

Genome Sequences of Alatato and Audell, two Phages that infect *Arthrobacter globiformis* and have Different Life Cycles

Sarah E Robinson¹, Erin Stanford¹, Shilen Abraham¹, Rebecca Aguilera¹, Nihal S Anantoji¹, Marco Barillas¹, Ryan Barstys¹, Emily Benson¹, Kayce R Boucher¹, Adrianna M Chacko¹, Niveditha Chandrakanth¹, Angel Chau¹, Angela Chen¹, John Coleman¹, Luong Tri Duc¹, Olivia Godfrey¹, Sasha N Hoang¹, Medhana Kethamreddy¹, Nisarg Kumar¹, Bao Ngan Le¹, Gregory T Lin¹, Kim B Nguyen¹, Madhura Pandit¹, Kelie D Shah¹, Mohini Sharma¹, Sophia Singh¹, Josh Smith¹, Lucy V Thomas¹, Alvin Varghese¹, Ian Wells¹, Richard S Pollenz^{1§}

¹Molecular Biosciences, University of South Florida, Tampa, Florida, United States

[§]To whom correspondence should be addressed: pollenz@usf.edu

Abstract

Based on gene content similarity, Alatato and Audell are grouped to phage clusters FB and FR, respectively. Alatato has a 39,780bp genome and 71 genes of which 34 have a predicted function. Alatato is predicted to be temperate based on the identification of an immunity cassette with genes that encode a tyrosine integrase, helix-turn-helix DNA binding proteins and an excise. Audell has a 42,187bp genome and 67 genes of which 32 have a predicted function with five enzymes involved in thymine synthesis. Putative lysis cassettes with genes encoding endolysins and transmembrane proteins were identifiable in both phages.

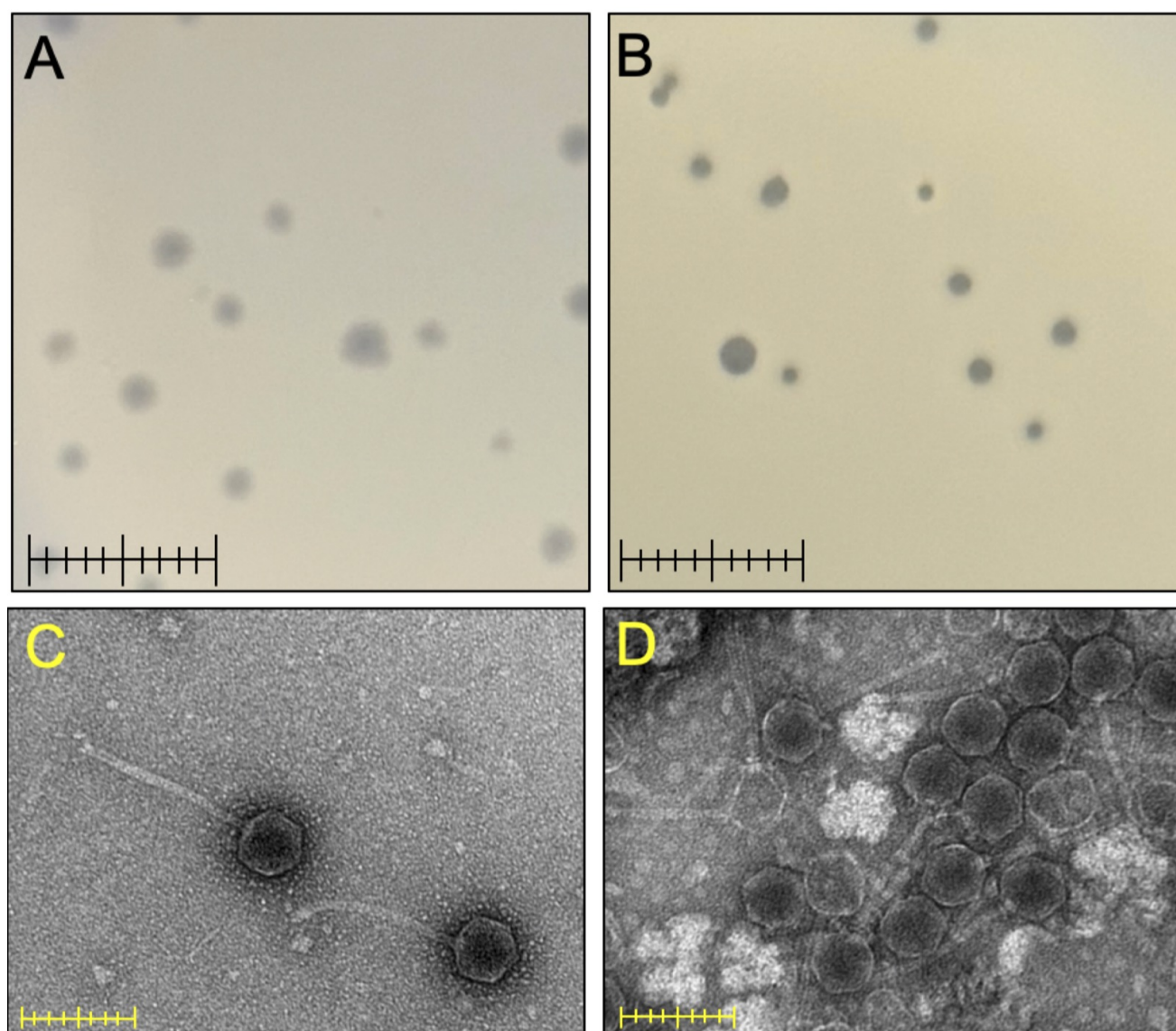


Figure 1. Plaques and particles of bacteriophages Alatato and Audell:

Plaque sizes of Alata (A) and Audell (B). Ruler = 1cm. Transmission electron micrographs of Alata (C) and Audell (D) stained using UryanLess (Electron Microscopy Sciences). Ruler = 200nm.

Description

There are an estimated 10^{31} phage particles in the biosphere (Rohwer et al., 2014) and to date the SEA- PHAGES program (Jordan et al., 2014) has isolated and annotated over 5,500 novel actinobacteriophages that are archived in the Actinobacteriophage Database at PhagesDB.org (Russell and Hatfull, 2017). Phages offer a vast repository of uncharacterized genes and the isolation and annotation of evolutionarily diverse phages can help advance the field of phage genomics and the application of phages to the fields of agriculture, food safety, and medicine (Shield et al., 2021; Wang and Zhao, 2022).

Alata and Audell were isolated from ~15g moist soil samples taken from different areas of Tampa, FL. GPS coordinates are shown in Table 1. Soil samples were mixed with equal parts peptone-yeast calcium media (PYCa), shaken at 250rpm for 2 hr. at 25°C and filter sterilized (0.2µm PES). 500ul of the sterile soil filtrates was mixed with 250ul of saturated *Arthrobacter globiformis* NRRL B-2979 and plaque assays performed by mixing the sample with 5ml of PYCa Top Agar and plating on PYCa agar plates. Plates were incubated at 30°C for 18-24 hrs. A single plaque that was separated a minimum of 2cm from others was picked with a sterile 1000ul pipette tip, resuspended in 100ul of Phage Buffer (PB; 10mM Tris pH 7.5; 1mM CaCl₂; 10mM MgSO₄; 68mM NaCl), vortexed and then serial diluted in PB. After a 10-minute incubation at room temperature, 10ul aliquots were used from each serial dilution to complete plaque assays as described above. A plaque was picked from the new set of plates, serial diluted and plated as above to assure a clonal population with uniform plaque morphology. High titer lysates ($> 5 \times 10^9$ PFU/ml) were prepared from these plates and genomic DNA was isolated using the Wizard DNA clean-up kit (A7280; Promega). Genomic DNA was used to create sequencing libraries with the NEB XLEP-P1 kit. Sequencing was performed by the Pittsburgh Bacteriophage Institute and the library run on an Illumina NextSeq 1000 instrument. Raw reads were trimmed with cutadapt 4.7 (using the option: `-nextseq-trim 30`) and filtered with skewer 0.2.2 (using the options: `-q 20 -Q 30 -n -l 50`) prior to assembly. Sequencing data for each phage is presented in Table 1. The resulting sequences were assembled using Consed (v29.0) with Unicycler (v5.0) and contigs checked for completeness, accuracy, and genome termini (Gordon et al., 1998; Russell, 2018). Default parameters were used for all software. The genomes of both Alata (5'-TGCCGCCCCGTA) and Audell (5'-TTTACGGCAT) have 3' sticky overhangs and were bioinformatically linearized such that base 1 is assigned in accord with other *Arthrobacter* phages (Russell, 2018). Both genomes were auto-annotated using DNA master (v5.23.6) (Pope and Jacobs-Sera, 2018) and the genes then manually validated for starts and functional calls. GeneMark (v2.5) (Besemer and Borodovsky, 2005) and Glimmer (v3.02) (Delcher et al., 2007) were utilized to assess start sites and coding potential. Starterator (v601; github.com/SEA-PHAGES/starterator) to summarize the starts across each family of phage genes. Evidence to support a gene product function was collected using HHpred (databases: PDB_mmCIF70_30_Mar; Pfam-A_v37; NCBI_Conserved_Domains(CD)_v3.19) (Söding, 2005; Marchler-Bauer, 2015), and NCBI BLAST (BLAST+2.13) (Altschul et al., 1990). Putative transmembrane domains (TMD) were identified using Deep TMHMM (v1.0.24) (Hallgren et al., 2022) and TOPCONS (v2.0) (Tsirigos et al., 2015). Information on each phage regarding isolation, characterization, and gene content is archived in Phamerator (Actino_draft database v578; Cresawn et al., 2011), and the Actinobacteriophage Database at PhagesDB.org (<https://phagesdb.org/phages/Alata/> and <https://phagesdb.org/phages/Audell/>).

Alata and Audell were assigned to clusters FB and FR, respectively based on gene content similarity of at least 35% to phages in the Actinobacteriophage database, PhagesDB (<https://phagesdb.org>) (Pope et al., 2017; Russell and Hatfull, 2017). Both phages are siphoviruses as determined by negative stain transmission electron microscopy and have icosahedral capsids and flexible tails (measurements presented in Figure 1). Alata produces cloudy plaques ranging from 1 mm to 2.5 mm (Figure 1). The establishment of lysogeny by phage Alata is predicted based on the identification of a putative immunity cassette inclusive of and flanked by genes encoding a tyrosine integrase (reverse gene 29) and an excise (forward gene 35). Within this flanked region, Alata reverse gene 33 and forward genes 32 and 34 all encode helix-turn-helix (HTH) DNA binding domain proteins that likely function as repressors based on high probability HHpred hits to the lambda CI repressor PDB 1LMB_3 (Beamer and Pabo, 1990). All of the other FB phages appear to have an immunity cassette in a similar region of the genome but there is some diversity in gene content and organization. While all contain an integrase grouped to the same pham as Alata, this gene is found in the forward direction in phages Kumotta and MargaretKali. All FB phages encode a protein grouped in the same pham as Alata HTH proteins gp31 and gp33, but only phage Shoya contains a forward gene encoding a protein ortholog to Alata gp32. Sarge and Shoya encode a protein grouped in the same pham as Alata gp34, while the other FB phages have a different gene in this region that also encodes an HTH protein but is grouped to a different pham based having <34% identity Alata gp34. Thus, all FB phages have genes that encode at least three distinct HTH proteins and since all of the genes are predicted to encode repressors, it is not possible to identify which might serve as the immunity repressor or cro protein without additional wet lab experiments. Alata also contains a putative lysis cassette with gene 21 encoding an endolysin flanked upstream by gene

19 encoding a protein with a single transmembrane domain (TMD) and downstream by gene 22 encoding a putative 2TMD holin.

Phage Audell shares 67%, 76% and 72% gene content to FR phages Neuville (draft), TMaxx (draft) and Annabelle, respectively. Audell produces clear plaques ranging from 1 mm to 2.5 mm (Figure 1) and is predicted to be lytic based on the absence of genes associated with lysogeny. Audell has a lysis cassette organization similar to Alata with gene 24 encoding an endolysin flanked upstream by gene 23 encoding a 1TMD protein and downstream by gene 25 encoding a 2TMD holin protein similar to Alata. Although the function of these proteins in lysis has not been evaluated in *Arthrobacter* phages, the lysis cassette organization is different from phages that infect the Gram positive *Gordonia rubripertincta* NRRL B-16540 strain, where the TMD encoded proteins are found downstream of the endolysin (Pollenz et al., 2022; McGarrah et al., 2023). Audell contains several genes that encode enzymes predicted to be involved in thymine synthesis including: dUTPase (gene 33), dihydrofolate reductase (gene 41), deoxycytidylate deaminase (gene 43), thymidylate kinase (gene 50) and ribonucleotide reductase (gene 57). These enzymes may be involved in thymine hypermodification that protect the phage genome from host defense systems (Lee et al., 2021).

Data availability

Alata whole Genome Shotgun project has been deposited in DDB/ENA/GenBank under the accession no. PV915846 and [SRX28150552](#). Audell whole Genome Shotgun project has been deposited in DDB/ENA/GenBank under the accession no. PV915817 and [SRX28150564](#). The versions described in this paper are the first version.

Table 1. Summary of relevant sequencing information, isolation parameters, genome characteristics and morphology measurements of bacteriophages Alata and Audell

Parameter	Bacteriophage	
phagesDB link	Alata	Audell
Location of sample collection, GPS Coordinates	Tampa, FL. 28.06 N, 82.4180 W	Tampa, FL. 27.9275 N, 82.1655 W
Single-end Read Size	100bp	100bp
Single-end Reads, #	2.6 x 10 ⁶	2.9 x 10 ⁶
Shotgun Coverage	6,536	6,874
Genome Size (bp)	39,780	42,187
G/C content (%)	63.8	58.6
Total Protein Encoding Genes	71	67
Genes with Functional Annotation	34	32
tRNA Genes	0	0
Cluster	FB	FR
Capsid Diameter	49.7 nm +/- 3.4 nm (n = 10)	50.5 nm +/- 2.8 nm (n = 12)
Tail Length	255 nm +/- 5.5 nm (n = 10)	230 nm +/- 4.3 nm (n = 12)

Acknowledgements: This work was made possible by HHMI Science Education Alliance and the University of Pittsburgh. D.A. Russell, D. Jacobs-Sera, and G.F. Hatfull are gratefully acknowledged for their help with genome sequencing and quality control and we also acknowledge support of the Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science program (Jordan et al., 2014).

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403-410. DOI: [10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Beamer LJ, Pabo CO. 1992. Refined 1.8 Å crystal structure of the λ repressor-operator complex. *Journal of Molecular Biology* 227: 177-196. DOI: [10.1016/0022-2836\(92\)90690-L](https://doi.org/10.1016/0022-2836(92)90690-L)
- Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. *Nucleic Acids Research* 33: W451-W454. DOI: [10.1093/nar/gki487](https://doi.org/10.1093/nar/gki487)
- Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. *BMC Bioinformatics* 12: 10.1186/1471-2105-12-395. DOI: [10.1186/1471-2105-12-395](https://doi.org/10.1186/1471-2105-12-395)
- Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* 23: 673-679. DOI: [10.1093/bioinformatics/btm009](https://doi.org/10.1093/bioinformatics/btm009)
- Gordon D, Abajian C, Green P. 1998. *Consed*: A Graphical Tool for Sequence Finishing. *Genome Research* 8: 195-202. DOI: [doi:10.1101/gr.8.3.195](https://doi.org/10.1101/gr.8.3.195)
- Hallgren J, Tsirigos KD, Pedersen MD, Almagro Armenteros JJ, Marcatili P, Nielsen H, Krogh A, Winther O. 2022. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. : 10.1101/2022.04.08.487609. DOI: [10.1101/2022.04.08.487609](https://doi.org/10.1101/2022.04.08.487609)
- Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, et al., Hatfull. 2014. A Broadly Implementable Research Course in Phage Discovery and Genomics for First-Year Undergraduate Students. *mBio* 5: 10.1128/mbio.01051-13. DOI: [10.1128/mbio.01051-13](https://doi.org/10.1128/mbio.01051-13)
- Lee YJ, Dai N, Müller SI, Guan C, Parker MJ, Fraser ME, et al., Weigele. 2021. Pathways of thymidine hypermodification. *Nucleic Acids Research* 50: 3001-3017. DOI: [10.1093/nar/gkab781](https://doi.org/10.1093/nar/gkab781)
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, et al., Bryant. 2014. CDD: NCBI's conserved domain database. *Nucleic Acids Research* 43: D222-D226. DOI: [10.1093/nar/gku1221](https://doi.org/10.1093/nar/gku1221)
- McGarrah CEE, Algarin-Martinez ED, Cavasini MED, Correa V, Danielson DF, Dean WR, et al., Pollenz. 2023. Isolation and Annotation of Azira, a CT Cluster Phage That Infects *Gordonia rubripertincta*. *Microbiology Resource Announcements* 12: 10.1128/mra.00347-23. DOI: [10.1128/mra.00347-23](https://doi.org/10.1128/mra.00347-23)
- Pollenz RS, Bland J, Pope WH. 2022. Bioinformatic characterization of endolysins and holin-like membrane proteins in the lysis cassette of phages that infect *Gordonia rubripertincta*. *PLOS ONE* 17: e0276603. DOI: [10.1371/journal.pone.0276603](https://doi.org/10.1371/journal.pone.0276603)
- Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al., Hatfull. 2017. Bacteriophages of *Gordonia* spp. Display a Spectrum of Diversity and Genetic Relationships. *mBio* 8: 10.1128/mbio.01069-17. DOI: [10.1128/mbio.01069-17](https://doi.org/10.1128/mbio.01069-17)
- Pope WH, Jacobs-Sera D. 2017. Annotation of Bacteriophage Genome Sequences Using DNA Master: An Overview. *Methods in Molecular Biology, Bacteriophages* : 217-229. DOI: [10.1007/978-1-4939-7343-9_16](https://doi.org/10.1007/978-1-4939-7343-9_16)
- Rohwer F, Youle M, Maughan H, Hisakawa N. 2014. *Life in Our Phage World. A centennial field guide to the Earth's most diverse inhabitants.* ISBN: 978-0-9904943-0-0. Wholon, San Diego, CA
- Russell DA. 2018. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. *Methods in Molecular Biology, Bacteriophages* : 109-125. DOI: [10.1007/978-1-4939-7343-9_9](https://doi.org/10.1007/978-1-4939-7343-9_9)
- Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. *Bioinformatics* 33: 784-786. DOI: [10.1093/bioinformatics/btw711](https://doi.org/10.1093/bioinformatics/btw711)
- Shield CG, Swift BMC, McHugh TD, Dedrick RM, Hatfull GF, Satta G. 2021. Application of Bacteriophages for Mycobacterial Infections, from Diagnosis to Treatment. *Microorganisms* 9: 2366. DOI: [10.3390/microorganisms9112366](https://doi.org/10.3390/microorganisms9112366)
- Soding J, Biegert A, Lupas AN. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33: W244-W248. DOI: [10.1093/nar/gki408](https://doi.org/10.1093/nar/gki408)
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. 2015. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. *Nucleic Acids Research* 43: W401-W407. DOI: [10.1093/nar/gkv485](https://doi.org/10.1093/nar/gkv485)
- Wang Z, Zhao X. 2022. The application and research progress of bacteriophages in food safety. *Journal of Applied Microbiology* 133: 2137-2147. DOI: [10.1111/jam.15555](https://doi.org/10.1111/jam.15555) PMID: 35353432

Yoon SH, Ha Sm, Lim J, Kwon S, Chun J. 2017. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie van Leeuwenhoek* 110: 1281-1286. DOI: [10.1007/s10482-017-0844-4](https://doi.org/10.1007/s10482-017-0844-4)

Funding: NA

Author Contributions: Sarah E Robinson: investigation, validation, visualization, writing - original draft, writing - review editing. Erin Stanford: investigation, validation, visualization, writing - original draft, writing - review editing. Shilen Abraham: investigation, validation, visualization, writing - original draft, writing - review editing. Rebecca Aguilera: investigation, validation, visualization, writing - original draft, writing - review editing. Nihal S Anantoji: investigation, validation, visualization, writing - original draft, writing - review editing. Marco Barillas: investigation, validation, visualization, writing - original draft, writing - review editing. Ryan Barstys: investigation, validation, visualization, writing - original draft, writing - review editing. Emily Benson: investigation, validation, visualization, writing - original draft, writing - review editing. Kayce R Boucher: investigation, validation, visualization, writing - original draft, writing - review editing. Adrianna M Chacko: investigation, validation, visualization, writing - original draft, writing - review editing. Niveditha Chandrakanth: investigation, validation, visualization, writing - original draft, writing - review editing. Angel Chau: investigation, validation, visualization, writing - original draft, writing - review editing. Angela Chen: investigation, validation, visualization, writing - original draft, writing - review editing. John Coleman: investigation, validation, visualization, writing - original draft, writing - review editing. Luong Tri Duc: investigation, validation, visualization, writing - original draft, writing - review editing. Olivia Godfrey: investigation, validation, visualization, writing - original draft, writing - review editing. Sasha N Hoang: investigation, validation, visualization, writing - original draft, writing - review editing. Medhana Kethamreddy: investigation, validation, visualization, writing - original draft, writing - review editing. Nisarg Kumar: investigation, validation, visualization, writing - original draft, writing - review editing. Bao Ngan Le: investigation, validation, visualization, writing - original draft, writing - review editing. Gregory T Lin: investigation, validation, visualization, writing - original draft, writing - review editing. Kim B Nguyen: investigation, validation, visualization, writing - original draft, writing - review editing. Madhura Pandit: investigation, validation, visualization, writing - original draft, writing - review editing. Kelie D Shah: investigation, validation, visualization, writing - original draft, writing - review editing. Mohini Sharma: investigation, validation, visualization, writing - original draft, writing - review editing. Sophia Singh: investigation, validation, visualization, writing - original draft, writing - review editing. Josh Smith: investigation, validation, visualization, writing - original draft, writing - review editing. Lucy V Thomas: investigation, validation, visualization, writing - original draft, writing - review editing. Alvin Varghese: investigation, validation, visualization, writing - original draft, writing - review editing. Ian Wells: investigation, validation, visualization, writing - original draft, writing - review editing. Richard S Pollenz: conceptualization, investigation, validation, visualization, methodology, writing - original draft, writing - review editing.

Reviewed By: Anonymous

History: Received August 5, 2025 **Revision Received** November 6, 2025 **Accepted** December 3, 2025 **Published Online** December 3, 2025 **Indexed** December 17, 2025

Copyright: © 2025 by the authors. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Citation: Robinson SE, Stanford E, Abraham S, Aguilera R, Anantoji NS, Barillas M, et al., Pollenz RS. 2025. Genome Sequences of Alata and Audell, two Phages that infect *Arthrobacter globiformis* and have Different Life Cycles. *microPublication Biology*. [10.17912/micropub.biology.001786](https://doi.org/10.17912/micropub.biology.001786)