

New reference genomes of honey bee-associated bacteria *Paenibacillus melissococcoides*, *Paenibacillus dendritiformis*, and *Paenibacillus thiaminolyticus*

Benjamin Dainat,¹ Simone Oberhaensli,² Florine Ory,¹ Vincent Dietemann^{1,3}

AUTHOR AFFILIATIONS See affiliation list on p. 3.

ABSTRACT We sequenced the genomes of recently discovered *Paenibacillus melissococcoides* (CCOS 2000) and of the type strains of closely related *P. thiaminolyticus* (DSM 7262) and *P. dendritiformis* (LMG 21716). The three genomes set the basis to unambiguous diagnostic of these honey bee associated *Paenibacillus* bacteria.

KEYWORDS brood pathogen, European foulbrood, honey bee health, *Paenibacillus dendritiformis*, *Paenibacillus melissococcoides*, *Paenibacillus thiaminolyticus*

Several bacteria species of the *Paenibacillus* genus are associated with the honey bee, *Apis mellifera*. Among them, *P. larvae* (1) causes American foulbrood, a highly contagious disease, which impedes colony development and can lead to its death. *P. melissococcoides* (2) and *P. dendritiformis* (3) were isolated from colonies affected by European foulbrood (4). *P. thiaminolyticus* was isolated from hive material (5, 6). Because of their close 16S rRNA genetic relatedness (2), *P. melissococcoides* could have been wrongly identified as *P. dendritiformis* or *P. thiaminolyticus* in previous studies.

Here, we present the genome sequences of *P. melissococcoides* CCOS 2000 and of the type strains *P. dendritiformis* LMG 21716 and *P. thiaminolyticus* DSM 7262. The latter bacteria were obtained from BCCM and DSMZ culture collections, respectively. *P. melissococcoides* was found in worker jelly droplets cultured on EFB Basal medium (7, 8) under anaerobic conditions for 4 days at 36°C. The sampled colony was located near Reutigen (46° 41' 39" N, 7° 37' 13" E), Switzerland.

The three bacteria were grown in 10 mL liquid Basal medium at 36°C overnight. High-molecular-weight genomic DNA was recovered using the GES method of DNA extraction (9) and assessed for quantity, quality, and purity using a Qubit 4.0 fluorometer (dsDNA HS Assay kit; Q32851, Thermo Fisher Scientific, Waltham, MA, USA), an Advanced Analytical FEMTO Pulse instrument (Genomic DNA 165 kb Kit; FP-1002-0275, Agilent, Santa Clara, CA, USA), and a Denovix DS-11 UV-Vis spectrophotometer. Multiplexed SMRTbell libraries were prepared according to PacBio guidelines, Part Number 101-696-100 Version 06 (March 2020). One microgram of gDNA in 100 µL was sheared in a g-TUBE (Covaris, Woburn, MA, USA), concentrated, and cleaned using AMPure PB beads. Samples were quantified and qualified to be in the range of 12–15 kb using the Qubit and the FEMTO instruments, respectively. Libraries were pooled using the PacBio microbial multiplexing calculator. Prior to and after size selection, the library pool was purified using AMPure PB beads. Size selection was performed with a BluePippin instrument (BLU0001; Sage Science, Beverly, MA, USA) using BluePippin with dye free, 0.75% Agarose Cassettes, and S1 Marker (Sage Science; BLF7510) wherein the selection cut-off was set at 6,000 bp. Library pool concentration and size was

Editor Irene L. G. Newton, Indiana University, Bloomington, Bloomington, Indiana, USA

Address correspondence to Benjamin Dainat, benjamin.dainat@agroscope.admin.ch.

The authors declare no conflict of interest.

See the funding table on p. 4.

Received 27 April 2023

Accepted 18 June 2023

Published 2 August 2023

Copyright © 2023 Dainat et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

again assessed using the Qubit and FEMTO instruments, respectively. The final library pools were on average 11.4 kb in size.

PacBio Sequencing primer v4 and Sequel DNA Polymerase 3.0 were annealed and bound, respectively, to the DNA template libraries. Libraries and Spike-In internal control were diffusion loaded at an on-plate concentration of 10 or 11 pM. Sequencing was performed in continuous long read (CLR) mode on the Sequel System with Sequencing kit 3.0, SMRTCells 1M v3, and a 2 h pre-extension followed by 600 min movie time.

CLRs were assembled with the microbial assembly pipeline (<https://www.pacb.com/products-and-services/analytical-software/smrt-analysis/>, SMRTlink v9.0.0.92188) with default parameter settings (except `microasm_coverage = 25` and `microasm_genome_size = 6.8 Mb`). The pipeline does automatic overlap identification, trimming, and rotation of circular sequences.

Raw data and assembly statistics are summarized in Table 1.

In a phylogenetic tree build using single-copy genes (Fig. 1), *P. melissococcoides* is located close to *P. dendritiformis* [ANI value 92.4%, calculated with fastANI v1.1 (12) using the genome assemblies] and *P. thiaminolyticus* (genome ANI value 91.1%) but distant from the known foulbrood pathogen *P. larvae*.

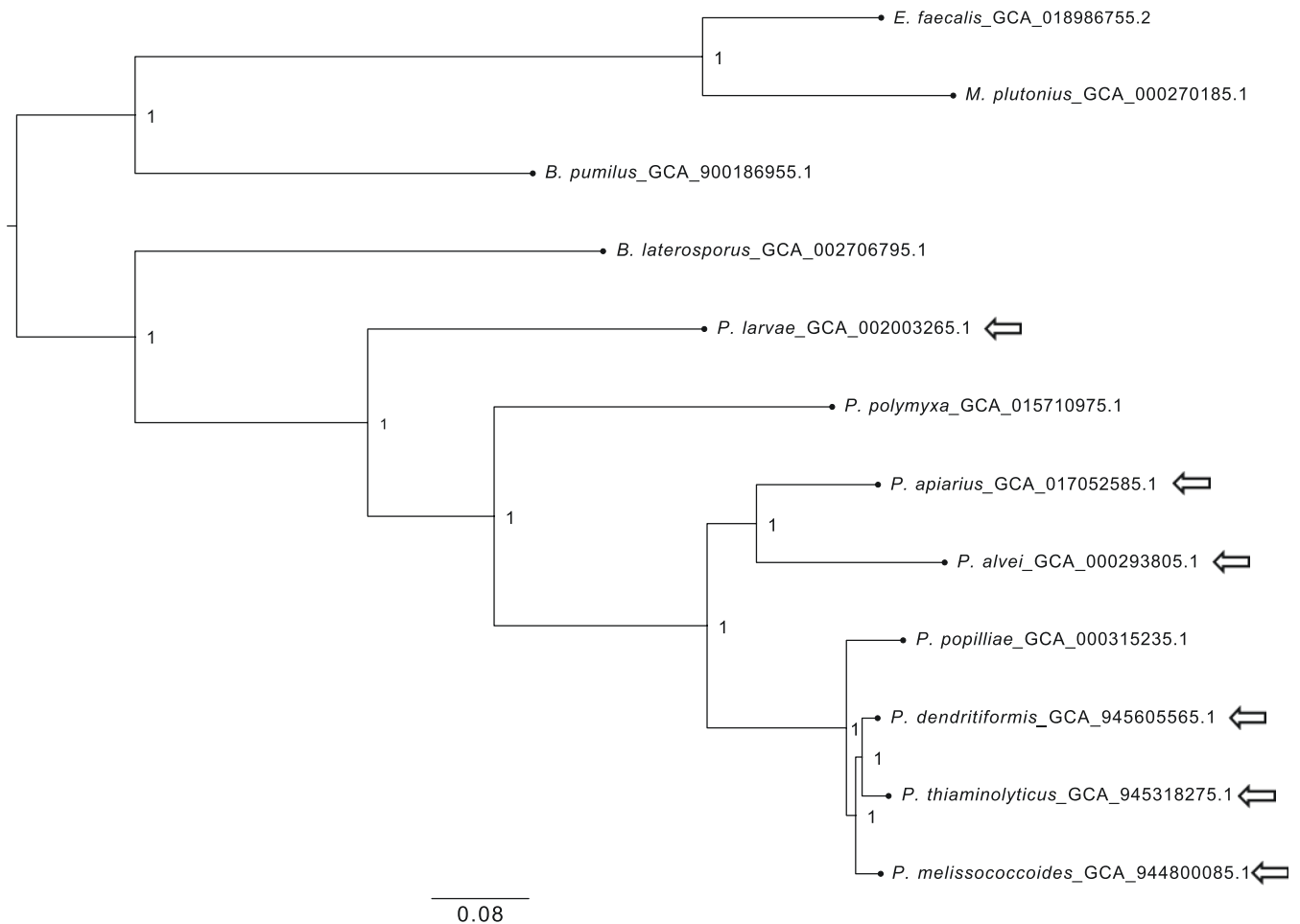


FIG 1 Species tree inferred from concatenated multiple sequence alignments of single-copy genes from *Paenibacillus* spp., *Bacillus* spp., and other bacteria that can be found in a honey bee colony. The tree was generated using Orthofinder v2.3.8 with option '-M msa' (13–16) and rooted using STRIDE. STAG support values (15) are indicated at internal nodes, the species *M. plutonius*, *B. pumilus*, and *E. faecalis* were used as outgroup to root the tree. Arrows indicate *Paenibacillus* species associated with honey bees.

TABLE 1 Sequencing statistics and genome data availability of *Paenibacillus melissococcoides*, *Paenibacillus dendritiformis*, and *Paenibacillus thiaminolyticus*^a

Parameter	Data for:		
	<i>P. melissococcoides</i>	<i>P. dendritiformis</i>	<i>P. thiaminolyticus</i>
No. of polymerase reads, N50	29,360, 71,000	24,865, 70,800	26,390, 50,900
No. of subreads, N50	157,996, 7,100	154,543, 6,700	107,539, 6,900
No. of contigs	21, two of them circular	1, linear	4, one circular
N50 (bp)	594,566	n.a. ^b	6,594,752
L50	4	n.a.	1
Assembly size (bases)	7,186,093	6,722,799	6,609,466
Contig type, length (bases), coverage	Circular 1: 16,000, 304x; Circular 2: 63,000, 264x; Coverage of the linears: 165x–248x, mean 174x	Linear, 152x	Circular : 6,609,466; Linear 1*: 11,000, 20x; Linear 2*: 2,000, 1x; Linear 3: 900, 18x * Linear contig 2 and 3 could be assembly artifacts based on either length or coverage
GC content average (%)	53	54.1	53
Bacterial strain deposition in a repository : accession no.	German Collection of Microorganisms and Cell Cultures DSM 113619, Belgian Coordinated Collections of Microorgan- isms LMG 32539, Culture Collection of Switzerland CCOS 2000	n.a.	n.a.
BioProject accession no.	PRJEB49674	PRJEB49674	PRJEB49674
BioSample accession no.	SAMEA14251711	SAMEA14509629	SAMEA14509628
Genome assembly accession no.	GCA_944800085	GCA_945605565	GCA_945318275
SRA accession no.	ERX9376482	ERX9376366	ERX9375765

^aGenome assemblies' completeness was assessed with BUSCO [v4.0.6 (10), with option `-auto-lineage-prok`]. From the bacillales_odb10 database (creation date: 24 April 2019, number of species: 409, number of BUSCOs: 450), four BUSCO genes are missing in the *P. melissococcoides* assembly, one is fragmented, and six are complete but duplicated. These numbers are comparable to the reference strains *P. dendritiformis* LMG 21716 (missing: 2, fragmented: 0, duplicated: 6) and *P. thiaminolyticus* DMZ 7262 (missing: 2, fragmented: 1, duplicated: 15). Since the data of LMG 21716 were assembled into a single chromosome, six duplicated BUSCOs are presumably typical for *Paenibacillus* spp. Assemblies were annotated using PGAP (https://www.ncbi.nlm.nih.gov/genome/annotation_prok/, 2022-02-10.build5872), and the EMBLmyGFF3 tool (11) was used to convert the GFF3 annotation files into EMBL format to allow data submission to all three databases of the International Nucleotide Sequence Database Collaboration (INSDC).

^bn.a.; not applicable.

ACKNOWLEDGMENTS

This work was funded by the Agroscope and the Gottfried and Julia Bangerter-Rhyner foundation.

We thank the Next Generation Sequencing Platform of the University of Bern for the high-throughput sequencing.

AUTHOR AFFILIATIONS

¹Swiss Bee Research Center, Agroscope Liebefeld, Bern, Switzerland

²Interfaculty Bioinformatics Unit and SIB Swiss Institute of Bioinformatics, University of Bern, Bern, Switzerland

³Department of Ecology and Evolution, Biophore, UNIL-Sorge, University of Lausanne, Lausanne, Switzerland

AUTHOR ORCIDs

Benjamin Dainat  <http://orcid.org/0000-0002-1740-7136>

Simone Oberhaensli  <http://orcid.org/0000-0002-2116-1685>

Florine Ory  <http://orcid.org/0000-0002-7460-7943>

Vincent Diemann  <http://orcid.org/0000-0002-0791-9939>

FUNDING

Funder	Grant(s)	Author(s)
Agroscope		Benjamin Dainat
Gottfried und Julia Bangerter-Rhyner-Stiftung (Bangerter-Stiftung)		Benjamin Dainat Vincent Dietemann Simone Oberhänsli Florine Ory

AUTHOR CONTRIBUTIONS

Benjamin Dainat, Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review and editing | Simone Oberhänsli, Validation, Visualization, Writing – review and editing, Data curation, Formal analysis, Methodology | Florine Ory, Conceptualization, Investigation, Visualization, Writing – review and editing, Formal analysis | Vincent Dietemann, Conceptualization, Investigation, Validation, Visualization, Writing – review and editing

DATA AVAILABILITY

See Table 1.

REFERENCES

- White GF. 1906. The bacteria of the Apiary with special reference to bee diseases. In Tech series U.S. Department of Agriculture: 1-50. Government Printing Office, Washington, DC. <https://doi.org/10.5962/bhl.title.87503>
- Ory F, Dietemann V, Guisolan A, von Ah U, Fleuti C, Oberhänsli S, Charrière J-D, Dainat B. 2023. *Paenibacillus melissococcoides* sp. nov., isolated from a honey bee colony affected by European foulbrood disease. *Int J Syst Evol Microbiol* 73. <https://doi.org/10.1099/ijsem.0.005829>
- Gaggia F, Baffoni L, Stenico V, Alberoni D, Buglione E, Lilli A, Di Gioia D, Porrini C. 2015. Microbial investigation on honey bee larvae showing atypical symptoms of European foulbrood. *Bull Insectology* 68:321–327.
- Burri R. 1906. Bakteriologische Untersuchungen Über die Faulbrut und Sauerbrut der Bienen. Sauerländer ed, Aarau.
- Nakamura LK. 1996. *Paenibacillus apiarius* sp. nov.. *Int J Syst Bacteriol* 46:688–693. <https://doi.org/10.1099/00207713-46-3-688>
- Shida O, Takagi H, Kadowaki K, Nakamura LK, Komagata K. 1997. Transfer of *Bacillus alginolyticus*, *Bacillus chondroitinus*, *Bacillus curdlanolyticus*, *Bacillus gluconolyticus*, *Bacillus kobensis*, and *Bacillus thiaminolyticus* to the genus *Paenibacillus* and emended description of the genus *Paenibacillus*. *Int J Syst Bacteriol* 47:289–298. <https://doi.org/10.1099/00207713-47-2-289>
- Grossar D, Kilchenmann V, Forsgren E, Charrière J-D, Gauthier L, Chapuisat M, Dietemann V. 2020. Putative determinants of virulence in *Melissococcus plutonius*, the bacterial agent causing European foulbrood in honey bees. *Virulence* 11:554–567. <https://doi.org/10.1080/21505594.2020.1768338>
- Forsgren E, Budge GE, Charrière J-D, Hornitzky MAZ. 2013. Standard methods for European foulbrood research. *J Apic Res* 52:1–14. <https://doi.org/10.3896/IBRA.1.52.1.12>
- Pitcher DG, Saunders NA, Owen RJ. 1989. Rapid extraction of bacterial genomic DNA with guanidium thiocyanate. *Lett Appl Microbiol* 8:151–156. <https://doi.org/10.1111/j.1472-765X.1989.tb00262.x>
- Manni M, Berkeley MR, Seppely M, Simão FA, Zdobnov EM. 2021. BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol Biol Evol* 38:4647–4654. <https://doi.org/10.1093/molbev/msab199>
- Norling M, Jareborg N, Dainat J. 2018. EMBLmyGFF3: a converter facilitating genome annotation submission to european nucleotide archive. *BMC Res Notes* 11:584. <https://doi.org/10.1186/s13104-018-3686-x>
- Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>
- Emms DM, Kelly S. 2015. Orthofinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* 16:157. <https://doi.org/10.1186/s13059-015-0721-2>
- Emms DM, Kelly S. 2017. STRIDE: species tree root inference from gene duplication events. *Mol Biol Evol* 34:3267–3278. <https://doi.org/10.1093/molbev/msx259>
- Emms DM, Kelly S. 2018. STAG: Species tree inference from all genes. *bioRxiv*. <https://doi.org/10.1101/267914>
- Emms DM, Kelly S. 2019. Orthofinder: phylogenetic orthology inference for comparative genomics. *Genome Biol* 20:238. <https://doi.org/10.1186/s13059-019-1832-y>