| Bacteriophages | Announcement

# Desert diversity: genome sequence of *Gordonia rubripertincta* cluster DJ phage Mossy and cluster DV phage Erutan

October Barnes,[1] Christopher J. Workman,[1] Noah C. Patterson,[1] Riley Oesch,[1] Katie L. Johnson,[1] Kaarin Goncz,[1] Joel Sharbrough,[1] Linda C. DeVeaux[1]

**AUTHOR AFFILIATION** See affiliation list on p. 2.

**ABSTRACT** Lytic bacteriophages Mossy and Erutan were directly isolated from desert soil on *Gordonia rubripertincta* and characterized by their morphologies and genomes. Mossy, part of the DJ cluster of Actinobacteriophage, has a genome of 61,183 bp. The genome of Erutan, part of the DV cluster, is 66,957 bp.

Phages from desert climates are relatively understudied in comparison to marine counterparts (1). We describe the genomes of two novel phages isolated from northern Chihuahuan desert soil.

Mossy and Erutan were isolated from semi-moist soil from New Mexico Tech campus in Socorro, New Mexico (Table 1) on *Gordonia rubripertincta* using standard procedures (2). Soil was suspended in peptone-yeast extract-calcium medium (PYCa), the suspensions filtered (0.2 µm pore size), and plated in PYCa top agar with *G. rubripertincta*. After 3 days of incubation (30°C), a plaque was selected and purified through three rounds of plating. Mossy produced plaques with clear centers surrounded by turbid rings and defined outer borders (diameter ~1 mm; $n = 10$; StDv = 0.05). Erutan produced clear plaques (diameter ~1.5 mm; StDv = 0.04) (https://phagesdb.org/). Both phages exhibited siphovirus morphologies according to transmission electron micrographs (https://phagesdb.org/) analyzed in ImageJ (3). Lysates were placed on 200-mesh formvar-covered, carbon-coated copper grids (EMS) stained with 2% uranyl acetate and imaged in a Hitachi HT7800 TEM at 100 kV. Mossy virions possess a flexible tail measuring 254 nm ($n = 13$; StDv = 38.2) and an icosahedral capsid (width ~66 nm; $n = 13$; StDv = 4.2). Erutan's capsid was slightly smaller (width ~59 nm; $n = 4$; StDv = 4.1), with a longer tail (length ~404 nm; $n = 5$; StDv = 9.5).

DNA was isolated from high-titer phage lysates using the Wizard DNA Cleanup Kit (Promega, Madison, USA), prepared for sequencing using the NEBNext Ultra II FS Kit, and sequenced on an Illumina MiSeq (v3 Reagents). This yielded ~448K 150-bp single-end reads for Mossy and ~611K 150-bp single-end reads for Erutan. Genomes were assembled using Newbler v2.9, and assemblies were verified using Consedv29 (4) (Table 1). Ends were determined as previously described (4).

Genomes were auto-annotated using Glimmer v3.02b (5) and GeneMarkS 2.5 p (6), and manually verified using PECAAN v20221109 (https://discover.kbrinsgd.org/evidence/summary), BLASTp (7), Phamerator v509 (https://phamerator.org ), and Starterator v509 (http://phages.wustl.edu/starterator). Functions were assigned using BLASTp (against the actinobacteriophages and NCBI non-redundant databases) and HHPred (against the PDB mmCIF70, Pfam-A, and NCBI Conserved Domain databases) (8), and screened for tRNAs by Aragorn (9) and tRNAscanSE (10). Default parameters were used for all bioinformatic analyses (11). Based on gene content similarity of at least 35% to
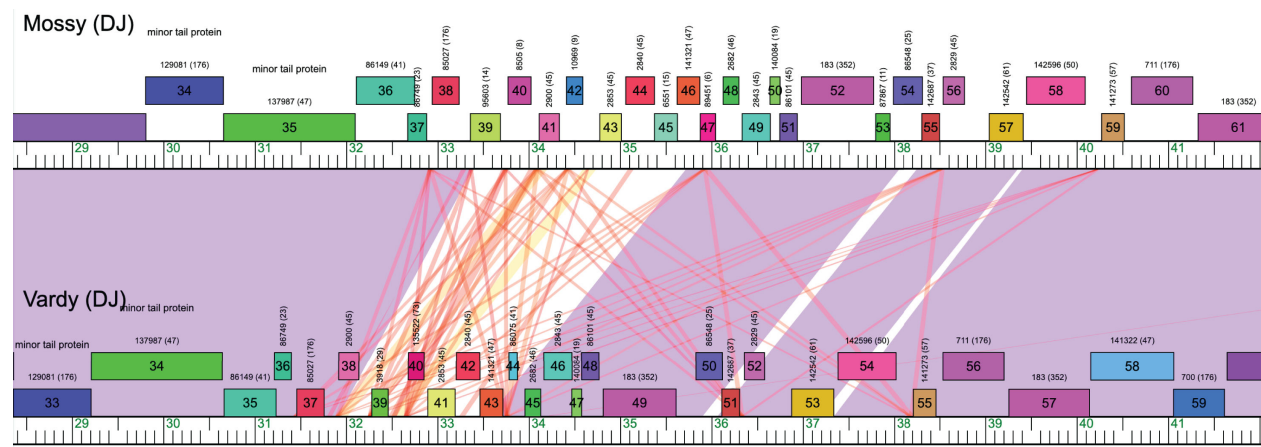
FIG 1 Repeated sequence motif found in cluster DJ phage genomes. Phamerator map of Mossy (top) and Vardy (bottom) phages, with genes indicated by numbered boxes. Purple connecting boxes indicate regions of high BLAST similarity between phage genomes, while white areas indicate sequences not shared across genomes. Red lines indicate short stretches of homologous sequence. In this case, the repeat motif (red lines) is highly abundant in both genomes but variable in its locations. The region with the highest abundance of the repeat motif also coincides with the region of Mossy's genome that is not shared with Vardy (or other cluster DJ phages).

TABLE 1 Phage characteristics and genome assembly statistics

| Phage | GPS coordinates | Length (bp) | Coverage | GC content (%) | End type | Cluster |
|---|---|---|---|---|---|---|
| Mossy | 34.068091 N, 106.909335 W | 61,183 | 1,069× | 51.9 | Non-redundant terminal repeats | DJ |
| Erutan | 34.067882 N, 106.909707 W | 66,957 | 1,359× | 58.4 | Circularly permuted | DV |

phages in PhagesDB (https://phagesdb.org/), Mossy and Erutan are assigned to phage clusters DJ and DV, respectively (12, 13).

Mossy's genome contains 90 predicted genes (25 assigned putative functions). As described for other cluster DJ phages, the middle region contains several repeated instances of a sequence motif that is predicted to be a translational start site (Fig. 1) (14). This repeat appears to be associated with rearrangements based on synteny comparisons with other cluster DJ phages (Fig. 1). Erutan's genome contains 95 predicted genes (28 with putative functions). One Erutan-specific gene (SEA_ERUTAN_5) aligns best with bacterial sequences (top hit: Caulobacteraceae, NBW15848.1, Max Score = 91.7, *e*-value = 7e−18). Three other cluster members (Gibbin, Tillicus, and Zany) also contain unique, putatively bacterially derived genes at this location, consistent with horizontal gene transfer, perhaps similar to that reported in *Rhodococcus* phages (15).

## AUTHOR AFFILIATION

[1]Department Biology, New Mexico Institute of Mining and Technology, Socorro, New Mexico, USA

## AUTHOR ORCIDs

Linda C. DeVeaux http://orcid.org/0000-0001-5040-7958

## AUTHOR CONTRIBUTIONS

October Barnes, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Christopher J. Workman, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Noah C. Patterson, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Riley Oesch, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Katie L. Johnson, Formal analysis, Investigation, Writing – original draft, Writing – review and editing | Kaarin Goncz, Formal analysis, Investigation, Methodology, Project administration, Supervision | Joel Sharbrough, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Supervision, Validation, Writing – review and editing | Linda C. DeVeaux, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Writing – review and editing

## DATA AVAILABILITY

Mossy available at GenBank Accession No. OR195046; Sequence Read Archive (SRA) No. SRX20165774. Erutan available at GenBank Accession No. OR475273 and SRA No. SRX20165762. Electron micrographs are available at http://phagesdb.org for Mossy and Erutan.

## REFERENCES

1. Pratama AA, van Elsas JD. 2018. The 'neglected' soil virome - potential role and impact. Trends Microbiol 26:649–662. https://doi.org/10.1016/j.tim.2017.12.004
2. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull GF. 2018. HHMI SEA-PHAGES phage discovery guide
3. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH image to imageJ: 25 years of image analysis. Nat Methods 9:671–675. https://doi.org/10.1038/nmeth.2089
4. Russell DA. 2018. Sequencing, assembling, and finishing complete bacteriophage genomes. Methods Mol Biol 1681:109–125. https://doi.org/10.1007/978-1-4939-7343-9_9
5. Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with glimmer. Bioinformatics 23:673–679. https://doi.org/10.1093/bioinformatics/btm009
6. Besemer J, Borodovsky M. 2005. GeneMark: web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res 33:W451–4. https://doi.org/10.1093/nar/gki487
7. 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
8. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. J Mol Biol 430:2237–2243. https://doi.org/10.1016/j.jmb.2017.12.007
9. 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
10. Chan PP, Lowe TM. 2019. tRNAscan-SE: searching for tRNA genes in genomic sequences. Methods Mol Biol 1962:1–14. https://doi.org/10.1007/978-1-4939-9173-0_1
11. Pope WH, Jacobs-Sera D. 2018. Annotation of bacteriophage genome sequences using DNA master: an overview. Methods Mol Biol 1681:217–229. https://doi.org/10.1007/978-1-4939-7343-9_16
12. Russell DA, Hatfull GF. 2017. PhagesDB: the actinobacteriophage database. Bioinformatics 33:784–786. https://doi.org/10.1093/bioinformatics/btw711
13. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, Russell DA, Warner MH, Hatfull GF, Science Education Alliance-Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES). 2017. Bacteriophages of Gordonia spp. display a spectrum of diversity and genetic relationships. mBio 8:e01069-17. https://doi.org/10.1128/mBio.01069-17
14. Henson M, Fratarcangeli M, Park S, Minot A, Gainey MD. 2022. Genome characteristics of vardy a cluster DJ actinobacteriophage isolated on *Gordonia rubripertincta* in Western North Carolina. Microbiol Resour Announc 11:e0070422. https://doi.org/10.1128/mra.00704-22
15. Summer EJ, Liu M, Gill JJ, Grant M, Chan-Cortes TN, Ferguson L, Janes C, Lange K, Bertoli M, Moore C, Orchard RC, Cohen ND, Young R. 2011. Genomic and functional analyses of *Rhodococcus equi* phages ReqiPepy6, ReqiPoco6, ReqiPine5, and ReqiDocb7. Appl Environ Microbiol 77:669–683. https://doi.org/10.1128/AEM.01952-10