highlighted in yellow in Fig. 1, will hereafter be referred to as the 'clade of interest'. '*P. apium*' AS1 is not included in the clade of interest for reasons outlined below. The relationships between strains within the clade of interest are difficult to determine – as is evidenced by the low bootstrap support for nodes within the clade – owing to the high degree of similarity between the 16S rRNA gene sequences (Table S2). However, the phylogeny clearly places these accessions with *B. apis*, and sister to '*P. apium*' AS1. It is interesting to note that, although the sequences all show similar degrees of divergence from *B. intestini* and *B. apis*, bootstrap support is quite high for the separation of *B. intestini* from the clade of interest (Fig. 1a).

### Core ortholog phylogeny

We used reciprocal BLAST to define core orthologs using the '*Parasaccharibacter*', *Saccharibacter* and *Bombella* genomes listed in File S1; *Gluconobacter oxydans* DSM3503 was used as an outgroup. In total, an alignment of 2270622 nucleotide positions was defined, with nucleotide positions represented by four or more taxa. This alignment was then used to reconstruct a core ortholog phylogeny (Fig. 1b) to better resolve phylogenetic relationships.

This core orthology phylogeny largely agrees with our 16S rRNA gene phylogeny with one exception: the relationships of *B. intestini* and '*P. apium*' AS1 have switched such that *B. intestini* is now sister to the clade of interest. Bootstrap support for the nodes in this tree are much higher than in the 16S rRNA gene tree, including within the clade of interest. Importantly, bootstrap support among the nodes making up the main backbone of the tree are all 100, indicating very high support for the basal nodes (Fig. 1b).

### Calculation of genomic similarity

Given the discrepancy between nomenclature and the phylogeny (the placement of '*Parasaccharibacter*' and *Saccharibacter* species within the *Bombella* group), and considering the short branch lengths within the clade of interest, we calculated gANI and AF to clarify species relationships (Fig. 2). The general criteria for two genomes to be considered the same species are AF>0.6 and gANI >96.5 [21]. Using these criteria, all genomes within the clade of interest should be considered the same species, and distinct from *B. intestini* and '*P. apium*' AS1 (Table S3), though part of the genus *Bombella*. Both the high degree of support for placement of *B. apis* within the clade of interest and the

gAF and gANI results strongly suggests that the genomes within the clade of interest belong to the same species and are all *B. apis* strains.

## DISCUSSION

Here, we used a combination of phylogenetic analysis (16S rRNA gene and core orthologs) and calculations of AF and gANI to determine relationships among symbionts of honey bees, bumble bees and environmental bacteria. The phylogenetic data largely agree with each other, and with the AF and gANI delimitations of species. The combination of these data indicate that the genomic accessions GCA\_000723565.1, GCA\_002079945.1, GCA\_002917995.1, GCA\_002917945.1, GCA\_002917985.1, GCA\_002150105.1 and GCA\_002150125.1 are all the same species, despite being named in NCBI as various strains of '*P. apium*' and *Saccharibacter* species. Given that the species name '*P. apium*' has been effectively but not validly published (i.e. in accordance with IJSEM standards; [14]), while *B. apis* has been validly published [22], we propose renaming the above accession numbers to reflect the proper genus and species assignment while maintaining their current strain designations (see Table 1 for full genus, species, and strain designations).

### Funding information

E.A.S. is supported by Agriculture and Food Research Initiative - Education and Workforce Development project accession no. 1019114 from the USDA National Institute of Food and Agriculture.

### Acknowledgements

The authors thank Amelia R.I. Lindsey, Delaney L. Miller and Audrey J. Parish for helpful feedback on drafts of this manuscript.

### Author contributions

Conceptualization: E.A.S. and I.L.G.N.; sequencing and assembly: A.J.P.; formal analysis: E.A.S. and D.W.R.; writing – original draft: E.A.S.; writing – review and editing: E.A.S., K.E.A., V.C.H., Q.S.M. and I.L.G.N.

### Conflicts of interest

The authors declare that there are no conflicts of interest.

### References

- Moran NA. Genomics of the honey bee microbiome. *Curr Opin Insect Sci* 2015;10:22–28.
- Moran NA, Hansen AK, Powell JE, Sabree ZL. Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. *PLoS One* 2012;7:e36393.
- Engel P, Martinson VG, Moran NA. Functional diversity within the simple gut microbiota of the honey bee. *Proc Natl Acad Sci U S A* 2012;109:11002–11007.
- Kesnerova L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, et al. Disentangling metabolic functions of bacteria in the honey bee gut. *PLoS Biol* 2017;15:e2003467.
- Akman Gunduz E, Douglas AE. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proc Biol Sci* 2009;276:987–991.
- Douglas AE. Microbial brokers of insect-plant interactions revisited. *J Chem Ecol* 2013;39:952–961.
- Koch H, Schmid-Hempel P. Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc Natl Acad Sci U S A* 2011;108:19288–19292.
- Martinson VG, Moy J, Moran NA. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl Environ Microbiol* 2012;78:2830–2840.
- Rokop ZP, Horton MA, Newton ILG. Interactions between cooccurring lactic acid bacteria in honey bee hives. *Appl Environ Microbiol* 2015;81:7261–7270.
- Tarpy DR, Mattila HR, Newton IL. Development of the honey bee gut microbiome throughout the queen-rearing process. *Appl Environ Microbiol* 2015;81:3182–3191.
- Corby-Harris V, Snyder LA, Schwan MR, Maes P, McFrederick QS, et al. Origin and effect of Alpha 2.2 Acetobacteraceae in honey bee larvae and description of *Parasaccharibacter apium* gen. nov., sp. nov. *Appl Environ Microbiol* 2014;80:7460–7472.
- Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, et al. 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* 2013;8:e83125.
- Vojvodic S, Rehan SM, Anderson KE. Microbial gut diversity of Africanized and European honey bee larval instars. *PLoS One* 2013;8:e72106.
- Corby-Harris V, Snyder L, Meador CA, Naldo R, Mott B, et al. *Parasaccharibacter apium*, gen. nov., sp. nov., improves honey bee (Hymenoptera: Apidae) resistance to *Nosema*. *J Econ Entomol* 2016;109:537–543.
- Miller DL, Smith EA, Newton ILG. A bacterial symbiont protects honey bees from fungal disease. *bioRxiv* 2020.
- Anderson KE, Ricigliano VA, Mott BM, Copeland DC, Floyd AS, et al. The queen's gut refines with age: longevity phenotypes in a social insect model. *Microbiome* 2018;6:108.
- Li L, Praet J, Borremans W, Nunes OC, Manaia CM, et al. *Bombella intestini* gen. nov., sp. nov., an acetic acid bacterium isolated from bumble bee crop. *Int J Syst Evol Microbiol* 2015;65:267–273.
- Pruesse E, Peplies J, Glöckner FO. SINA: accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;28:1823–1829.
- Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 2006;22:2688–2690.
- Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast fourier transform. *Nucleic Acids Res* 2002;30:3059–3066.
- Varghese NJ, Mukherjee S, Ivanova N, Konstantinidis KT, Mavrommatis K, et al. Microbial species delineation using whole genome sequences. *Nucleic Acids Res* 2015;43:6761–6771.
- Yun J-H, Lee J-Y, Hyun D-W, Jung M-J, Bae J-W, et al. *Bombella apis* sp. nov., an acetic acid bacterium isolated from the midgut of a honey bee. *Int J Syst Evol Microbiol* 2017;67:2184–2188.
- Corby-Harris V, Anderson KE. Draft genome sequences of four parasaccharibacter apium strains isolated from honey bees. *Genome Announc* 2018;6:e00165.
- Chouaia B, Gaiarsa S, Crotti E, Comandatore F, Degli Esposti M, et al. Acetic acid bacteria genomes reveal functional traits for adaptation to life in insect guts. *Genome Biol Evol* 2014;6:912–920.
- Veress A, Wilk T, Kiss J, Olasz F, Papp PP. Draft genome sequences of *Saccharibacter* sp. strains 3. A. 1 and M18 isolated from honey and a honey bee (*Apis mellifera*) stomach. *Genome Announc* 2017;5:e00744-17.
- Smith EA, Martin-Eberhardt SA, Miller DL, Parish AJ, Newton ILG, et al. Draft genome sequence of a *Bombella Apis* strain isolated from honey bees. *Microbiol Resour Announc* 2019;8:e01329-19.