Check for updates

Complete Genome Sequences of Mycobacteriophages HarryOW and Peeb

Fernando E. Nieto-Fernandez,[a,b] Christos Noutsos,[a] John Kleopoulos,[b] Olubusola Babalola,[c] Belle L. Connaught,[c] Balquees Shafique,[c] Shannon Farnum,[c] Hammad Nawaz,[c] Raymond Catapano,[c] Rita M. Reddy,[a] Jorge Morales,[d] Patricia Roccanova,[b] Shirley Barrera[b]

aDepartment of Biological Sciences, SUNY Old Westbury, Old Westbury, New York, USA
bSUNY OW Science and Technology Entry Program, Old Westbury, New York, USA
cWestbury High School, Old Westbury, New York, USA
dThe City College of New York, New York, New York, USA

ABSTRACT HarryOW and Peeb are *Mycobacterium smegmatis* mc² 155 *Siphoviridae* temperate phages with 52,935 and 41,876 base pairs in genome length, respectively. HarryOW belongs to the A1 subcluster and Peeb to the G1 subcluster. They were isolated and annotated by students from the SUNY Old Westbury Science and Technology Entry Program.

HarryOW and Peeb were isolated from soils located in Westbury, New York (40.760026 N, 73.589 W). Isolation was done by enrichment with *Mycobacterium smegmatis* mc² 155 following two cycles of purification and amplification in 7H9 top agar at 37°C (1). Phages were collected from a high-titer lysate by high-speed centrifugation; mounted on carbon-stabilized, Formvar-coated copper transmission electron microscopy (TEM) grids stained with uranyl acetate; and imaged with a JEOL JEM-2100 TEM (Fig. 1) (1). DNA was extracted using a Promega DNA Wizard kit with a modified protocol (1). Sequencing was done at the North Carolina State University Genomic Science Laboratory. Genomic libraries were prepared using an Illumina Truseq Nano DNA library kit. Pooled libraries were run on an Illumina MiSeq system yielding 670,230 and 661,512 single-end 150-base pair (bp) reads with average coverages of $1,783\times$ and $2,189\times$ for HarryOW and Peeb, respectively (Table 1). The raw reads were assembled using Newbler and checked for completeness and genomic termini with Consed at the University of Pittsburgh (2).
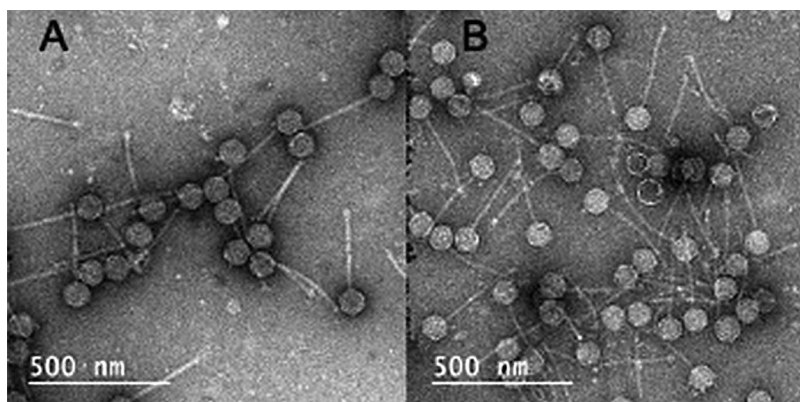
FIG 1 TEM samples mounted on Formvar-coated copper grids stained with uranyl acetate and imaged with a JEOL JEM-2100 at $\times$20,000. (A) HarryOW, icosahedral capsid ($\varnothing$ = 103 nm, tail = 283 nm). (B) Peeb, icosahedral capsid ($\varnothing$ = 93 nm, tail = 304 nm).

TABLE 1 Genome features and accession numbers for HarryOW and Peeb

Phage name	GenBank accession no.	SRA accession no.	Avg coverage (×)	Cluster	Genome length (bp)	GC content (%)	No. of genes
HarryOW	MT776812	SRX10050390	1,783	A1	52,935	63.8	94
Peeb	MW291028	SRX10050392	2,189	G1	41,876	66.6	62

Their genomes were annotated using DNA Master (http://cobamide2.bio.pitt.edu/computer.htm), and gene prediction tools GLIMMER v3.0 (3), GeneMark v2.5 (4), and Starterator v1.1 were used to determine gene start sites. tRNA and transfer-messenger RNA (tmRNA) predictions were made using ARAGORN v1.2.38 (5) and tRNAscan-SE v3.0 (6). Functional assignments were made using BLAST v2.9 (7), HHpred (8), and Phamerator (9). Manual annotation was done in parallel with PECAAN (10). Default settings were used for all programs. Phages in a cluster share sequence similarity over 50% of their genome length (11).

Peeb is a G1 cluster phage with 41,876 base pairs (bp), a GC content of 66.6%, 62 genes, and an 11-bp 3′ sticky overhang terminus. A total of 23 of the 62 genes are of known function. Peeb is 99% genetically identical to Schiebel (GenBank accession no. MH045569) with whom it shares 94.32% gene content and 58 gene phams. The integration cassette genes of Peeb are SEA_PEEB_32, a tyrosine integrase; SEA_PEEB_33, an immunity repressor; and SEA_PEEB_34, an excise. SEA_PEEB_32 and SEA_PEEB_33 are the only two genes in the Peeb genome coded on the reverse strand. SEA_PEEB_54 is a mycobacteriophage mobile element 1 (MPME1) found in other members of this gene pham and of this cluster (12).

HarryOW is an A1 cluster mycobacteriophage with 52,935 bp, a GC content of 63.8%, 94 genes, and an 11-bp 3′ sticky overhang terminus. A total of 35 of the 94 genes are of known function. Forty genes are on the forward strand, mostly on the left arm of the genome, except for genes SEA_HARRYOW_90, SEA_HARRYOW_91, and SEA_HARRYOW_93. HarryOW is 98% genetically identical to Rutherferd (GenBank accession no. NC_052461), sharing 80.02% gene content similarity and 74 phams. As in many of the A cluster phages, the immunity repressor SEA_HARRYOW_76 is not located adjacent to the serine integrase SEA_HARRYOW_37, and the DNA primase has two overlapping reading frames, namely, SEA_HARRYOW_54 and SEA_HARRYOW_55. One of the three minor tail protein genes, SEA_HARRYOW_6, is located in the left arm of the genome, upstream of the lysin A SEA_HARRYOW_10, lysin B SEA_HARRYOW_11, and the terminase SEA_HARRYOW_12. The other two minor tail protein genes, namely, SEA_HARRYOW_26 and SEA_HARRYOW_28, are located after the tape measure gene SEA_HARRYOW_25.

Data availability. The genome sequences for Peeb and HarryOW have been deposited in GenBank and the Sequence Read Archive (see Table 1).

REFERENCES

1. Poxleitner M, Pope W, Jacobs-Sera D, Sivanathan V, Hatfull G. 2018. Phage discovery guide. Howard Hughes Medical Institute, Chevy Chase, MD.
2. 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.
3. 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.
4. Lukashin A, Borodovsky M. 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res 26:1107–1115. https://doi.org/10.1093/nar/26.4.1107.
5. 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.
6. Lowe TM, Eddy SR. 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res 25:955–964. https://doi.org/10.1093/nar/25.5.955.
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. 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.

9. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. 2011. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. BMC Bioinformatics 12:395. https://doi.org/10.1186/1471-2105-12-395.

10. Rinehart CA, Gaffney BL, Smith JR, Wood JD. 2016. PECAAN: Phage Evidence Collection and Annotation Network user guide. Western Kentucky University Bioinformatics and Information Science Center, Bowling Green, KY. https://pecaan.kbrinsgd.org/.

11. Hatfull GF. 2014. Molecular genetics of mycobacteriophages. Microbiol Spectr 2:1–36.

12. Sampson T, Broussard GW, Marinelli LJ, Jacobs-Sera D, Ray M, Ko C-C, Russell D, Hendrix RW, Hatfull GF. 2009. Mycobacteriophages BPs, Angel and Halo: comparative genomics reveals a novel class of ultra-small mobile genetic elements. Microbiology (Reading) 155:2962–2977. https://doi.org/10.1099/mic.0.030486-0.