




# Whole-Genome Sequences of a *Lactobacillus melliventris* Strain and Its *Myovirus* Temperate Phage, pHBH004, Isolated from the Digestive Tract of *Apis mellifera* in Switzerland

Andrew F. Brown,<sup>a,b</sup> Judith Pfister,<sup>a</sup> Javier E. Fernandez,<sup>b,c</sup> Valentina Donà,<sup>c</sup> Victor Rodriguez,<sup>a</sup> Gina Retschnig,<sup>a</sup> Alexandra Collaud,<sup>c</sup> Peter Neumann,<sup>a</sup>  Vincent Perreten<sup>c</sup>

<sup>a</sup>Institute of Bee Health, Vetsuisse Faculty, University of Bern, Bern, Switzerland

<sup>b</sup>Graduate School of Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland

<sup>c</sup>Division of Molecular Bacterial Epidemiology and Infectious Diseases, Institute of Veterinary Bacteriology, Vetsuisse Faculty, University of Bern, Bern, Switzerland

**ABSTRACT** The complete genome sequence of *Lactobacillus melliventris* strain IBH004, isolated from the gut of a honeybee worker (*Apis mellifera*) and containing two plasmids and a temperate phage, was determined using hybrid assembly of Oxford Nanopore and Illumina reads. Phage-sequence relationships were identified from the coding sequences, and a proteomic tree was constructed.

**L** *actobacillus* spp. are important commensals for Western honeybees (*Apis mellifera*). They regulate opportunistic bacteria and fungi (1, 2), aid in digestion (3), and produce essential micronutrients (4). Here, we present the circular genome of *Lactobacillus melliventris* IBH004 and its temperate phage.

The mid- and hindguts of an adult *A. mellifera* worker in Switzerland were streaked onto MRS agar (Sigma-Aldrich) containing 20 g/L fructose and incubated anaerobically at 30°C for 24 h. A single colony purified three times under the same conditions was identified as a *Lactobacillus* species by matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) mass spectrometry (Bruker) and cryopreserved. Genomic DNA was obtained from a colony lawn by acid guanidinium thiocyanate-phenol-chloroform extraction (5) and purified using CleanNA beads. DNA sequencing was performed using a NEBNext Ultra II directional DNA library prep kit with TruSeq adapters on an Illumina NovaSeq 6000 system (2 × 150-bp paired-end format) at Eurofins Genomics (Germany); sequencing was also conducted using the MinION MK1b device on a R9.4.1 SpotON flow cell using the 1D ligation kit (SQK-LSK109) (Oxford Nanopore Technology [ONT]). The raw Illumina reads were quality controlled using FastQC v0.11.7 (6) and quality filtered and trimmed using Prinseq v0.6 (7). The ONT reads (28,226 reads; read length  $N_{50}$ , 14,169 bases) were base called using Guppy v4.4.1 (8) and size filtered, keeping reads of >5,000 bp, using Geneious Prime v2022.0.2 (Biomatters Ltd.). The resulting ONT reads (13,158 reads; mean read length, 10,633 bp; total 177,638,807 bases; 90× coverage) were *de novo* assembled using Flye v1.3 (9) with a 2,000-bp minimum overlap length. The ONT assemblies were corrected using quality-filtered paired-end Illumina reads (12,089,792 reads, total 1,778,546,634 bases, 907× coverage) with Polypolish v0.5.0 (10), confirmed to be circular and rotated to DnaA using Circlator v1.5.5-docker5 (11), and annotated using PGAP v6.3 (12). Default parameters were used for all software unless otherwise specified.

The assembly of IBH004 contained a circular chromosome of 1,965,960 bp and two circular plasmids of 9,000 bp (pIBH004-1) and 6,359 bp (pIBH004-2), with a total GC content of 36.5%, 1,845 protein-coding sequences (CDSs), 14 pseudogenes, and 70 RNA-encoding genes (4 rRNAs [5S, 16S, 23S], 55 tRNAs, 3 noncoding RNAs [ncRNAs]). IBH004 was identified as *L. melliventris* based on the digital DNA-DNA-hybridization (dDDH) (d4, 78.2%) and orthologous CDS average nucleotide identity (ANI) (97.6%) values obtained between the

**Editor** Steven R. Gill, University of Rochester School of Medicine and Dentistry

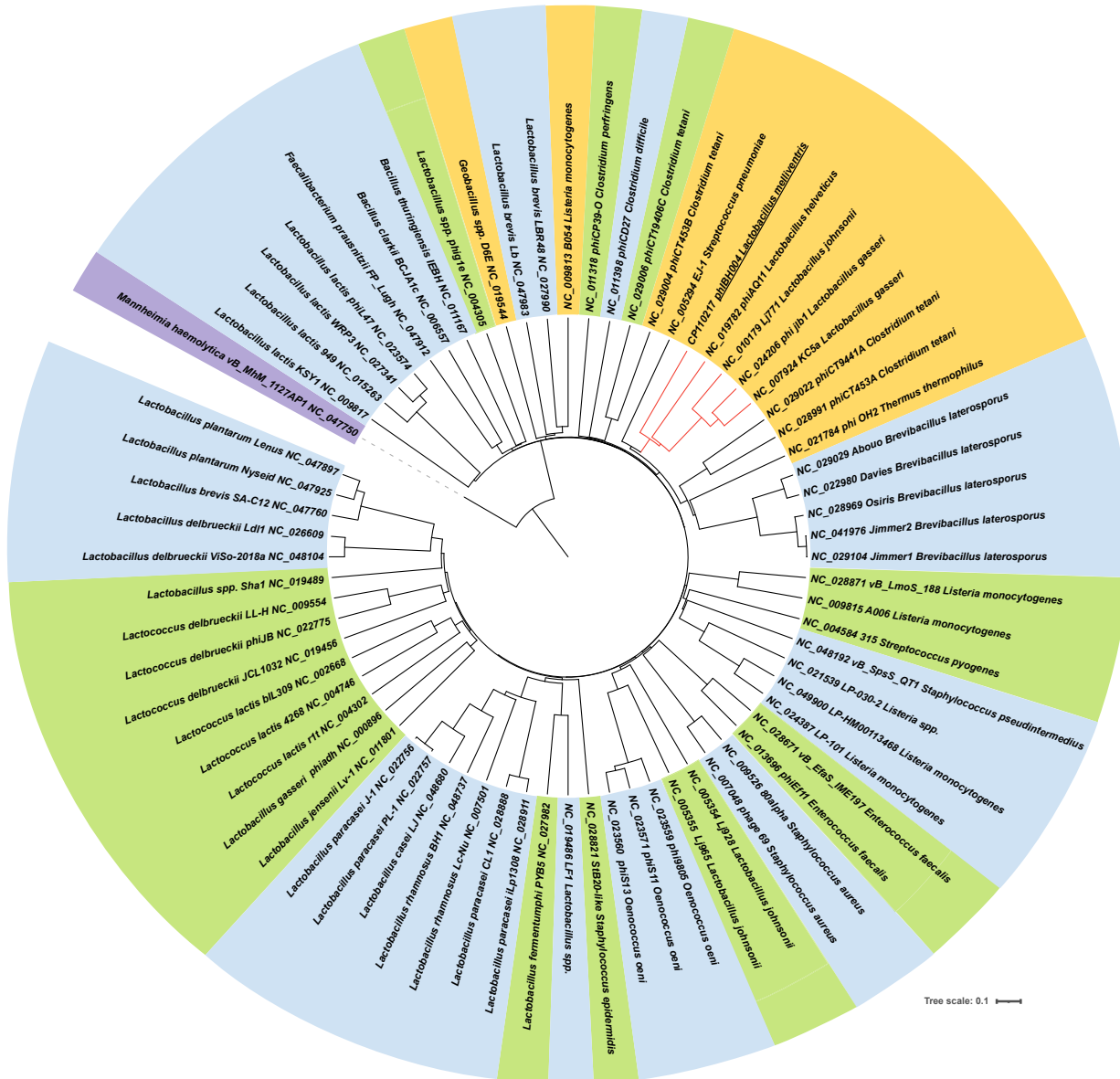
**Copyright** © 2023 Brown et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Vincent Perreten, [vincent.perreten@unibe.ch](mailto:vincent.perreten@unibe.ch).

The authors declare no conflict of interest.

**Received** 18 January 2023

**Accepted** 4 February 2023



**FIG 1** Virus proteomic tree (VIPTree) of prophage phiBH004 from *Lactobacillus melliventris* strain IBH004 and 70 related phages obtained using PHASTER (17, 18). The GenBank accession number and name of each phage are provided and juxtaposed with their corresponding bacterial host; orange shading indicates phages of the Myoviridae family, green indicates Siphoviridae, purple Peduoviridae, and blue unassigned phage families, as obtained using the VIPTree v3.3 server (21). The *Peduoviridae* phage vB\_MhM\_1127AP1 of *Mannheimia haemolytica* was used to root the tree. Branch lengths are indicated using a linear scale and indicate relationships. The red lines indicate the *Lactobacillus* spp. Myoviridae phages belonging to the same cluster as phiBH004 (underlined).

genomes of IBH004 and the type strain, *L. melliventris* Hma8 (GenBank accession number [JXLI000000000](https://genbank.ncbi.nlm.nih.gov/GenBank/FASTA/JXLI000000000)) (13), which were above the species delimitation thresholds of 70% and 95% to 96%, respectively (14), as calculated using GGDC v3.0 (<https://tygs.dsmz.de/>) (15) and OrthoANlu (<http://www.ezbiocloud.net/tools/ani>) (16).

An intact 42,508-bp prophage (phiBH004) (GC content, 37.5%; 65 CDSs; 2 tRNAs) was identified using PHASTER (score of 150) (<https://phaster.ca>) in the chromosome of IBH004 at position 490469 to 532976 (17, 18) and was flanked with attachment sites (attL/attR, CGGAGACTGAG). Using Virfam (<http://biodev.extra.cea.fr/virfam>), phiBH004 was classified as Myoviridae type 1, cluster1 (19). PHASTER revealed protein hits with 70 bacteriophages. Their genomic sequences were accessed from GenBank (20), uploaded to the VIPTree v3.3 server (21), and visualized using iTOL v6 (22). The resulting proteomic VIPTree generated by tBLASTx comparisons (21) placed phiBH004 with other *Lactobacillus* spp. Myoviridae phages (Fig. 1). Read mapping the ONT reads against the chromosome revealed a mean coverage of the

prophage sequence that was 4.5 times higher (326×) than that of the rest of the chromosome (71×), indicating the presence of free temperate phage in the bacterial culture.

**Data availability.** The genome sequence of *L. melliventris* strain IBH004 has been deposited in GenBank under accession numbers [CP110217](https://doi.org/10.1093/nar/gkw569) (chromosome, including prophage phiBH004), [CP110218](https://doi.org/10.1093/nar/gkw569) (pIBH004-1), and [CP110219](https://doi.org/10.1093/nar/gkw569) (pIBH004-2). The BioProject and BioSample accession numbers are [PRJNA764829](https://doi.org/10.1093/nar/gkw569) and [SAMN21529766](https://doi.org/10.1093/nar/gkw569). The raw reads were deposited in the SRA under accession numbers [SRX12291800](https://doi.org/10.1093/nar/gkw569) (ONT) and [SRX12291801](https://doi.org/10.1093/nar/gkw569) (Illumina).

## ACKNOWLEDGMENTS

This project was financed by the Ricola Foundation, Nature and Culture (REF-38-899), the Vinetum foundation, and internal funds of the Institute of Veterinary Bacteriology of the University of Bern (REF-660-50).

We thank Felipe Hernandez Coutinho for his generous advice on data visualization.

## REFERENCES

- Vásquez A, Forsgren E, Fries I, Paxton RJ, Flaberg E, Szekely L, Olofsson TC. 2012. Symbionts as major modulators of insect health: lactic acid bacteria and honeybees. *PLoS One* 7:e33188. <https://doi.org/10.1371/journal.pone.0033188>.
- Tejerina MR, Cabana MJ, Benitez-Ahrendts MR. 2021. Strains of *Lactobacillus* spp. reduce chalkbrood in *Apis mellifera*. *J Invertebr Pathol* 178:107521. <https://doi.org/10.1016/j.jip.2020.107521>.
- Lee FJ, Rusch DB, Stewart FJ, Mattila HR, Newton ILG. 2015. Saccharide breakdown and fermentation by the honey bee gut microbiome. *Environ Microbiol* 17:796–815. <https://doi.org/10.1111/1462-2920.12526>.
- 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* 111:11509–11514. <https://doi.org/10.1073/pnas.1405838111>.
- 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>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Schmieder R, Edwards R. 2011. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. <https://doi.org/10.1093/bioinformatics/btr026>.
- Wick RR, Judd LM, Holt KE. 2019. Performance of neural network base calling tools for Oxford Nanopore sequencing. *Genome Biol* 20:129. <https://doi.org/10.1186/s13059-019-1727-y>.
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>.
- Wick RR, Holt KE. 2022. Polypolish: short-read polishing of long-read bacterial genome assemblies. *PLoS Comput Biol* 18:e1009802. <https://doi.org/10.1371/journal.pcbi.1009802>.
- Hunt M, Silva DN, Otto TD, Parkhill J, Keane JA, Harris SR. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol* 16:294. <https://doi.org/10.1186/s13059-015-0849-0>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Ellegaard KM, Tamarit D, Javelind E, Olofsson TC, Andersson SGE, Vásquez A. 2015. Extensive intra-phyloclade diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* 16:284. <https://doi.org/10.1186/s12864-015-1476-6>.
- Chun J, Oren A, Ventosa A, Christensen H, Arahal D, da Costa M, Rooney A, Yi H, Xu X, de Meyer S, Trujillo M. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68:461–466. <https://doi.org/10.1099/ijsem.0.002516>.
- Meier-Kolthoff JP, Carbasse JS, Peinado-Olarte RL, Göker M. 2022. TYGS and LPSN: a database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res* 50:D801–D807. <https://doi.org/10.1093/nar/gkab902>.
- Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 110:1281–1286. <https://doi.org/10.1007/s10482-017-0844-4>.
- Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, Wishart DS. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 44:W16–W21. <https://doi.org/10.1093/nar/gkw387>.
- Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. 2011. PHAST: a fast phage search tool. *Nucleic Acids Res* 39:W347–W352. <https://doi.org/10.1093/nar/gkr485>.
- Lopes A, Tavares P, Petit MA, Guérois R, Zinn-Justin S. 2014. Automated classification of tailed bacteriophages according to their neck organization. *BMC Genomics* 15:1027. <https://doi.org/10.1186/1471-2164-15-1027>.
- Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2013. GenBank. *Nucleic Acids Res* 41:D36–D42. <https://doi.org/10.1093/nar/gks1195>.
- Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. 2017. ViP-Tree: the viral proteomic tree server. *Bioinformatics* 33:2379–2380. <https://doi.org/10.1093/bioinformatics/btx157>.
- Letunic I, Bork P. 2021. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 49:W293–W296. <https://doi.org/10.1093/nar/gkab301>.