# High-Quality Draft Genome Sequence of *Kibdelosporangium philippinense*, Generated by Hybrid Assembly of Short and Long Sequencing Reads

Eugenia A. Fedorov,[a] Medina Omeragic,[a] Kristina F. Shalygina,[a] Ashlyn C. Farwell,[a] ⓘ Kyle S. MacLea[a,b,c]

[a]Biology Program, Department of Life Sciences, University of New Hampshire, Manchester, New Hampshire, USA
[b]Biotechnology Program, Department of Life Sciences, University of New Hampshire, Manchester, New Hampshire, USA
[c]Department of Life Sciences, University of New Hampshire, Manchester, New Hampshire, USA

**ABSTRACT** The glycopeptide antibiotic-producing soil actinobacterium *Kibdelosporangium philippinense* A80407 (=ATCC 49844) was sequenced using Illumina and Nanopore sequencing methodologies, and a hybrid genome assembly was generated for this type strain, with a total predicted genome length of 12,054,556 bp, 10,953 protein-coding sequences, 79 RNAs, 298 pseudogenes, and a G+C content of 65.13%.

*K*ibdelosporangium philippinense A80407$^T$ (=ATCC 49844$^T$ = NRRL 18198$^T$) is an unsequenced glycopeptide antibiotic-producing actinobacterium (1). Of the sequenced members of its genus, only *Kibdelosporangium phytohabitans* has a high-quality genome sequence available. Since *Kibdelosporangium* spp. produce numerous interesting antimicrobials (2–8) and have large genomes, averaging about 12 Mbp, sequencing other members of the genus is a priority.

For this study, freeze-dried *K. philippinense* ATCC 49844$^T$ (ATCC, Manassas, VA, USA) was rehydrated and subcultured on International Streptomyces Project 2 (ISP2) medium (BD, Franklin Lakes, NJ, USA), with a single colony grown in ISP2 broth (Teknova, Hollister, CA, USA) at 30°C/1 atm for 240 h. Genomic DNA (gDNA) was isolated using the FastDNA Spin kit for soil (matrix A; MP Biochemicals, Irvine, CA, USA) for Illumina sequencing and the Nanobind CBB Big DNA kit (Circulomics, Baltimore, MD, USA) for Nanopore sequencing. The KAPA HyperPlus kit (KR1145, v.3.16; Wilmington, MA, USA) was used to generate the Illumina sequencing library. The Hubbard Center for Genome Studies (HCGS; Durham, NH, USA) sequenced this library on a HiSeq 2500 instrument, producing 250-bp paired-end fragments. Trimmomatic v.0.38 (9) was used to trim the resulting reads (settings: paired-end mode with a window size of 4, quality requirement of 15, and minimum read length of 36). A long-read Nanopore library was generated using the ligation sequencing kit (LSK109; ONT, Oxford, UK), run on a MinION R9.4.1 flow cell/GridION instrument (HCGS), and base called using the Guppy v.5.0.13 base caller (10) in SUP mode. Adapters were trimmed using Porechop v.0.2.4 (read setting: 1000; the Nanopore reads as analyzed by QUAST [11], yielded an $N_{50}$ value of 17,277 bp), and the reads were filtered using Filtlong v.0.2.1 (settings, top 80% quality and length 1,000 bp). SPAdes v.3.13.0 (12) was used to assemble 2,926,840 trimmed short reads and 239,141 long reads with default bacterial assembly parameters in hybrid mode. Small contigs (<200 bp) and contigs with low coverage (<9.5×) were removed.

The assembly contained 27 contigs—the largest being 4,055,460 bp—with an $N_{50}$ value of 2,839,569 bp (11). More than 94% of the genome size (12,054,556 bp) is comprised by just 4 contigs, and only 7 contigs of the assembly are larger than 500 bp. However, 200- to 500-bp contigs matching (via BLAST [13]) closely related species (*Kibdelosporangium* sp. strain MJ126-NF4 [GenBank accession number GCF_000826545.1],

*Amycolatopsis aidingensis* [GCF_018885265.1], *Amycolatopsis* sp. strain CA-230715 [GCF_018736145.1], *Saccharopolyspora pogona* [GCF_014697215.1], *Saccharopolyspora erythraea* [GCF_002564065.1], and *Saccharopolyspora spinosa* [GCF_014490055.1]) were retained. It is unclear if the 4 largest contigs represent chromosomes; however, one of the *Kibdelosporangium* sp. genomes sequenced (*Kibdelosporangium phytohabitans*) does have a hybrid Illumina-PacBio assembly (GCF_001302585.1) that resolves into a single 11.75-Mbp chromosome. Therefore, we expect that these contigs represent large chunks of a single chromosome. We also calculated a G+C content of 65.13%. Although there is no previously published value, our calculation is close to that of the type member of the genus, *Kibdelosporangium aridum*, as determined chemically (66% ([2]) or *in silico* (66.2% [14]).

The genome is 99.1% complete according to an analysis using Benchmarking Universal Single-Copy Orthologs (BUSCO) v.5.2.2 (default parameters) (15), with an average genome coverage of 60.72×. The Prokaryotic Genome Assembly Pipeline (PGAP) (16) annotated 10,953 protein-coding genes, 298 pseudogenes, 79 RNAs, including 3 complete copies of the rRNA genes (also seen in the single-chromosome *K. phytohabitans* assembly above), 3 noncoding RNAs (ncRNAs), and 67 tRNAs. *Kibdelosporangium* spp. produce antimicrobials, and their genomes contain numerous biosynthetic gene clusters (BGCs) related to these metabolic pathways (1–6, 8, 14, 17). Forty-two BGCs were found in *K. philippinense* using antiSMASH 6.0 (18), including 26 nonribosomal peptide synthetase and/or polyketide clusters.

**Data availability.** The *Kibdelosporangium philippinense* ATCC 49844[T] whole-genome shotgun sequence (WGS) project has been deposited at DDBJ/ENA/GenBank under accession number JAJVCN000000000. The raw data, found under BioProject accession number PRJNA790681, were submitted to the NCBI Sequence Read Archive (SRA) under three experiment accession numbers: SRX13551729 (Illumina fastq files), SRX13555714 (Nanopore reads in fastq format), and SRX13556647 (Nanopore reads in fast5 format).

## REFERENCES

1. Mertz FP, Yao RCY. 1988. Kibdelosporangium philippinense sp. nov. isolated from soil. Int J Syst Evol Microbiol 38:282–286. https://doi.org/10.1099/00207713-38-3-282.

2. Shearer MC, Colman PM, Ferrin RM, Nisbet LJ, Nash CH. 1986. New genus of the Actinomycetales: Kibdelosporangium aridum gen. nov., sp. nov. Int J Syst Evol Microbiol 36:47–54. https://doi.org/10.1099/00207713-36-1-47.

3. Shearer MC, Giovenella AJ, Grappel SF, Hedde RD, Mehta RJ, Oh YK, Pan CH, Pitkin DH, Nisbet LJ. 1986. Kibdelins, novel glycopeptide antibiotics I. Discovery, production, and biological evaluation. J Antibiot (Tokyo) 39:1386–1394. https://doi.org/10.7164/antibiotics.39.1386.

4. Doolin LE, Gale RM, Godfrey OW, Hamill RL, Mahoney DF, Yao RC. 1993. A80407 antibiotics. 25 May 1993. US patent 5213797A.

5. Singh SB, Dayananth P, Balibar CJ, Garlisi CG, Lu J, Kishii R, Takei M, Fukuda Y, Ha S, Young K. 2015. Kibdelomycin is a bactericidal broad-spectrum aerobic antibacterial agent. Antimicrob Agents Chemother 59:3474–3481. https://doi.org/10.1128/AAC.00382-15.

6. Ogasawara Y, Torrez-Martinez N, Aragon AD, Yackley BJ, Weber JA, Sundararajan A, Ramaraj T, Edwards JS, Melançon CE. 2015. High-quality draft genome sequence of Actinobacterium kibdelosporangium sp. MJ126-NF4, producer of type II polyketide azicemicins, using Illumina and PacBio technologies. Genome Announc 3:e00114-15. https://doi.org/10.1128/genomeA.00114-15.

7. Mingma R, Duangmal K, Ōmura S, Takahashi Y, Matsumoto A. 2017. Three novel species of the genus Kibdelosporangium; Kibdelosporangium kanagawaense sp. nov., Kibdelosporangium rhizosphaerae sp. nov. and Kibdelosporangium rhizovicinum sp. nov. Int J Syst Evol Microbiol 67:1758–1765. https://doi.org/10.1099/ijsem.0.001861.

8. Safaei N, Nouioui I, Mast Y, Zaburannyi N, Rohde M, Schumann P, Müller R, Wink J. 2021. Kibdelosporangium persicum sp. nov., a new member of

the Actinomycetes from a hot desert in Iran. Int J Syst Evol Microbiol 71. https://doi.org/10.1099/ijsem.0.004625.

9. 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.

10. Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. Genome Biol 20:129. https://doi.org/10.1186/s13059-019-1727-y.

11. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075. https://doi.org/10.1093/bioinformatics/btt086.

12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477. https://doi.org/10.1089/cmb.2012.0021.

13. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410. https://doi.org/10.1016/S0022-2836(05)80360-2.

14. Ju K-S, Gao J, Doroghazi JR, Wang K-KA, Thibodeaux CJ, Li S, Metzger E,

Fudala J, Su J, Zhang JK, Lee J, Cioni JP, Evans BS, Hirota R, Labeda DP, van der Donk WA, Metcalf WW. 2015. Discovery of phosphonic acid natural products by mining the genomes of 10,000 actinomycetes. Proc Natl Acad Sci U S A 112:12175–12180. https://doi.org/10.1073/pnas.1500873112.

15. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212. https://doi.org/10.1093/bioinformatics/btv351.

16. Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Ciufo S, Li W. 2013. Prokaryotic Genome Annotation Pipeline. https://www.ncbi.nlm.nih.gov/books/NBK174280/.

17. Wei B, Du A-Q, Zhou Z-Y, Lai C, Yu W-C, Yu J-B, Yu Y-L, Chen J-W, Zhang H-W, Xu X-W, Wang H. 2021. An atlas of bacterial secondary metabolite biosynthesis gene clusters. Environ Microbiol 23:6981–6992. https://doi.org/10.1111/1462-2920.15761.

18. Blin K, Shaw S, Kloosterman AM, Charlop-Powers Z, van Wezel GP, Medema MH, Weber T. 2021. antiSMASH 6.0: improving cluster detection and comparison capabilities. Nucleic Acids Res 49:W29–W35. https://doi.org/10.1093/nar/gkab335.