Current Biology, Volume *30*

Supplemental Information

Vast Differences in Strain-Level Diversity in the Gut Microbiota of Two Closely Related Honey Bee Species

Kirsten M. Ellegaard, Shota Suenami, Ryo Miyazaki, and Philipp Engel

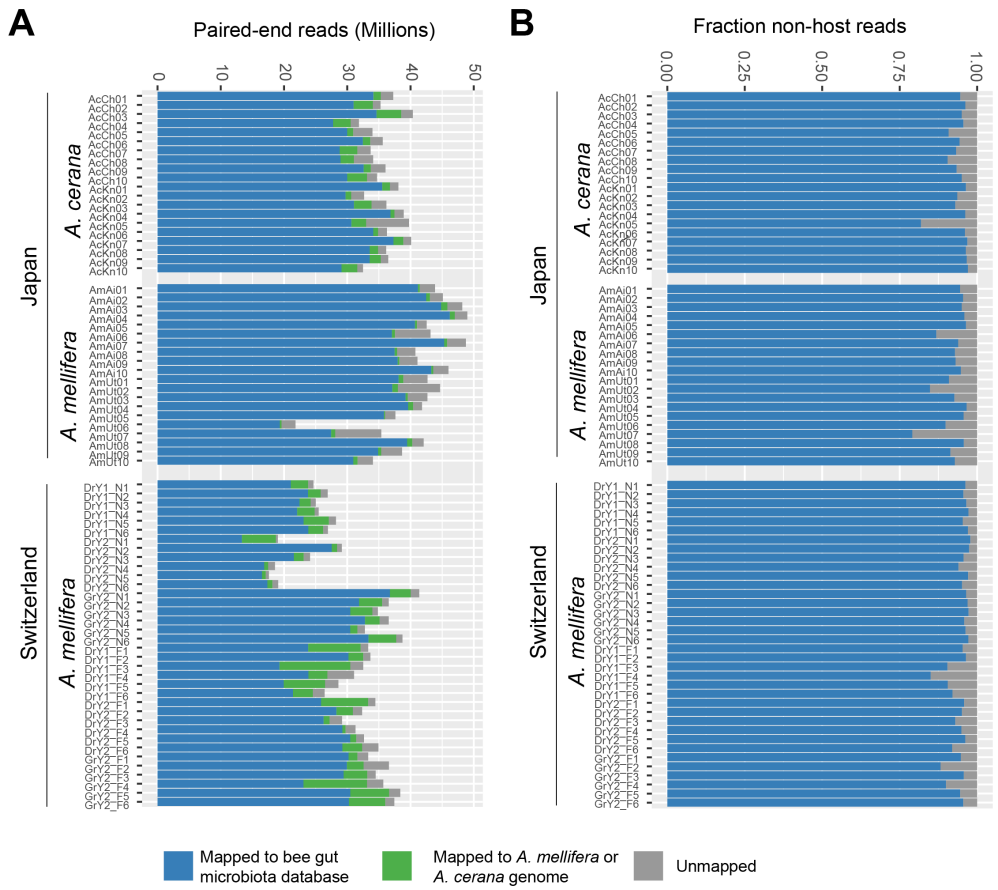

Figure S1. Mapping results on novel genomic database. Related to Figure 1B.

(A) Total number of reads (paired-end) mapped to the honey bee gut microbiota database (blue), to the host genome database (green) and unmapped (grey). **(B)** The relative fraction of reads mapped to the honey bee gut microbiota database (blue) and unmapped reads (grey), excluding host-derived reads.

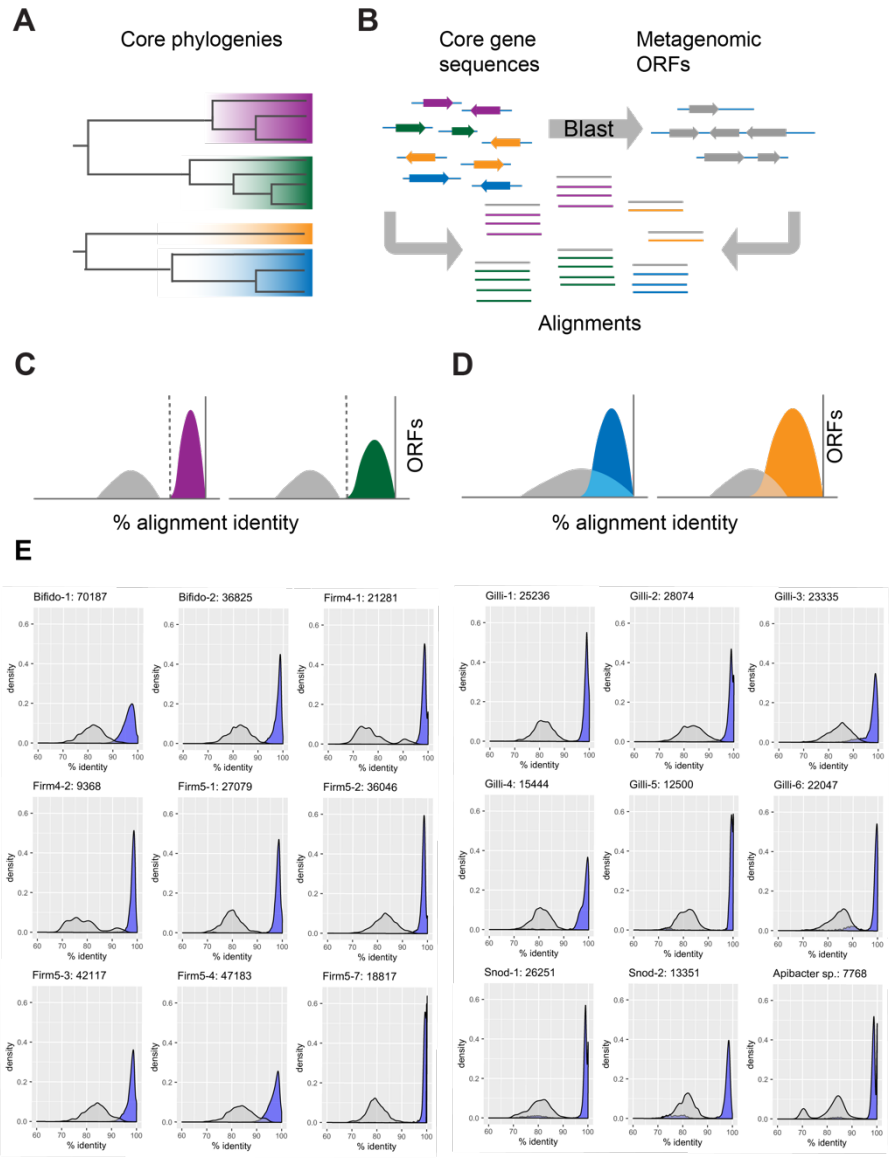

Figure S2. Bioinformatic pipeline for SDPs validation, and SDP validation plots for the current study. Related to Figure 2 and STAR methods.

(A) Candidate SDPs are inferred from isolate genomes, based on core genome phylogenies and pairwise ANI. In this example, two schematic core genome phylogenies are shown, each with two candidate SDPs, as indicated with colors. **(B)** Core genes are extracted from isolate genomes (colored arrows), and aligned separately for each core gene family and SDP (colored lines). Additionally, metagenomic ORFs are recruited to each SDP by blasting the core gene sequences against the metagenomic ORFs (grey arrows). Recruited ORFs are added

individually to the core gene alignments (grey lines), and their maximum percentage identity to the core genes is calculated. **(C,D)** Density distribution plots of the maximum percentage identity of all recruited ORFs. The colored distributions correspond to recruited ORFs with a best blast hit to the candidate SDP being evaluated, grey distributions correspond to recruited ORFs with a closer hit to another SDP in the database. An SDP is considered confirmed if the two distributions are largely non-overlapping **(C)**. **(E)** Density distribution plots for all confirmed SDPs in the current study. Recruited ORFs with best blast-hits (based on percentage identity) to the SDP being evaluated are shown in color, hits to other SDPs in the database are shown in grey. The total number of non-redundant recruited ORFs is indicated in the panel titles for each SDP.

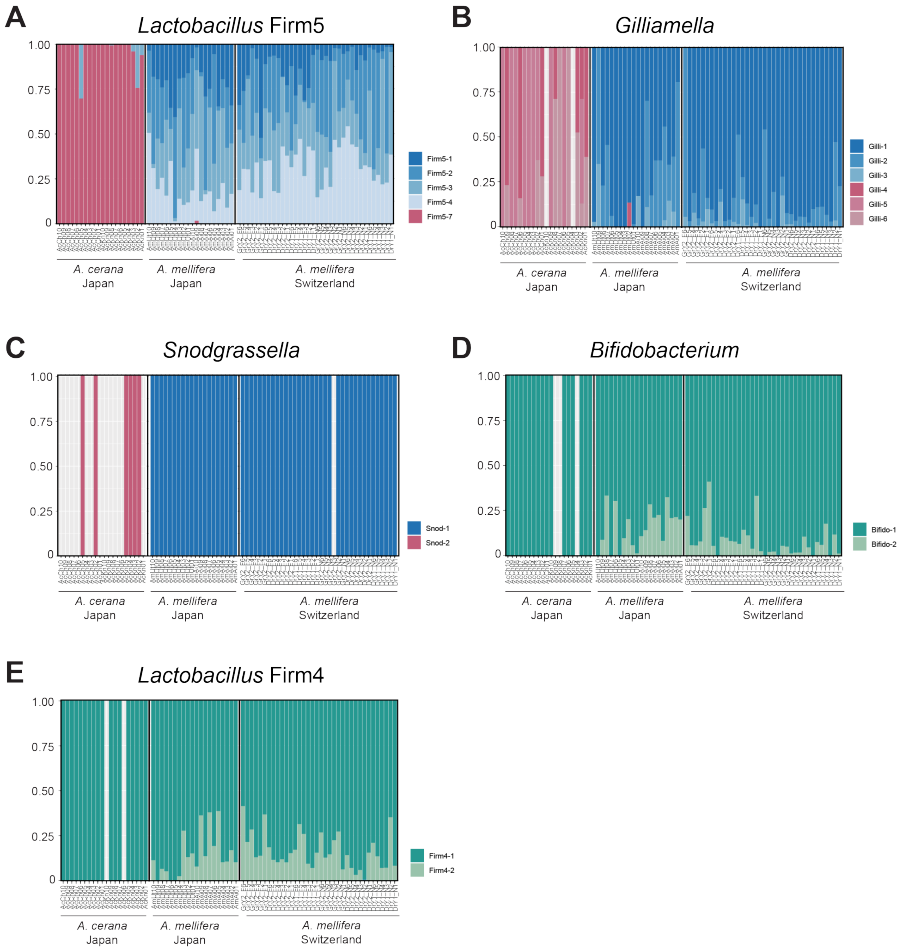

Figure S3. **Relative abundance of SDPs for all core phylotypes, including swiss samples. Related to Figure 2.**

Barplots displaying relative abundance of confirmed SDPs within all five core phylotypes colonizing both *A. mellifera* and *A. cerana*.

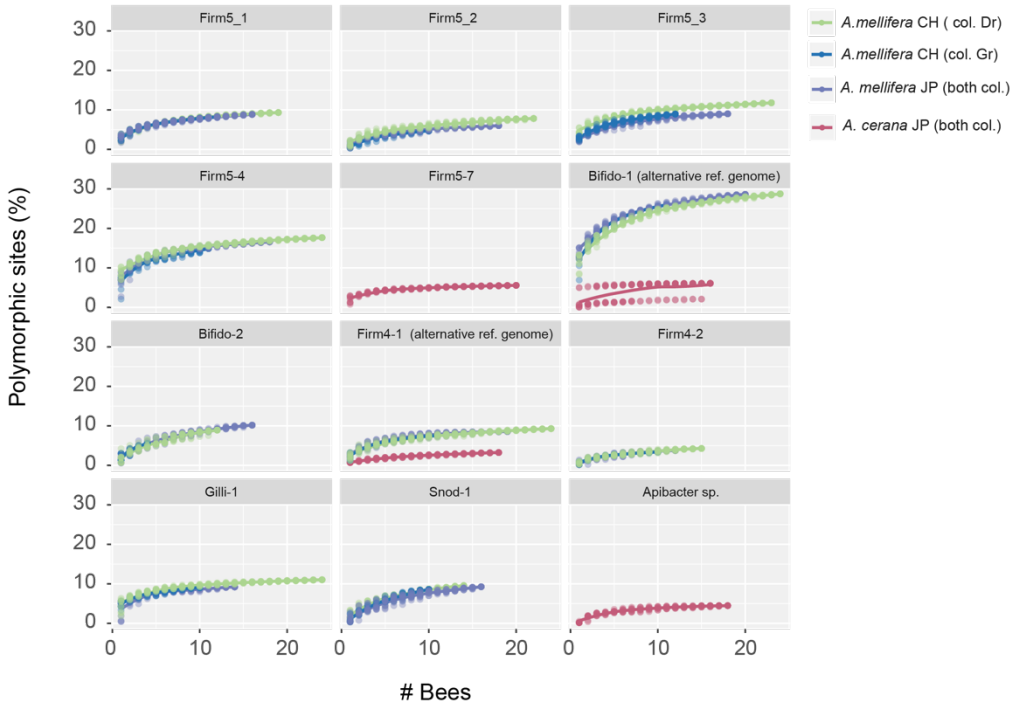

Figure S4. Cumulative curves of the fraction of SNVs relative to the number of samples. Related to Figure 3.

Curves were generated for SDPs where at least 10 samples had sufficient coverage for SNV profiling (min 20x terminus coverage), in at least one colony. Ten random sampling orders were generated per SDP (with individual data points represented by dots). For swiss samples, separate curves were generated for each colony. For Japanese samples, the data were pooled per host (due to the lower number of samples available per colony). For “Bifido-1” and “Firm4-1”, the curves represent the SNV profiling results using a database with a genome isolate representative from the alternate host (Compared to reference genomes used in the main database).

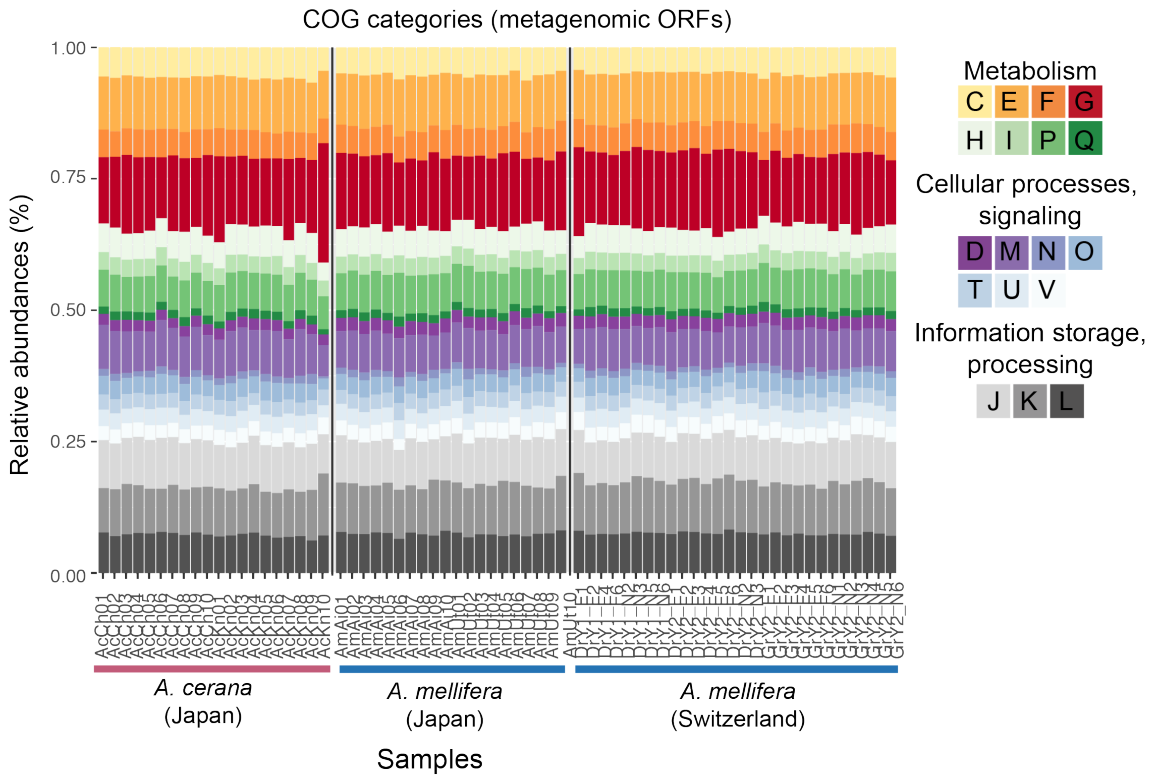

Figure S5. Relative abundance of COG annotations across all metagenomes. Related to Figure 4.

The relative abundance of general functional COG categories are shown for all metagenomes analyzed in the current study, including Swiss samples.

Supplemental References

S1. Ellegaard, K.M., and Engel, P. (2018). New reference genome sequences for 17 bacterial strains of the honey bee gut microbiota. Microbiology Resource Announcements *7*, e00834-00818.

S2. Ellegaard, K.M., Tamarit, D., Javelind, E., Olofsson, T.C., Andersson, S.G., and Vasquez, A. (2015). Extensive intra-phylotype diversity in lactobacilli and bifidobacteria from the honeybee gut. BMC Genomics *16*, 284.

S3. Bottacini, F., Milani, C., Turroni, F., Sanchez, B., Foroni, E., Duranti, S., Serafini, F., Viappiani, A., Strati, F., Ferrarini, A., et al. (2012). Bifidobacterium asteroides PRL2011 genome analysis reveals clues for colonization of the insect gut. PLoS One *7*, e44229.

S4. Milani, C., Lugli, G.A., Duranti, S., Turroni, F., Bottacini, F., Mangifesta, M., Sanchez, B., Viappiani, A., Mancabelli, L., Taminiau, B., et al. (2014). Genomic encyclopedia of type strains of the genus Bifidobacterium. Appl Environ Microbiol *80*, 6290-6302.

S5. Chen, X., E, Z., Gu, D., Lv, L., and Li, Y. (2015). Complete Genome Sequence of Bifidobacterium actinocoloniiforme Type Strain DSM 22766T, Isolated from Bumblebee Digestive Tracts. Genome Announc *3*.

S6. Ellegaard, K.M., Brochet, S., Bonilla-Rosso, G., Emery, O., Glover, N., Hadadi, N., Jaron, K.S., van der Meer, J.R., Robinson-Rechavi, M., Sentchilo, V., et al. (2019). Genomic changes underlying host specialization in the bee gut symbiont Lactobacillus Firm5. Mol Ecol.

S7. Kwong, W.K., Mancenido, A.L., and Moran, N.A. (2014). Genome Sequences of Lactobacillus sp. Strains wkB8 and wkB10, Members of the Firm-5 Clade, from Honey Bee Guts. Genome Announc *2*.

S8. Kwong, W.K., Engel, P., Koch, H., and Moran, N.A. (2014). Genomics and host specialization of honey bee and bumble bee gut symbionts. Proc Natl Acad Sci U S A *111*, 11509-11514.

S9. Ludvigsen, J., Porcellato, D., L'Abee-Lund, T.M., Amdam, G.V., and Rudi, K. (2017). Geographically widespread honeybee-gut symbiont subgroups show locally distinct antibiotic-resistant patterns. Mol Ecol *26*, 6590-6607.

S10. Zheng, H., Nishida, A., Kwong, W.K., Koch, H., Engel, P., Steele, M.I., and Moran, N.A. (2016). Metabolism of Toxic Sugars by Strains of the Bee Gut Symbiont Gilliamella apicola. mBio *7*.

S11. Steele, M.I., Kwong, W.K., Whiteley, M., and Moran, N.A. (2017). Diversification of Type VI Secretion System Toxins Reveals Ancient Antagonism among Bee Gut Microbes. MBio *8*.

S12. Kwong, W.K., Steele, M.I., and Moran, N.A. (2018). Genome Sequences of Apibacter spp., Gut Symbionts of Asian Honey Bees. Genome Biol Evol *10*, 1174-1179.

S13. Engel, P., Vizcaino, M.I., and Crawford, J.M. (2015). Gut symbionts from distinct hosts exhibit genotoxic activity via divergent colibactin biosynthesis pathways. Appl Environ Microbiol *81*, 1502-1512.

S14. Segers, F.H., Kesnerova, L., Kosoy, M., and Engel, P. (2017). Genomic changes associated with the evolutionary transition of an insect gut symbiont into a blood-borne pathogen. ISME J *11*, 1232-1244.

S15. Bonilla-Rosso, G., Paredes Juan, C., Das, S., Ellegaard, K.M., Emery, O., Garcia-Garcera, M., Glover, N., Hadadi, N., van der Meer, J.R., SAGE class 2017-18, et al.

(2019). Acetobacteraceae in the honey bee gut comprise two distant clades with diverging metabolism and ecological niches. bioRxiv.

S16. Li, L., Illeghems, K., Van Kerrebroeck, S., Borremans, W., Cleenwerck, I., Smagghe, G., De Vuyst, L., and Vandamme, P. (2016). Whole-Genome Sequence Analysis of Bombella intestini LMG 28161T, a Novel Acetic Acid Bacterium Isolated from the Crop of a Red-Tailed Bumble Bee, Bombus lapidarius. PLoS One *11*, e0165611.

S17. Roh, S.W., Nam, Y.D., Chang, H.W., Kim, K.H., Kim, M.S., Ryu, J.H., Kim, S.H., Lee, W.J., and Bae, J.W. (2008). Phylogenetic characterization of two novel commensal bacteria involved with innate immune homeostasis in Drosophila melanogaster. Appl Environ Microbiol *74*, 6171-6177.

S18. Servin-Garciduenas, L.E., Sanchez-Quinto, A., and Martinez-Romero, E. (2014). Draft Genome Sequence of Commensalibacter papalotli MX01, a Symbiont Identified from the Guts of Overwintering Monarch Butterflies. Genome Announc *2*.

S19. Tamarit, D., Ellegaard, K.M., Wikander, J., Olofsson, T., Vasquez, A., and Andersson, S.G. (2015). Functionally Structured Genomes in Lactobacillus kunkeei Colonizing the Honey Crop and Food Products of Honeybees and Stingless Bees. Genome Biol Evol *7*, 1455-1473.