nature microbiology

Supplementary information

https://doi.org/10.1038/s41564-026-02486-1

Defining the essential genome of diverse phages with phage Tn-seq

In the format provided by the authors and unedited

Supplementary Information

Defining the essential genome of diverse phages with phage Tn-seq

Natalie Kyte[1-3#], Manuela Fuchs[1#], Leah M. Smith[1-3*] and Peter C. Fineran[1-4*]

[1]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand

[2]Maurice Wilkins Centre for Molecular Biodiscovery, University of Otago, Dunedin, New Zealand

[3]Genetics Otago, University of Otago, Dunedin, New Zealand

[4]Bioprotection Aotearoa, University of Otago, Dunedin, New Zealand

#These authors contributed equally

*For correspondence: leah.smith@otago.ac.nz and peter.fineran@otago.ac.nz

Supplementary Tables

Supplementary Table 1. Transposon insertion sequencing statistics for JS26:

Total reads are the number of reads following de-multiplexing. Reads with Tn*5* tag and % with Tn*5* tag are those containing the 12 nt transposon 'tag' generated during sequencing (2 mismatches allowed). Reads mapped and % mapped are those mapping to the *Prodigiosinella* phage JS26. Unique insertion sites (UIS) equals the total number of unique insertions. Length/UIS is the average number of unique insertions per nt for the JS26 genome (63.97 KB).

File	Total Tn Reads in sample	Reads Matched to Tn	% Matched to Tn	Reads Mapped to Phage Genome	% Mapped to Phage Genome	Unique Insertion Sites (UIS)	Genome length/UIS
Pre-enrichment combined	246,597	216,906	87.96	185,987	85.75	4465	14.33
Post-enrichment combined	1,241,819	1,116,482	89.9	982,648	88.01	3,765	17.0
Pre_enriched_ JS26_1	46,850	39,778	84.91	32,633	82.04	2252	28.41
Pre_enriched_ JS26_2	61,091	52,963	86.70	43,465	82.07	2041	31.34
Pre_enriched_ JS26_3	138,656	124,165	89.55	109,889	88.50	1217	52.56
Post_enrichment_ JS26_1	404,292	377,475	93.4	332,597	88.11	2,538	25.2
Post_enrichment_ JS26_2	466,885	407,338	87.2	334,874	82.21	2,378	26.9
Post_enrichment_ JS26_3	370,642	331,669	89.5	315,177	95.03	1,424	44.9

Supplementary Table 2. Transposon insertion sequencing statistics for PCH45:

Total reads are the number of reads following de-multiplexing. Reads with Tn5 tag and % with Tn5 tag are those containing the 12 nt transposon 'tag' generated during sequencing (2 mismatches allowed). Reads mapped and % mapped are those mapping to the *Prodigiosinella* phage PCH45. Unique insertion sites (UIS) equals the total number of unique insertions. Length/UIS is the average number of unique insertions per nt for the PCH45 genome (212.81 KB).

File	Total Tn Reads in sample	Reads Matched to Tn	% Matched to Tn	Reads Mapped to Phage Genome	% Mapped to Phage Genome	Unique Insertion Sites (UIS)	Genome length/UIS
Pre-enrichment UvsX negative combined	145,981	123,388	84.52	1,310	1.06	266	800.03
Pre-enrichment UvsX positive combined	200,713	163,960	81.69	6,507	3.97	451	471.86
Post-enrichment UvsX negative combined	1,205,388	1,112,938	92.33	437,838	39.34	1,445	147.27
Post-enrichment UvsX positive combined	13,043,913	11,749,195	90.07	9,426,289	80.23	9,851	21.60
Pre_enrichment_UvsX_Neg_PCH45_1	25,948	15,065	58.06	139	0.92	116	1,834.54
Pre_enrichment_UvsX_Neg_PCH45_2	43,167	39,088	90.55	318	0.81	129	1,649.67
Pre_enrichment_UvsX_Neg_PCH45_3	76,866	69,235	90.07	853	1.23	87	2,446.06
Pre_enrichment_UvsX_Pos_PCH45_1	141,047	111,019	78.71	3,386	3.05	283	751.97
Pre_enrichment_UvsX_Pos_PCH45_2	31,382	26,929	85.81	2,516	9.34	155	1,372.95
Pre_enrichment_UvsX_Pos_PCH45_3	28,284	26,012	91.97	605	2.33	115	1,850.50
Post_enrichment_UvsX_Neg_PCH45_1	794,155	736,509	92.74	209,162	28.40	946	224.95
Post_enrichment_UvsX_Neg_PCH45_2	139,804	129,716	92.78	56,972	43.92	800	266.01
Post_enrichment_UvsX_Neg_PCH45_3	271,429	246,713	90.89	171,704	69.60	309	688.70
Post_enrichment_UvsX_Pos_PCH45_1	9,803,228	9,027,161	92.08	7,339,452	81.30	4,393	48.44
Post_enrichment_UvsX_Pos_PCH45_2	1,261,438	1,071,535	84.95	847,265	79.07	4,783	44.49
Post_enrichment_UvsX_Pos_PCH45_3	1,979,247	1,650,499	83.39	1,239,572	75.10	6,995	30.42

Supplementary Table 3. Bacterial strains used in this study

Strain name	Description	Reference
E. coli ST18	*Escherichia coli* ST18; auxotrophic donor for conjugation	Thoma and Schobert[1] and Jackson *et al.* [2]
E. coli DH5α	*Escherichia coli* DH5α; cloning strain	Gibco/BRL
E. coli DH10β	*Escherichia coli* DH10β	Durfee *et al.* [3]
E. coli ΔRM	*Escherichia coli* K-12 MG1655 ΔRM	Maffei *et al.* [4]
P. confusarubida LacA	*Prodigiosinella confusarubida* (formerly *Serratia* sp. ATCC 39006) Lac- EMS mutant, denoted WT in this paper	Thomson *et al*. [5] and Hampton *et al*. [6]

Supplementary Table 4: Bacteriophages used in this study

Phage name	Description	Reference
JS26	*Prodigiosinella confusarubida* (ATCC 39006) siphovirus JS26	Malone *et al.*[7]
PCH45	*Prodigiosinella confusarubida* (ATCC 39006) nucleus-forming jumbo phage PCH45	Malone *et al.*[8]
LC53	*Prodigiosinella confusarubida* (ATCC 39006) phage LC53 containing 5-ara-hC modified DNA	Mahler *et al.*[9]
Bas46	*Escherichia coli* BASEL phage Bas46 containing 5-ara-ara-hC modified DNA	Maffei *et al.*[4]

Supplementary Notes

Supplementary Note 1. Predicted functions of JS26 genes mentioned in the manuscript

Gp46 is a predicted Rz-like spanin involved in cell lysis for phage release[10], Gp57 and Gp58 are predicted tail assembly chaperones and Gp62 and Gp63 are predicted pre-tape measure and tail length tape measure proteins, respectively. Gp48 and Gp68 have no predicted function; however, Gp68 is conserved across several phages, suggesting it may be essential in other genera. Another cluster of genes (gp74 to gp84) encoding proteins that were absent in mature virions was also classified as essential, which correlates with their predicted role in DNA replication.

The gp28-gp33 region is predicted to encode proteins involved in DNA transcriptional regulation, DNA modification and potentially DNA repair. Gp30 is a predicted winged helix-turn-helix (wHTH) transcriptional regulator, Gp32 is related to Dcm methyltransferases and Gp33 is an exonuclease recombination-associated protein. Gp28 and Gp29 have no predicted function.

Supplementary Note 2. Proteomic support for PCH45 virion protein essentiality

Essential virion proteins identified by phage Tn-seq and detected by proteomics included proteins involved in formation of the head (i.e., major capsid protein (Gp33), internal head vertex proteins (Gp23 and Gp24), head maturation protease (Gp81) and virion-associated head protein (Gp162)) and tail (i.e., tail tube (Gp154), tail sheath (Gp155), tail fibres (Gp56 and Gp60) and tail tip (Gp19)). Similarly, PCH45 Gp111 is an essential tail fibre protein and a homologue of ΦPA3 Gp164 (a *Chimalliviridae*), which localises to the phage “bouquet” of assembled particles during phage maturation in the cell[11]. In addition, there was evidence of injected essential proteins, such as virion-associated RNA polymerase (vRNAP) subunits (Gp3, Gp87 and Gp88).

Supplementary Note 3. Conserved and lineage-specific essential PCH45 genes

Conserved essential genes

Although 15 of the *chimallin*-encoding core genes were detected by proteomics, other essential core genes were not. As expected their role was during phage replication within the bacterial host and include the proteinaceous shell of the phage nucleus ChmA (gp202; cg4), RNA binding protein ChmC (gp209; cg10), nuclear shell import proteins PicA/Imp1 (gp211; cg12) and Imp3 (gp204; cg6) and the large terminase[12-15]. Although not assigned within the core genome, the genes encoding the small terminase (gp77) and portal protein (gp206) were also essential. These results are consistent with conserved jumbo phage genes functioning in critical steps in the viral lifecycle and reinforces the efficacy of Tn-seq in essential gene identification.

Conserved non-essential genes

Comparison with core gene sets also allowed us to identify non-essential core genes. Within the *chimallin*-encoding core genes, one non-essential (gp76) gene had no known function, whereas the non-essential gp198 gene encodes an SH3 family protein. Within the entire set of core genes across all jumbo phages, 12 were non-essential in PCH45. In some cases, multiple paralogues were annotated, with one

being non-essential. For example, gp56, gp57 and gp58 encode tail fibre proteins that are similar to cg42, with only gp58 being non-essential. Likewise, genes encoding head maturation proteins gp81 and gp220 are related to cg60, with gp220 being non-essential. A thymidylate kinase (gp50; cg32) involved in DNA biosynthesis[16,17] was non-essential, potentially due to functional redundancy with a host-encoded thymidylate kinase. Although UvsX localises to the phage nucleus, it was non-essential, which is consistent with the dispensable nature of a related recombinase in T4[18]. Although not part of the core genome, the tubulin homologue PhuZ (gp187) was non-essential in PCH45, consistent with its dispensable role in phiKZ[19].

Essential non-core genes

We also identified 30 essential PCH45 proteins that are not members of the core jumbo phage genome. Gp9 has no predicted function but is a homologue of ΦPA3 Gp078, which localises to the phage nucleus[11]. The essential gp107-gp111 operon was detected by proteomics and includes the predicted tail fibre (gp111) and tail assembly chaperones (gp107 and gp110). In addition, cell lysis cassette genes were essential, including gp152, a putative lytic transglycosylase, and gp54, which is predicted to encode an Rz lysis/spanin[20,21].

References

1 Thoma, S. & Schobert, M. An improved *Escherichia coli* donor strain for diparental mating. *FEMS Microbiol Lett* **294**, 127-132 (2009). https://doi.org/10.1111/j.1574-6968.2009.01556.x

2 Jackson, S. A., Fellows, B. J. & Fineran, P. C. Complete Genome Sequences of the *Escherichia coli* Donor Strains ST18 and MFDpir. *Microbiol Resour Announc* **9** (2020). https://doi.org/10.1128/MRA.01014-20

3 Durfee, T. *et al.* The complete genome sequence of *Escherichia coli* DH10B: insights into the biology of a laboratory workhorse. *J Bacteriol* **190**, 2597-2606 (2008). https://doi.org/10.1128/JB.01695-07

4 Maffei, E. *et al.* Systematic exploration of *Escherichia coli* phage-host interactions with the BASEL phage collection. *PLoS Biol* **19**, e3001424 (2021). https://doi.org/10.1371/journal.pbio.3001424

5 Thomson, N. R., Crow, M. A., McGowan, S. J., Cox, A. & Salmond, G. P. Biosynthesis of carbapenem antibiotic and prodigiosin pigment in *Serratia* is under quorum sensing control. *Mol Microbiol* **36**, 539-556 (2000). https://doi.org/10.1046/j.1365-2958.2000.01872.x

6 Hampton, H. G. *et al.* Functional genomics reveals the toxin-antitoxin repertoire and AbiE activity in *Serratia*. *Microb Genom* **6** (2020). https://doi.org/10.1099/mgen.0.000458

7 Malone, L. M., Hampton, H. G., Morgan, X. C. & Fineran, P. C. Type I CRISPR-Cas provides robust immunity but incomplete attenuation of phage-induced cellular stress. *Nucleic Acids Res* **50**, 160-174 (2022). https://doi.org/10.1093/nar/gkab1210

8 Malone, L. M. *et al.* A jumbo phage that forms a nucleus-like structure evades CRISPR-Cas DNA targeting but is vulnerable to type III RNA-based immunity. *Nat Microbiol* **5**, 48-55 (2020). https://doi.org/10.1038/s41564-019-0612-5

9 Mahler, M. *et al.* Phage arabinosyl-hydroxy-cytosine DNA modifications result in distinct evasion and sensitivity responses to phage defense systems. *Cell Host & Microbe* **33**, 1173-1190.e1179 (2025). https://doi.org/10.1016/j.chom.2025.06.005

10 Berry, J., Rajaure, M., Pang, T. & Young, R. The Spanin Complex Is Essential for Lambda Lysis. *Journal of Bacteriology* **194**, 5667-5674 (2012). https://doi.org/10.1128/jb.01245-12

11 Enustun, E. *et al.* Identification of the bacteriophage nucleus protein interaction network. *Nature Structural & Molecular Biology* **30**, 1653-1662 (2023). https://doi.org/10.1038/s41594-023-01094-5

12 Chaikeeratisak, V. *et al.* Assembly of a nucleus-like structure during viral replication in bacteria. *Science* **355**, 194-197 (2017).

13 Kokontis, C., Klein, T. A., Silas, S. & Bondy-Denomy, J. Multi-interface licensing of protein import into a phage nucleus. *Nature* **639**, 456-462 (2025). https://doi.org/10.1038/s41586-024-08547-x

14 Yakunina, M. *et al.* A non-canonical multisubunit RNA polymerase encoded by a giant bacteriophage. *Nucleic Acids Res* **43**, 10411-10420 (2015). https://doi.org/10.1093/nar/gkv1095

15 Morgan, C. J. *et al.* An essential and highly selective protein import pathway encoded by nucleus-forming phage. *bioRxiv* (2024). https://doi.org/10.1101/2024.03.21.585822

16 Tran Ngoc, Q., Tabor, S. & Richardson Charles, C. Genetic Requirements for Sensitivity of Bacteriophage T7 to Dideoxythymidine. *Journal of Bacteriology* **196**, 2842-2850 (2014). https://doi.org/10.1128/jb.01718-14

17 Chace, K. V. & Hall, D. H. Isolation of mutants of bacteriophage T4 unable to induce thymidine kinase activity. *J Virol* **12**, 343-348 (1973). https://doi.org/10.1128/jvi.12.2.343-348.1973

18 Hinton, D. M. & Nossal, N. G. Cloning of the bacteriophage T4 uvsX gene and purification and characterization of the T4 uvsX recombination protein. *J Biol Chem* **261**, 5663-5673 (1986).

19 Guan, J. *et al.* Bacteriophage genome engineering with CRISPR–Cas13a. *Nature Microbiology* **7**, 1956-1966 (2022). https://doi.org/10.1038/s41564-022-01243-4

20 Fokine, A., Miroshnikov, K. A., Shneider, M. M., Mesyanzhinov, V. V. & Rossmann, M. G. Structure of the bacteriophage phi KZ lytic transglycosylase gp144. *J Biol Chem* **283**, 7242-7250 (2008). https://doi.org/10.1074/jbc.M709398200

21 Berry, J., Summer, E. J., Struck, D. K. & Young, R. The final step in the phage infection cycle: the Rz and Rz1 lysis proteins link the inner and outer membranes. *Mol Microbiol* **70**, 341-351 (2008). https://doi.org/10.1111/j.1365-2958.2008.06408.x