





Draft Genome Sequences of New Isolates and the Known Species of the Family *Microbacteriaceae* Associated with Plants

 Oleg V. Vasilenko,^a Irina P. Starodumova,^a Lubov V. Dorofeeva,^a Sergey V. Tarlachkov,^{a,b} Natalia V. Prisyazhnaya,^a Vladimir N. Chizhov,^c Sergei A. Subbotin,^d Marcel Huntemann,^e Alicia Clum,^e Kecia Duffy,^e Manoj Pillay,^e Krishnaveni Palaniappan,^e Neha Varghese,^e I-Min A. Chen,^e Dimitrios Stamatis,^e  T. B. K. Reddy,^e Ronan O'Malley,^e Chris Daum,^e Nicole Shapiro,^e Natalia Ivanova,^e Nikos C. Kyrpides,^e Tanja Woyke,^e William B. Whitman,^f Lyudmila I. Evtushenko^a

^aAll-Russian Collection of Microorganisms (VKM), G. K. Skryabin Institute of Biochemistry and Physiology of Microorganisms, Pushchino, Russia

^bBranch of Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Pushchino, Russia

^cCenter of Parasitology, A. N. Severtsov Institute of Ecology and Evolution, Moscow, Russia

^dCalifornia Department of Food and Agriculture, Sacramento, California, USA

^eDOE Joint Genome Institute, Walnut Creek, California, USA

^fDepartment of Microbiology, University of Georgia, Athens, Georgia, USA

ABSTRACT Draft genome sequences of 11 bacteria belonging to the family *Microbacteriaceae* were obtained using Illumina technology. The genomes of these strains have sizes from 3.14 to 4.30 Mb with their genomic DNA characterized as having high G+C contents (above 65%). These genomic data will be useful for natural taxonomy and comparative genomic studies of bacterial strains of the family *Microbacteriaceae*.

The family *Microbacteriaceae* Park et al. 1995 emend. Stackebrandt et al. 1997 emend. Zhi et al. 2009 comprises a large group of predominantly aerobic Gram-positive bacteria with high G+C contents in their genomic DNA that are distinguished from other actinobacteria by a combination of their unusual group B cell wall peptidoglycan and unsaturated respiratory menaquinones (1, 2). Currently, the family comprises 51 recognized genera and about 275 recognized species (<http://www.bacterio.net/microbacteriaceae.html>). There are about 80 sequenced genomes for type strains of the family *Microbacteriaceae* (<https://www.ezbiocloud.net/taxon?tn=Microbacteriaceae>). Efforts to obtain complete genome sequences of all type strains belonging to the family *Microbacteriaceae* are critical to the future microbial systematics of this group and for constructing the unified tree of life.

The type strains of six known species of the genera *Agreia*, *Clavibacter*, *Okibacterium*, *Plantibacter*, and *Rathayibacter* were isolated from various sources associated with plants (Table 1) (3–7). In addition, the novel strains *Agreia* sp. VKM Ac-1783, *Agreia* sp. VKM Ac-2052, *Plantibacter* sp. VKM Ac-1784, *Plantibacter* sp. VKM Ac-1787, and *Rathayibacter oskolensis* VKM Ac-2121^T were isolated from plant galls induced by nematodes and from the healthy plant without any symptoms of disease (Table 1) (8–10). All strains were deposited in the All-Russian Collection of Microorganisms (<http://www.vkm.ru/Catalogue.htm>).

Biomass for DNA extraction was obtained after growth in liquid peptone-yeast medium (11) at 28°C for 18 to 20 h on a rotary shaker. The genomic DNA was extracted using guanidinium thiocyanate and Triton X-100, followed by purification with Cleanup Standard BC022 (Evrogen, Russia). The draft genome sequences of the strains were generated at the DOE Joint Genome Institute (JGI) (Walnut Creek, CA, USA). The Illumina shotgun libraries with an insert size of 262 to 287 bp were constructed and

Received 27 July 2018 Accepted 20 August 2018 Published 20 September 2018

Citation Vasilenko OV, Starodumova IP, Dorofeeva LV, Tarlachkov SV, Prisyazhnaya NV, Chizhov VN, Subbotin SA, Huntemann M, Clum A, Duffy K, Pillay M, Palaniappan K, Varghese N, Chen I-MA, Stamatis D, Reddy TBK, O'Malley R, Daum C, Shapiro N, Ivanova N, Kyrpides NC, Woyke T, Whitman WB, Evtushenko LI. 2018. Draft genome sequences of new isolates and the known species of the family *Microbacteriaceae* associated with plants. *Microbiol Resour Announc* 7:e01051-18. <https://doi.org/10.1128/MRA.01051-18>.

Editor Catherine Putonti, Loyola University Chicago

This is a work of the U.S. Government and is not subject to copyright protection in the United States. Foreign copyrights may apply.

Address correspondence to Oleg V. Vasilenko, ovvasilenko@gmail.com, or Irina P. Starodumova, iri-starodumova@yandex.ru. O.V.V. and I.P.S. contributed equally to this work.

TABLE 1 Genome sequence statistical information and DDBJ/ENA/GenBank accession numbers

Organism	Plant source	Nematode	Total no. of reads post-trimming (Mb)	Total size post-trimming (Mb)	No. of contigs (scaffolds)	Coverage (X)	N ₅₀ (bp)	Genome size (Mb)	G+C content (%)	No. of proteins	SRA no.	GenBank accession no.
<i>A. pratensis</i> VKM Ac-2510 ^T	Phyllosphere of grasses	No	5,153,882	773.1	10	214	860,306	3.58	65.4	3,418	SRR5832269	FXAY000000000
<i>Agreia</i> sp. VKM Ac-1783	<i>Elymus repens</i>	<i>Anguina</i>	5,567,834	835.2	7	220	1,002,356	3.76	65.8	3,551	SRR5832516	FXWI000000000
<i>Agreia</i> sp. VKM Ac-2052	<i>Calamagrostis neglecta</i>	<i>Heteroanguina agropyri</i>	7,310,900	1,096.6	19	280	442,588	3.88	65.2	3,654	SRR5832511	FUYG000000000
<i>C. michiganensis</i> subsp. <i>michiganensis</i> VKM Ac-1403 ^T	<i>Lycopersicon esculentum</i>	No	6,669,630	993.8	14 (13)	300	2,018,315	3.30	72.7	3,045	SRR5216688	FVZG000000000
<i>O. fritillariae</i> VKM Ac-2059 ^T	<i>Fritillaria ruthenica</i>	No	5,278,668	791.8	4	250	2,068,692	3.14	67.8	2,808	SRR5832512	FUZP000000000
<i>P. flavus</i> VKM Ac-2504 ^T	Phyllosphere of grasses	No	4,884,328	732.6	12 (11)	169	544,526	4.30	69.1	4,002	SRR5832515	FXAP000000000
<i>Plantibacter</i> sp. VKM Ac-1784	<i>E. repens</i>	<i>A. agropyri</i>	5,096,142	764.4	8	184	674,883	4.12	69.2	3,829	SRR5832517	FXWJ000000000
<i>Plantibacter</i> sp. VKM Ac-1787	<i>Cousinia onopordioides</i>	<i>Mesoanguina picridis</i>	5,710,806	850.9	5	211	2,657,854	4.02	69.5	3,719	SRR5832510	FUZO000000000
<i>R. rathayi</i> VKM Ac-1601 ^T	<i>Dactylis glomerata</i>	<i>Anguina</i> sp.	5,193,410	779.0	46 (45)	244	321,699	3.22	69.3	3,126	SRR5832513	OCNL000000000
<i>R. iranicus</i> VKM Ac-1602 ^T	<i>Triticum aestivum</i>	<i>A. tritici</i>	5,315,024	797.3	50 (49)	236	188,097	3.38	67.2	3,268	SRR5832514	QGDV000000000
<i>R. oskolensis</i> VKM Ac-2121 ^T	<i>Androsace koso-poljanskii</i>	No	5,903,488	879.6	5	222	1,280,863	3.95	71.6	3,628	SRR5832669	FXBM000000000

sequenced using the Illumina HiSeq 2000 platform to produce 2×150 -bp paired-end reads. All raw Illumina sequence data were filtered using BBDuk (<http://sourceforge.net/projects/bbmap>), which removes known Illumina artifacts and PhiX sequences. Reads with more than one N, quality scores (before trimming) averaging less than 8, or reads shorter than 51 bp (after trimming) were discarded. The remaining reads were mapped to masked versions of human, cat, and dog references using BBMAP (<http://sourceforge.net/projects/bbmap>) and discarded if their identities exceeded 95%. Sequence masking was performed with BBMask (<http://sourceforge.net/projects/bbmap>). The following steps were then performed for assembly: (i) artifact-filtered Illumina reads were assembled using SPAdes version 3.6.2 (with the parameters “--cov-cutoff auto --phred-offset 33 -t 8 -m 40 --careful -k 25,55,95 --12”) (12); and (2) assembly contigs with lengths of <1 kb were discarded. The genome was annotated using the JGI Microbial Genome Annotation Pipeline (13). Statistical information for the draft genome sequences is given in Table 1.

Further genome-wide comparative analyses of these strains and phylogenetically closely related bacteria belonging to the family *Microbacteriaceae* will enhance an understanding of the borderlines of prokaryotic species and facilitate insight into the molecular mechanisms involved in interactions between plants and bacteria.

Data availability. The whole-genome shotgun projects reported here have been deposited in DDBJ/ENA/GenBank under the accession numbers listed in Table 1. The versions reported here are the first versions.

ACKNOWLEDGMENTS

These sequence data were produced by the U.S. Department of Energy Joint Genome Institute (DOE JGI) (<http://www.jgi.doe.gov>) in collaboration with the All-Russian Collection of Microorganisms (VKM) (<http://www.vkm.ru/index.htm>) in the frame of the Genomic Encyclopedia of Bacteria and Archaea Project Phase III (GEBA III) (14). The employees of DOE JGI, VKM, and the Department of Microbiology, University of Georgia, are acknowledged for their support of this global project.

The genome sequencing was supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. The reported study was also partially funded by RFBR through research project 16-34-01048mol_a (N.V.P.) and the Russian Federation Federal Agency of Scientific Organizations' Bioresource Collections Program (L.I.E.).

REFERENCES

- Evtushenko LI. 2012. Family XI. *Microbacteriaceae* Park, Suzuki, Yim, Lee, Kim, Yoon, Kim, Kho, Goodfellow and Komagata 1995, 418^{VP} (Effective publication: Park, Suzuki, Yim, Lee, Kim, Yoon, Kim, Kho, Goodfellow and Komagata 1993, 312.) emend. Rainey, Ward-Rainey and Stackebrandt 1997, 485, p 807–813. In Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K-I, Ludwig W, Whitman WB (ed), *Bergey's manual of systematic bacteriology*, 2nd ed, vol 5. Springer, New York, NY.
- Evtushenko LI, Takeuchi M. 2006. The family *Microbacteriaceae*, p 1020–1098. In Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E (ed), *The prokaryotes: a handbook on the biology of bacteria. Archaea, Bacteria, Firmicutes, Actinomycetes*, 3rd ed, vol 3. Springer, New York, NY.
- Behrendt U, Ulrich A, Schumann P, Naumann D, Suzuki K. 2002. Diversity of grass-associated *Microbacteriaceae* isolated from the phyllosphere and litter layer after mulching the sward; polyphasic characterization of *Subtercola pratensis* sp. nov., *Curtobacterium herbarum* sp. nov. and *Plantibacter flavus* gen. nov., sp. nov. *Int J Syst Evol Microbiol* 52: 1441–1454. <https://doi.org/10.1099/00207713-52-5-1441>.
- Davis MJ, Gillaspie AG, Vidaver AK, Harris RW. 1984. *Clavibacter*: a new genus containing some phytopathogenic coryneform bacteria, including *Clavibacter xyli* subsp. *xyli* sp. nov., subsp. nov. and *Clavibacter xyli* subsp. *cyndodontis* subsp. nov., pathogens that cause ratoon stunting disease of sugarcane and bermudagrass stunting disease. *Int J Syst Bacteriol* 34: 107–117. <https://doi.org/10.1099/00207713-34-2-107>.
- Evtushenko LI, Dorofeeva LV, Krausova VI, Gavrish EY, Yashina SG, Takeuchi M. 2002. *Okibacterium fritillariae* gen. nov., sp. nov., a novel genus of the family *Microbacteriaceae*. *Int J Syst Evol Microbiol* 52:987–993. <https://doi.org/10.1099/00207713-52-3-987>.
- Schumann P, Behrendt U, Ulrich A, Suzuki K. 2003. Reclassification of *Subtercola pratensis* Behrendt et al. 2002 as *Agreia pratensis* comb. nov. *Int J Syst Evol Microbiol* 53:2041–2044. <https://doi.org/10.1099/ijs.0.02664-0>.
- Zgurskaya HI, Evtushenko LI, Akimov VN, Kalakoutskii LV. 1993. *Rathayibacter* gen. nov., including the species *Rathayibacter rathayi* comb. nov., *Rathayibacter tritici* comb. nov., *Rathayibacter iranicus* comb. nov., and six strains from annual grasses. *Int J Syst Bacteriol* 43:143–149. <https://doi.org/10.1099/00207713-43-1-143>.
- Behrendt U, Ulrich A, Schumann P. 2012. Genus XX. *Plantibacter* Behrendt, Ulrich, Schumann, Naumann and Suzuki 2002, 1451^{VP}, p 943–949. In Goodfellow M, Kämpfer P, Busse H-J, Trujillo ME, Suzuki K-I, Ludwig W, Whitman WB (ed), *Bergey's manual of systematic bacteriology*, 2nd ed, vol 5. Springer, New York, NY.
- Dorofeeva LV, Starodumova IP, Krausova VI, Prisyazhnaya NV, Vinokurova NG, Lysanskaya VY, Tarlachkov SV, Evtushenko LI. 2018. *Rathayibacter oskolensis* sp. nov., a novel actinobacterium from *Androsace kosopoljanskii* ovcz. (Primulaceae) endemic to the central Russian upland. *Int J Syst Evol Microbiol* 68:1442–1447. <https://doi.org/10.1099/ijsem.0.002681>.
- Starodumova IP, Prisyazhnaya NV, Dorofeeva LV, Evtushenko LI. 2015. Novel actinobacteria of the genus *Agreia* isolated from plant galls in-

- duced by plant-parasitic nematodes [in Russian]. *Russian J Immunol* 9:613–615.
11. Naumova IB, Kuznetsov VD, Kudrina KS, Bezzubenkova AP. 1980. The occurrence of teichoic acids in streptomycetes. *Arch Microbiol* 126: 71–75. <https://doi.org/10.1007/BF00421893>.
 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. Mavromatis K, Ivanova NN, Chen I-MA, Szeto E, Markowitz VM, Kyrpides NC. 2009. The DOE-JGI standard operating procedure for the annotations of microbial genomes. *Stand Genomic Sci* 1:63–67. <https://doi.org/10.4056/sigs.632>.
 14. Whitman WB, Woyke T, Klenk H-P, Zhou Y, Lilburn TG, Beck BJ, De Vos P, Vandamme P, Eisen JA, Garrity G, Hugenholtz P, Kyrpides NC. 2015. Genomic encyclopedia of bacterial and archaeal type strains, phase III: the genomes of soil and plant-associated and newly described type strains. *Stand Genomic Sci* 10:26. <https://doi.org/10.1186/s40793-015-0017-x>.