First Draft Genome Sequence of a *Polymyxa* Genus Member, *Polymyxa betae,* the Protist Vector of Rhizomania

Alain Decroës,ª Magnalena Calusinska,ᵇ Philippe Delfosse,ᵇ Claude Bragard,ª Anne Legrèveª

ªEarth and Life Institute, Applied Microbiology, Phytopathology, Université catholique de Louvain, Louvain-la-Neuve, Belgium
ᵇEnvironmental Research and Innovation Department, Luxembourg Institute of Science and Technology, Belvaux, Luxembourg

ABSTRACT *Polymyxa betae* belongs to the *Plasmodiophorida* (Phytomyxea, Rhizaria). Here, we report the first draft genome sequence of a member of the *Polymyxa* genus, which includes two obligate root endoparasite species, vectors of important soilborne plant viruses. The genome assembly was represented by 1,001 contigs, with a cumulated length of 27,085,946 bp.

*P*olymyxa betae* and *Polymyxa graminis* belong to the *Plasmodiophorida,* a protist order which includes obligate root endoparasites of plants (1, 2). Although *Polymyxa* sp. infections are symptomless, these species are considered a real threat because of their vectorial capacity for viruses (3–5). Molecular studies of these organisms are difficult due to their biotrophic nature. Recently, they have been boosted with the publication of *Plasmodiophora brassicae* and *Spongospora subterranea* genomes and transcriptomes (6–9). Here, we present the first full draft genome sequence of *P. betae*, the vector of *Beet necrotic yellow vein virus*, the agent of rhizomania, *Beet soilborne mosaic virus, Beet soilborne virus*, and *Beet virus Q*.

Three-week-old *Beta vulgaris* of the fully sequenced monogerm DH line KWS2320 (10) were inoculated with zoospores of a *P. betae* single sporosorus (aggregation of resting spores) strain, and grown for 14 weeks in an automated immersion system (11). Heavily infested roots were harvested, surface disinfected with 1 N NaOH and 0.1% SDS, and the sporosori were extracted in a Virtis blender. After sequential filtration steps (500, 200, and 100 μm), sporosori were disinfected with 70% ethyl alcohol (EtOH) and 0.5% antiseptic liquid (Dettol), rinsed three times in sterile distilled water, resuspended in 100 μg/ml cycloheximide (Carl-Roth), 50 μg/ml tetracycline (Sigma-Aldrich), and 100 μg/ml gentamicin sulfate (Sigma-Aldrich) and kept at 4°C. Sporosori were then purified by two 3-h-long centrifugations at 4,500 × *g* on a 0 to 30% continuous gradient of HS-40 colloidal silica (Ludox; Sigma-Aldrich), treated with 50 U/ml RQ1 DNase (Promega) at 37°C for 1 h, and rinsed five times. The resulting material was ground in liquid nitrogen with mortar and pestle, and DNA was extracted following a cetyltrimethylammonium bromide (CTAB) extraction procedure (12).

One library was prepared with the TruSeq Nano DNA sample preparation kit and paired-end sequenced using an Illumina MiSeq instrument and the reagent kit version 3, with a sequencing length of 2 × 250 bp, yielding 12,292 Mb. After adapter trimming with Trimmomatic version 0.32 (13), 49,092,018 reads were quality trimmed and assembled using CLC Genomics Workbench version 9.5.2, resulting in 12,450 contigs (≥1 kbp). Contigs were binned using MetaBAT version 0.32.4 (14), resulting in 47 bins, which were assessed for genome completeness and contamination and assigned bin higher-level taxonomy with CheckM version 1.0.11 (15). Bacterial metagenome-assembled genomes (MAGs) and putative viral (phage) genomes were discarded following further BLAST annotation against the NCBI nr database. Raw reads originating from *B. vulgaris* KWS2320 leaf DNA (SRA accession numbers SRR869626 and

SRR869627) were mapped to the resulting contigs with CLC Genomics Workbench version 9.5.2 to further remove plant DNA contamination. The remaining contigs were visualized with ICoVeR (16), and based on the GC content, coding density, contig coverage, and the tetranucleotide composition, a subset of 1,001 contigs was selected. This final potentially free-of-contamination assembly had a total length of 27,085,946 bp, a coverage of 225×, and an N_{50} value of 61,339 bp (L_{50}, 106). The GC content is 45.1%, a percentage similar to that of *S. subterranea* (45.7%) but lower than that of *P. brassicae* (58.5%) (6, 8). Genome completeness was assessed using CEGMA version 2.5 (17). Of the 248 CEGMA eukaryotic core genes, 91.3% were completely (89.1%) or partially (2%) recovered in the assembly. The k-mer genome size estimation using the 17-mer coverage distribution inferred with Jellyfish version 2.0 (18) led to a complete genome size of 29,353,583 bp.

Data availability. This whole-genome shotgun project has been deposited in DDBJ/ENA/GenBank under the accession number RBZT00000000 and BioProject number PRJNA494670. Raw reads are available under SRA accession number SRR8105394. The version described in this paper is the first version, RBZT01000000.

REFERENCES

1. Neuhauser S, Kirchmair M, Bulman S, Bass D. 2014. Cross-kingdom host shifts of phytomyxid parasites. BMC Evol Biol 14:33. https://doi.org/10.1186/1471-2148-14-33.
2. Neuhauser S, Bulman S, Kirchmair M. 2010. Plasmodiophorids: the challenge to understand soil-borne, obligate biotrophs with a multiphasic life cycle, p. 51–78. *In* Gherbawy Y, Voigt K (ed), Molecular identification of fungi. Springer Berlin Heidelberg, Berlin, Germany.
3. Barr DJS. 1979. Morphology and host range of Polymyxa graminis, Polymyxa betae, and Ligniera pilorum from Ontario and some other areas. Can J Plant Pathol 1:85–94. https://doi.org/10.1080/07060667909501468.
4. Kanyuka K, Ward E, Adams MJ. 2003. Polymyxa graminis and the cereal viruses it transmits: a research challenge. Mol Plant Pathol 4:393–406. https://doi.org/10.1046/j.1364-3703.2003.00177.x.
5. Tamada T, Kondo H. 2013. Biological and genetic diversity of plasmodiophorid-transmitted viruses and their vectors. J Gen Plant Pathol 79:307–320. https://doi.org/10.1007/s10327-013-0457-3.
6. Ciaghi S, Neuhauser S, Schwelm A. 2018. Draft genome resource for the potato powdery scab pathogen Spongospora subterranea. Mol Plant Microbe Interact 31:1227–1229. https://doi.org/10.1094/MPMI-06-18-0163-A.
7. Rolfe SA, Strelkov SE, Links MG, Clarke WE, Robinson SJ, Djavaheri M, Malinowski R, Haddadi P, Kagale S, Parkin IAP, Taheri A, Borhan MH. 2016. The compact genome of the plant pathogen Plasmodiophora brassicae is adapted to intracellular interactions with host Brassica spp. BMC Genomics 17:272. https://doi.org/10.1186/s12864-016-2597-2.
8. Schwelm A, Fogelqvist J, Knaust A, Jülke S, Lilja T, Bonilla-Rosso G, Karlsson M, Shevchenko A, Dhandapani V, Choi SR, Kim HG, Park JY, Lim YP, Ludwig-Müller J, Dixelius C. 2015. The Plasmodiophora brassicae genome reveals insights in its life cycle and ancestry of chitin synthases. Sci Rep 5:11153. https://doi.org/10.1038/srep11153.
9. Schwelm A, Dixelius C, Ludwig-Müller J. 2016. New kid on the block—the clubroot pathogen genome moves the plasmodiophorids into the genomic era. Eur J Plant Pathol 145:531–542. https://doi.org/10.1007/s10658-015-0839-9.
10. Dohm JC, Minoche AE, Holtgräwe D, Capella-Gutiérrez S, Zakrzewski F, Tafer H, Rupp O, Sörensen TR, Stracke R, Reinhardt R, Goesmann A, Kraft T, Schulz B, Stadler PF, Schmidt T, Gabaldón T, Lehrach H, Weisshaar B, Himmelbauer H. 2014. The genome of the recently domesticated crop plant sugar beet (Beta vulgaris). Nature 505:546–549. https://doi.org/10.1038/nature12817.
11. Legrève A, Delfosse P, Vanpee B, Goffin A, Maraite H. 1998. Differences in temperature requirements between Polymyxa sp. of Indian origin and Polymyxa graminis and Polymyxa betae from temperate areas. Eur J Plant Pathol 104:195–205. https://doi.org/10.1023/A:1008612903927.
12. Brandfass C, Karlovsky P. 2008. Upscaled CTAB-based DNA extraction and real-time PCR assays for Fusarium culmorum and F. graminearum DNA in plant material with reduced sampling error. Int J Mol Sci 9:2306–2321. https://doi.org/10.3390/ijms9112306.
13. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170.
14. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. PeerJ 3:e1165. https://doi.org/10.7717/peerj.1165.
15. 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 25:1043–1055. https://doi.org/10.1101/gr.186072.114.
16. Broeksema B, Calusinska M, McGee F, Winter K, Bongiovanni F, Goux X, Wilmes P, Delfosse P, Ghoniem M. 2017. ICoVeR—an interactive visualization tool for verification and refinement of metagenomic bins. BMC Bioinformatics 18:233. https://doi.org/10.1186/s12859-017-1653-5.
17. Parra G, Bradnam K, Korf I. 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. Bioinformatics 23:1061–1067. https://doi.org/10.1093/bioinformatics/btm071.
18. Marçais G, Kingsford C. 2011. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27:764–770. https://doi.org/10.1093/bioinformatics/btr011.