








Complete Genome Sequences of 44 *Arthrobacter* Phages

Karen K. Klyczek,^a  Deborah Jacobs-Sera,^b Tamarah L. Adair,^c Sandra D. Adams,^d Sarah L. Ball,^e Robert C. Benjamin,^f J. Alfred Bonilla,^a Caroline A. Breitenberger,^e Charles J. Daniels,^e Bobby L. Gaffney,^h Melinda Harrison,ⁱ Lee E. Hughes,^f Rodney A. King,^h Gregory P. Krukonijs,^j A. Javier Lopez,^g Kirsten Monsen-Collar,^d Marie C. Pizzorno,^j  Claire A. Rinehart,^h SEA-PHAGES Program, PHIRE Program, Amanda K. Staples,^h Emily L. Stowe,^j Rebecca A. Garlena,^b Daniel A. Russell,^b  Steven G. Cresawn,^k  Welkin H. Pope,^b  Graham F. Hatfull^b

^aDepartment of Biology, University of Wisconsin—River Falls, River Falls, Wisconsin, USA

^bDepartment of Biological Sciences, University of Pittsburgh, Pittsburgh, Pennsylvania, USA

^cDepartment of Biology, Baylor University, Waco, Texas, USA

^dDepartment of Biology, Montclair State University, Montclair, New Jersey, USA

^eCenter for Life Sciences Education, The Ohio State University, Columbus, Ohio, USA

^fDepartment of Biological Sciences, University of North Texas, Denton, Texas, USA

^gDepartment of Biological Sciences, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

^hDepartment of Biology, Western Kentucky University, Bowling Green, Kentucky, USA

ⁱDepartment of Science, Cabrini University, Radnor, Pennsylvania, USA

^jBiology Department, Bucknell University, Lewisburg, Pennsylvania, USA

^kDepartment of Biology, James Madison University, Harrisonburg, Virginia, USA

ABSTRACT We report here the complete genome sequences of 44 phages infecting *Arthrobacter* sp. strain ATCC 21022. These phages have double-stranded DNA genomes with sizes ranging from 15,680 to 70,707 bp and G+C contents from 45.1% to 68.5%. All three tail types (belonging to the families *Siphoviridae*, *Myoviridae*, and *Podoviridae*) are represented.

Bacteriophages are the most abundant biological entities in the biosphere and represent a large reservoir of undiscovered genetic information (1). Genome comparisons of 46 phages infecting a single *Arthrobacter* host revealed extensive genetic diversity (2), with phage clusters distinctly separated from phages isolated on *Mycobacterium smegmatis* (3). Here, we report the genome sequences of 44 additional phages isolated on *Arthrobacter* sp. strain ATCC 21022 (4). These phages were recovered from filtered soil samples by either direct plating (phages Temper16 and Wheelbite) or enrichment (the other 42 phages) and originate from diverse geographical locations (Table 1). Electron microscopy showed that 30 of the phages belong to the family *Siphoviridae*, 7 with prolate heads (cluster AM) and 23 with isometric capsids (Table 1). Eleven phages belong to the family *Myoviridae* and three to the family *Podoviridae* (Table 1).

Genomic DNA was isolated from plaque-purified phages and sequenced using the Illumina MiSeq platform. The 150-bp single-end reads were assembled into complete genomes using Newbler and Consed, with at least 180-fold coverage. Genome lengths range from 15,680 bp to 70,707 bp, with G+C contents from 45.1% to 68.5% (Table 1), compared to the host genome's G+C content of 63.4% (4). Thirty of the genomes have defined termini, with either 3' single-stranded DNA extensions or long (~1.5-kb) direct terminal repeats; the other 14 genomes are circularly permuted, and the first coordinate was assigned near the start of the predicted terminase gene (Table 1). Genomes were annotated using DNAMaster (<http://cobamide2.bio.pitt.edu>), and coding sequences (between 26 and 114) were predicted using GeneMark (5) and Glimmer (6). Putative protein functions were determined using BLASTp (7) and HHpred (8). Molivia

Received 28 November 2017 **Accepted** 14 December 2017 **Published** 1 February 2018

Citation Klyczek KK, Jacobs-Sera D, Adair TL, Adams SD, Ball SL, Benjamin RC, Bonilla JA, Breitenberger CA, Daniels CJ, Gaffney BL, Harrison M, Hughes LE, King RA, Krukonijs GP, Lopez AJ, Monsen-Collar K, Pizzorno MC, Rinehart CA, SEA-PHAGES Program, PHIRE Program, Staples AK, Stowe EL, Garlena RA, Russell DA, Cresawn SG, Pope WH, Hatfull GF. 2018. Complete genome sequences of 44 *Arthrobacter* phages. *Genome Announc* 6:e01474-17. <https://doi.org/10.1128/genomeA.01474-17>.

Copyright © 2018 Klyczek 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 Graham F. Hatfull, gfh@pitt.edu.

TABLE 1 Properties of 44 *Arthrobacter* phages

Phage name	Cluster	GenBank accession no.	Genome length (bp)	G+C content (%)	No. of CDSs ^a	Genome end type ^b	Virion tail morphology	Virion head morphology	Location of isolation
Canowicakte	AK	MF140400	43,914	61.2	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Radnor, PA
Christian	AK	MF140404	43,081	60.7	61	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	West Deptford, NJ
Dino	AK	MF140407	43,562	61.1	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Lewisburg, PA
Greenhouse	AK	KX688103	43,977	60.8	61	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Denton, TX
Huntingdon	AK	MG210949	43,891	60.7	61	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Huntingdon Valley, PA
Lucy	AK	KX576641	42,944	60.7	60	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Immaculata, PA
MeganNoll	AK	MG198782	44,258	60.9	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Pittsburgh, PA
Nubia	AK	MF140424	44,045	60.7	61	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Waco, TX
Oxynfrius	AK	KX688102	44,163	60.8	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Denton, TX
PitaDog	AK	MF140425	42,963	60.7	61	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	New Castle, DE
RcigaStruga	AK	KX576640	43,891	60.7	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Radnor, PA
Suppi	AK	KX621004	43,914	61.2	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Mays Landing, NJ
Temper16	AK	MF668285	43,950	61.9	60	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Columbus, OH
Urla	AK	MG198779	43,940	61	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Pittsburgh, PA
Vallejo	AK	KX621005	43,607	61	62	13-base 3' ext.	<i>Siphoviridae</i>	Isometric	Radnor, PA
LiSara	AL	MF140418	60,137	64.7	96	Circ. perm.	<i>Siphoviridae</i>	Isometric	Lewisburg, PA
Shrooms	AL	MF140427	59,707	64.5	98	Circ. perm.	<i>Siphoviridae</i>	Isometric	Aquilla, TX
Wheelbite	AL	MF140434	59,879	64.8	94	Circ. perm.	<i>Siphoviridae</i>	Isometric	Argyle, TX
Arcadia	AM	MF189170	58,059	45.2	97	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Radnor, PA
Cheesy	AM	MF324906	58,739	45.2	101	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	River Falls, WI
Correa	AM	MF189171	57,401	45.2	96	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Radnor, PA
Elsa	AM	MF189172	58,059	45.2	97	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Exton, PA
Heisenberger	AM	MF189173	58,208	45.1	100	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Owensboro, KY
Nason	AM	MF189174	58,059	45.1	97	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Radnor, PA
Tribby	AM	MF189175	59,084	45.2	102	9-base 3' ext.	<i>Siphoviridae</i>	Prolate	Lewisburg, PA
Swenson	AN	MF140429	15,680	59.9	26	11-base 3' ext.	<i>Siphoviridae</i>	Isometric	Pittsburgh, PA
Molivia	AQ	MF185731	58,247	53.5	98	1,467-bp DTR	<i>Siphoviridae</i>	Isometric	Radnor, PA
Chocolat	AR	KX670787	69,798	61.7	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Philadelphia, PA
Chubster	AR	KX670786	70,258	61.7	112	Circ. perm.	<i>Myoviridae</i>	Isometric	Lansdowne, PA
Colucci	AR	MF185718	70,707	61.7	114	Circ. perm.	<i>Myoviridae</i>	Isometric	Mississauga, ON
Conboy	AR	KX522650	70,096	61.7	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Radnor, PA
EdgarPoe	AR	KX855961	70,176	61.6	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Philadelphia, PA
HumptyDumpty	AR	KX855962	69,978	61.7	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Radnor, PA
JayCookie	AR	MF668274	70,362	61.7	112	Circ. perm.	<i>Myoviridae</i>	Isometric	River Falls, WI
Kabreeze	AR	MF185721	70,035	61.7	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Radnor, PA
RosiePosie	AR	MF185723	70,396	61.7	112	Circ. perm.	<i>Myoviridae</i>	Isometric	Immaculata, PA
Scavito	AR	MF185724	70,123	61.6	112	Circ. perm.	<i>Myoviridae</i>	Isometric	West Chester, PA
Tophat	AR	MF185725	70,091	61.6	111	Circ. perm.	<i>Myoviridae</i>	Isometric	Lewisburg, PA
Abidatro	AS	MF140397	39,122	68.5	66	12-base 3' ext.	<i>Siphoviridae</i>	Isometric	Montclair, NJ
ElephantMan	AU	MF038791	58,405	49.9	93	9-base 3' ext.	<i>Siphoviridae</i>	Isometric	Columbus, OH
Niktson	AU	MF038790	58,405	49.9	93	9-base 3' ext.	<i>Siphoviridae</i>	Isometric	Columbus, OH
Adat	AV	MF668266	45,428	45.7	56	1,551-bp DTR	<i>Podoviridae</i>	Isometric	Columbus, OH
GurgleFerb	AV	MF668273	45,426	45.7	56	1,550-bp DTR	<i>Podoviridae</i>	Isometric	Columbus, OH
Nellie	AV	MF668279	45,428	45.8	57	1,551-bp DTR	<i>Podoviridae</i>	Isometric	Columbus, OH

^aCDSs, protein-coding sequences.^bext., extension; DTR, direct terminal repeat; Circ. perm., circularly permuted.

is the only genome with tRNA genes, and six were predicted by Aragorn (9) and tRNAscan-SE (10).

All of the newly isolated phages share sequence similarity to various extents with previously described *Arthrobacter* phages and warrant either inclusion into extant clusters (AK, AL, AM, AN, AQ, AR, and AU) or the formation of new clusters (AS and AV) by joining with previously designated singleton genomes (Table 1). Cluster AS is formed by the similarity of phage Abidatro to the former singleton Galaxy, and they share 93.4% average nucleotide identity spanning ~90% of their genome lengths; both encode integrase and immunity repressor proteins and are likely temperate. The *attP* sites are unusually displaced over 2 kb from the integrase genes, and the intervening region is variable between the two genomes. None of the other phages have genes associated with lysogeny and are predicted to be lytic.

Phages Adat, GurgleFerb, and Nellie share >90% average nucleotide identity spanning ~84% of their genome lengths with Jasmine, forming cluster AV. All of these

phages have a podoviral morphology and are the only *Podoviridae* members described among over 5,000 morphologically examined actinobacteriophages (11). However, phages Adat, GurgleFerb, and Nellie lack the putative tail spike protein and lipolytic esterase genes of Jasmine (18 and 19), which are replaced with a putative rhamnoga-lacturonan acetyltransferase gene, perhaps reflecting host range differences.

Accession number(s). Nucleotide sequence accession numbers are provided in Table 1.

ACKNOWLEDGMENTS

This work was supported by Howard Hughes Medical Institute grant 54308198 to G.F.H.

We thank PHIRE Program students at the University of Pittsburgh and SEA-PHAGES Program students at Baylor University, Bucknell University, Cabrini University, Carnegie Mellon University, Montclair State University, The Ohio State University, University of North Texas, University of Wisconsin—River Falls, and Western Kentucky University (https://seaphages.org/media/GA_authors/GA7_Students_11-9-17.pdf) for their contributions to phage discovery and genomics. We thank the Howard Hughes Medical Institute for support for the SEA-PHAGES Program.

REFERENCES

- Hendrix RW. 2002. Bacteriophages: evolution of the majority. *Theor Popul Biol* 61:471–480. <https://doi.org/10.1006/tpbi.2002.1590>.
- Klyczek KK, Bonilla JA, Jacobs-Sera D, Adair TL, Afram P, Allen KG, Archambault ML, Aziz RM, Bagnasco FG, Ball SL, Barrett NA, Benjamin RC, Blasi CJ, Borst K, Braun MA, Broomell H, Brown CB, Brynell ZS, Bue AB, Burke SO, Casazza W, Cautela JA, Chen K, Chimalakonda NS, Chudoff D, Connor JA, Cross TS, Curtis KN, Dahlke JA, Deaton BM, Degroote SJ, DeNigris DM, DeRuff KC, Dolan M, Dunbar D, Egan MS, Evans DR, Fahnestock AK, Farooq A, Finn G, Fratus CR, Gaffney BL, Garland RA, Garrigan KE, Gibbon BC, Goedde MA, Guerrero Bustamante CA, Harrison M, Hartwell MC, Heckman EL. 2017. Tales of diversity: genomic and morphological characteristics of forty-six *Arthrobacter* phages. *PLoS One* 12:e0180517. <https://doi.org/10.1371/journal.pone.0180517>.
- Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, Jacobs WR, Hendrix RW, Lawrence JG, Hatfull GF, Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science, Phage Hunters Integrating Research and Education, Mycobacterial Genetics Course. 2015. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. *Elife* 4:e06416. <https://doi.org/10.7554/eLife.06416>.
- Russell DA, Hatfull GF. 2016. Complete genome sequence of *Arthrobacter* sp. ATCC 21022, a host for bacteriophage discovery. *Genome Announc* 4(2):e00168-16. <https://doi.org/10.1128/genomeA.00168-16>.
- Besemer J, Borodovsky M. 2005. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Res* 33:W451–W454. <https://doi.org/10.1093/nar/gki487>.
- Delcher AL, Harmon D, Kasif S, White O, Salzberg SL. 1999. Improved microbial gene identification with Glimmer. *Nucleic Acids Res* 27:4636–4641. <https://doi.org/10.1093/nar/27.23.4636>.
- 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](https://doi.org/10.1016/S0022-2836(05)80360-2).
- Söding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res* 33:W244–W248. <https://doi.org/10.1093/nar/gki408>.
- Laslett D, Canback B. 2004. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32:11–16. <https://doi.org/10.1093/nar/gkh152>.
- Schattner P, Brooks AN, Lowe TM. 2005. The tRNAscan-SE, snoscan and snoGPS Web servers for the detection of tRNAs and snoRNAs. *Nucleic Acids Res* 33:W686–W689. <https://doi.org/10.1093/nar/gki366>.
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33:784–786. <https://doi.org/10.1093/bioinformatics/btw711>.