**REPORT TEMPLATE:  
Collaborative Annotation Teams (CATs): CLUSTER REPORTS**

**PURPOSE**:   
To capture our current understanding of the characteristics of a given cluster, which can be used to support the process of genome annotations, the preparation of genome announcements, and inform and contribute towards future studies examining phages of this cluster.

**INSTRUCTIONS:**

* Use the sample/template Cluster Report below to develop a Cluster Report for your cluster.  
  *Note: Text in green is information specific to the cluster in the report. To use this template, replace the text in green with information relevant to the cluster you are examining. The text in black provides framing for the types of information that should be provided and can be edited as necessary.*
* To facilitate your ability to gather the relevant information for your cluster, the Cresawn lab has developed two Observable Notebooks that use data from Phamerator and phagesDB.
  + The Observable Notebook, *Subclusters*, provides a data overview of the cluster: <https://observablehq.com/d/5e5bc78c9b3ae2ed>.
  + The Observable Notebook, *Pham Matrices*, allows for an in-depth look at the gene organization for phages within a cluster: <https://observablehq.com/@cresawn-labs/pham-matrix>.
* Work collaboratively to develop a single report, guided by the SMART members in your group.
* At the end of the session,
  + Email the completed report to [sea@hhmi.org](mailto:sea@hhmi.org)

Over the coming months, the Cluster Reports will be reviewed and then published on QUBES to make advances in our understanding of phages available to the community, recognizing you, the authors, with a citable publication.

**A 2025 Report on the Characteristics of Phages Assigned to Cluster N.**

**Authors**: Denise Monti1, Viknesh Sivanathan1

**Facilitator**: (typically, the member of the SMART Team assigned to this cluster, please include the facilitator’s affiliation below)

**Affiliations**: 1Howard Hughes Medical Institute, Chevy Chase, MD  
  
\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Isolation and Morphological Characteristics**

As of June 2025, 46 phages have been assigned to Actinobacteriophage cluster N. Of those, 44 are annotated and two are in draft form. Cluster N consists of phages that were isolated from environmental samples collected in 8 different states, using *Mycobacterium smegmatis* mc2155 as the host. A sampling of transmission electron micrographs reveals these phages to have siphoviral morphologies, while photographs of plaque assays reveal small- to medium-sized clear plaques. Examples of phages displaying these morphologies include phages Butters and Pipsqueaks.

**Genometrics**  
The genomes of phages in this cluster have an average length of 42,908 bp (range: 40,580 bp – 44,872 bp) with an average GC content of 66.2% (range: 65.8% – 66.8%). The genomes have 13-base sticky overhang ends (CCCGCCGCAATGG).

**Gene Content**

1. A survey of all manually annotated genomes in this cluster (i.e., excluding draft genomes) reveals the number of genes called per genome to range between 63 – 76 genes, which represent a total of 235 gene phamilies (or phams).
2. Of the 235 phams represented in the cluster, 23 are fully conserved across the cluster. These phams are assigned putative functions that are involved in structure and assembly (terminase, portal protein, capsid maturation protein, major and minor capsid proteins, tape measure, tail assembly chaperones, and minor tail proteins).

1. Another 212 phams have varying levels of conservation. Notable here are 73 phams that are each only called once in this cluster, while another 24 phams are each only called twice in this cluster. Two of these latter phams are assigned the functions of RelB-like antitoxin and RelE-like toxin in phages Fulbright and Xeno.
2. Only two Cluster N phages identified tRNA genes. Phages Scitech and Tortoise12 identified a 78nt tRNA-Other (nnn).
3. Cluster N genomes are organized with genes involved in structure and assembly in the first third of the genome. Genes involved in lysis are also located in the first third of the genome. Genes associated with lysogeny are in the center of the genome, and genes associated with recombination/replication are at the beginning of the last third of the genome.
4. The majority of genes are transcribed in the forward direction, with the exception of several genes, including the integrase and repressor, that are in the middle of the genome and are transcribed in the reverse direction.

**Genes Associated with Lysis**

*Use the dropdown menu in the Pham Matrices Observable Notebook to explore genes related to lysis. In Observable, phages can be reorganized based on synteny by sliding the name of the phage in the list of selected phages. Once the phages have been reorganized, identify the borders of a putative lysis cassette by identifying the lysin/endolysin and holin genes, as well as transmembrane proteins in the vicinity. Be sure to hover over transmembrane proteins in the notebook to ensure they have not been assigned a function (e.g., tape measure protein).*

**If a putative lysis cassette is identified, consider the following text:**

Proteins required for phage lysis in Gram-positive bacteria include lysin (e.g., lysin A, lysin B, endolysin) and holin. Other transmembrane proteins may also be involved in the process of phage lysis (Pollenz, 2022). The Observable Notebook, *Pham Matrices*, was used to explore the arrangement of genes neighboring those known to be involved in phage lysis (Cobb et al., 2025). Among non-draft Cluster N phages, all phages have an annotated Lysin A gene encoded in 4 different phams. All phages also have a highly conserved, annotated holin gene encoded in 4 different phams. Both the lysin and holin are flanked by two highly conserved genes that encode transmembrane proteins.

We defined the putative lysis cassette as the region of the genome that contains the lysin and holin genes flanked by two conserved genes encoding putative transmembrane proteins based on shared architecture (Appendix A). Grouping the phages based on shared phams among the lysin genes reveals 11 different arrangements of genes that may be important for phage lysis (Figure 1). Four phages identified a Lysin B gene in addition to Lysin A. Among the 7 phages with a reverse gene neighboring the gene for a putative transmembrane protein downstream of holin, the reverse genes are in the same phamily. DeepTMHMM identified 1 transmembrane domain in 6 of the 7 phages with the reverse gene and this gene was annotated as a hypothetical protein in those phages. In phage PhancyPhin, the reverse gene does not have an identified membrane-spanning sequence and the gene was annotated as a minor tail protein.

**A diagram of a structure

AI-generated content may be incorrect.**

**Genes Associated with Lysogeny**

*Use the dropdown menu in the Pham Matrices Observable Notebook to explore genes related to lysogeny. In Observable, phages can be reorganized based on synteny by sliding the name of the phage in the list of selected phages. Once the phages have been reorganized, describe the gene arrangements for genes associated with lysogeny.*

The Observable Notebook, *Pham Matrices*, was used to explore the arrangement of genes neighboring those known to be involved in phage lysogeny (Cobb et al., 2025).

**If no genes associated with lysogeny are identified, the following text should be considered:**

Careful evaluation of the genome did not reveal any genes associated with lysogeny, including those with assigned functions such as integrase, repressor, excise, Cro, or partitioning protein. No lysogen reports are published in QUBES for Cluster M phages (SEA Faculty Group QUBES accessed June 7, 2025). OR Attempts to raise lysogens for phage XXX were unsuccessful (Cite the specific QUBES Lysogen Report).

**If genes associated with lysogeny are identified, the following text should be considered:**

1. All phages in the cluster identified a tyrosine integrase and a repressor with one exception, phage Rebel, which does not have annotated integrase or repressor genes (Appendix B). This suggests the majority of phages in Cluster N are likely temperate.
2. No lysogen reports are published in QUBES for Cluster M phages (SEA Faculty Group QUBES accessed June 7, 2025). OR Attempts to raise lysogens for phages XXX were successful (CITE QUBES Lysogen Report).
3. Analysis of the phages in Cluster N reveals five different arrangements of genes associated with lysogeny (Figure 2). The majority of phages in the cluster (86%) have a gene encoding a tyrosine integrase, followed by genes encoding an immunity repressor and an excise protein. In four genomes, a hypothetical protein or DNA-binding protein function was assigned to the gene downstream of the genes for the integrase and the repressor. Notably, the hypothetical protein and DNA-binding protein genes are in the same phamily as the highly conserved excise gene. One phage (Rebel) did not have annotated genes for the integrase or repressor. In phage Rebel, a single gene with an assigned DNA-binding function was annotated in the location of excise. This gene is in the same phamily as the excise identified in 38 of 44 Cluster N phages.

**A diagram of a dna structure

AI-generated content may be incorrect.**

**References:** *Include prior publications that describe the phages of this cluster, including genome announcements and cluster manuscripts. We recommend searching in PubMed using the cluster name in quotes and the term phage to filter the results (e.g., “Cluster N” and phage) and reviewing the papers associated with the cluster in phagesdb.org (Phages-Cluster List – Select the Cluster/subcluster of Interest – Scroll to see Listed Publications). If no publications are available for a listed category, consider including the text, “No prior publications available.”*

**Cluster Report**

No prior publications available.

**Genome Announcements**

1. Caratenuto, R. A., 3rd, Ciabattoni, G. O., DesGranges, N. J., Drost, C. L., Gao, L., Gipson, B., Kahler, N. C., Kirven, N. A., Melehani, J. C., Patel, K., Rokes, A. B., Seth, R. A., West, M. C., Alhout, A. A., Akoto, F. F., Capogna, N., Cudkevich, N., Graham, L. H., Grapel, M. S., Haleem, M. M., … Ware, V. C. (2019). Genome Sequences of Six Cluster N Mycobacteriophages, Kevin1, Nenae, Parmesanjohn, ShrimpFriedEgg, Smurph, and SpongeBob, Isolated on Mycobacterium smegmatis mc2155. Microbiology resource announcements, 8(22), e00399-19. https://doi.org/10.1128/MRA.00399-19
2. Kotturi, H., Sahi, U., Kedy, C., & Ali, A. K. (2021). Complete Genome Sequence of Mycobacteriophage Fulbright. Microbiology resource announcements, 10(11), e00123-21. <https://doi.org/10.1128/MRA.00123-21>
3. Mageeney, C. M., Mohammed, H. T., Dies, M., Anbari, S., Cudkevich, N., Chen, Y., Buceta, J., & Ware, V. C. (2020). *Mycobacterium* Phage Butters-Encoded Proteins Contribute to Host Defense against Viral Attack. *mSystems*, *5*(5), e00534-20. https://doi.org/10.1128/mSystems.00534-20
4. Ramirez Rendon, A. D., Diaz-Sanchez, A., Pool, R. J., Ahmed, F., Mendez, A., Medellin, R. C., Mendoza, F. A., Anderson, R., Sanchez, M. A., Ramos, J. R., Sadana, R., & Saha, S. (2021). Genome Sequences of Mycobacterium smegmatis Phages Purgamenstris and PhancyPhin. Microbiology resource announcements, 10(9), e01288-20. <https://doi.org/10.1128/MRA.01288-20>

**G3 Publications**

1. Mohammed, H. T., Mageeney, C., Korenberg, J., Graham, L., & Ware, V. C. (2023). Characterization of novel recombinant mycobacteriophages derived from homologous recombination between two temperate phages. *G3 (Bethesda, Md.)*, *13*(12), jkad210. <https://doi.org/10.1093/g3journal/jkad210>

**Cluster Papers**

1. Dedrick, R. M., Jacobs-Sera, D., Bustamante, C. A., Garlena, R. A., Mavrich, T. N., Pope, W. H., Reyes, J. C., Russell, D. A., Adair, T., Alvey, R., Bonilla, J. A., Bricker, J. S., Brown, B. R., Byrnes, D., Cresawn, S. G., Davis, W. B., Dickson, L. A., Edgington, N. P., Findley, A. M., Golebiewska, U., … Hatfull, G. F. (2017). Prophage-mediated defence against viral attack and viral counter-defence. *Nature microbiology*, *2*, 16251. https://doi.org/10.1038/nmicrobiol.2016.251

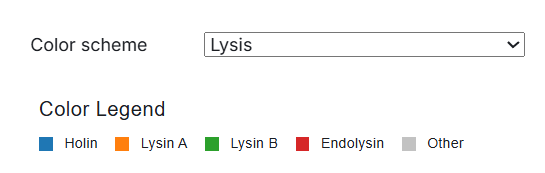
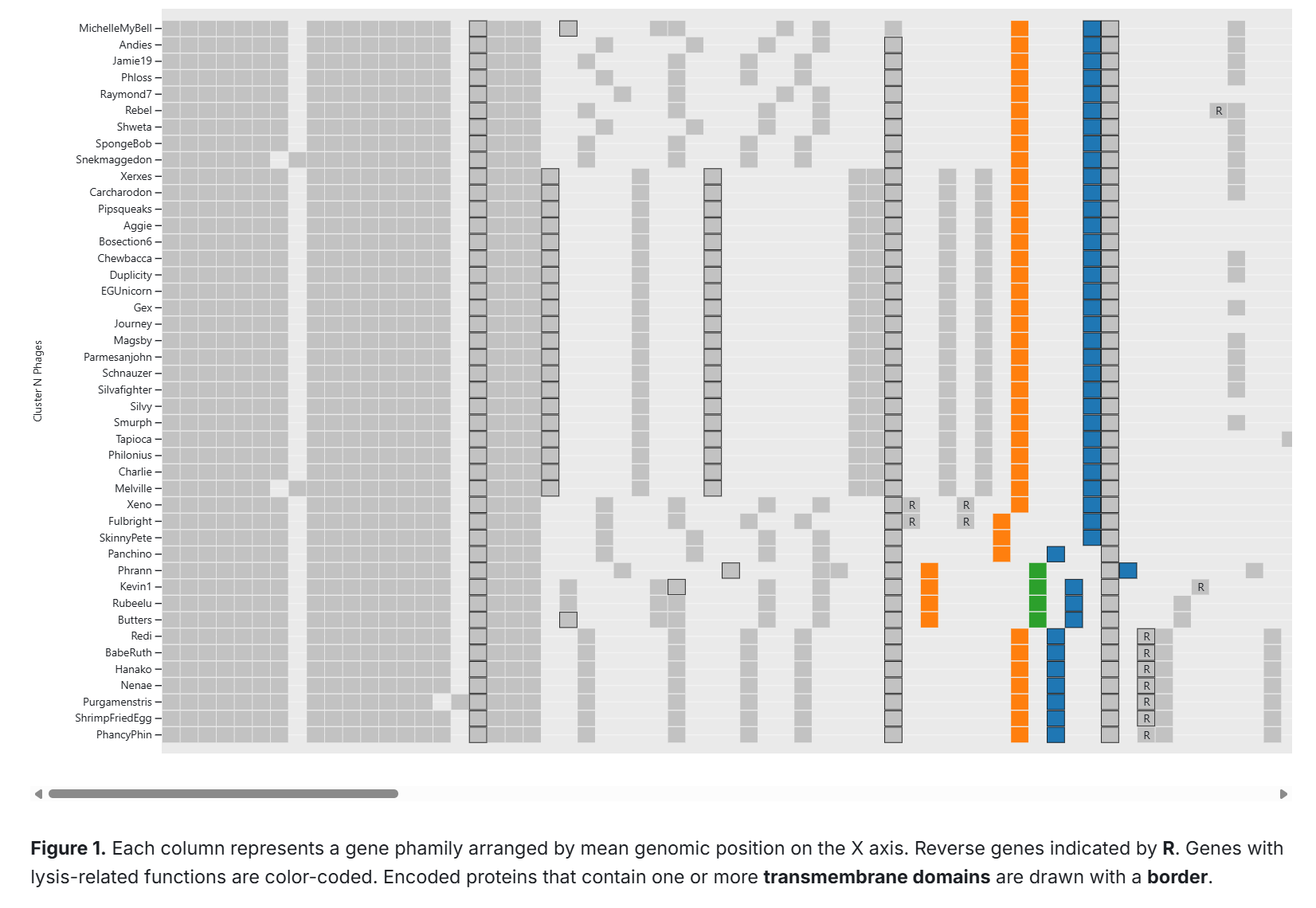
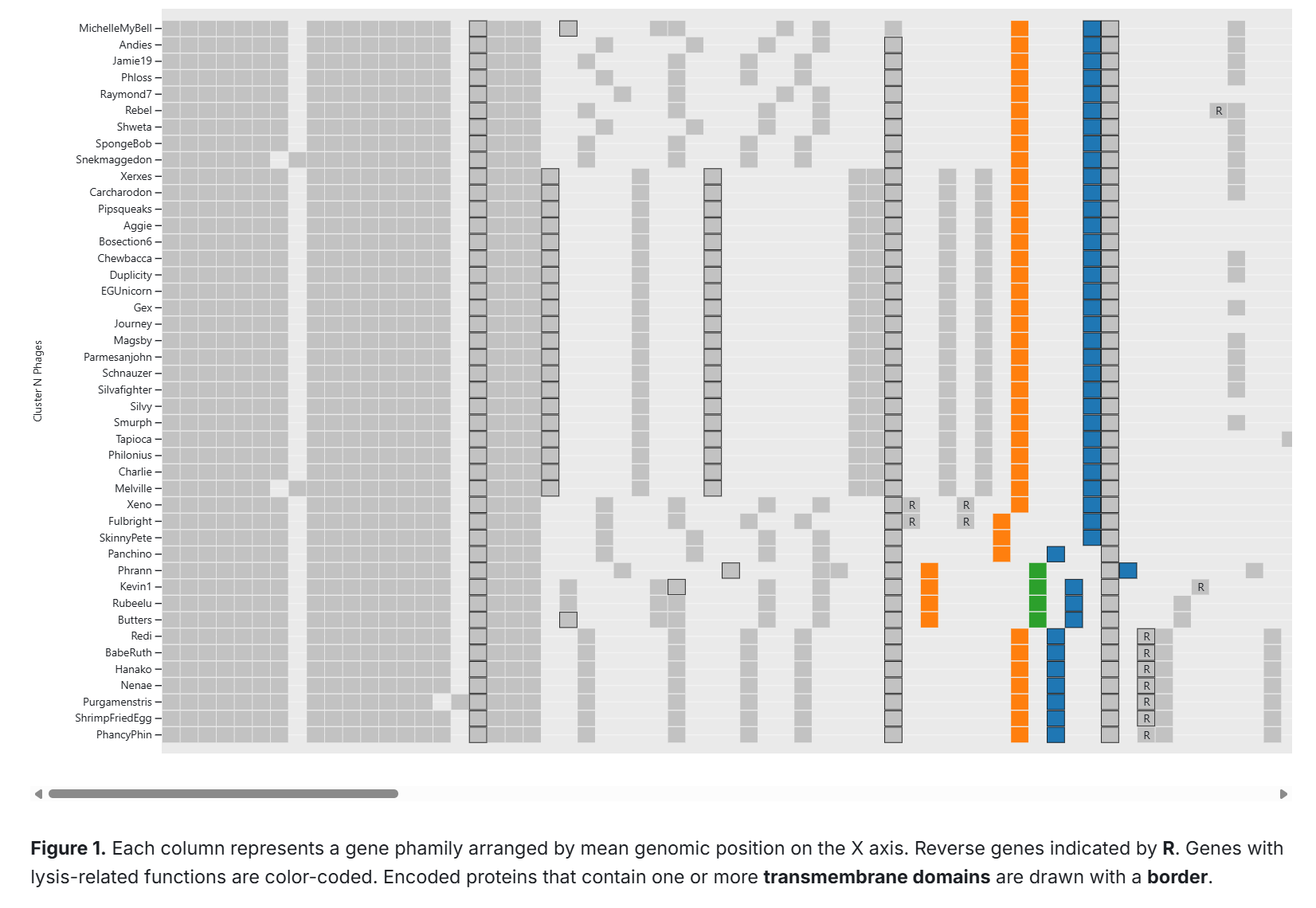
**Observable**

1. Bendele, M., Cobb, I., and Cresawn, S. *Subclusters*. Accessed June 2, 2025. Observable. <https://observablehq.com/d/5e5bc78c9b3ae2ed>
2. Cobb, I., Cooper, K., Bendele, M., Fisher R., Jones, Z., Shifflett, Z., Cresawn, S., *Pham Matrices*. Accessed June 2, 2025. Observable. <https://observablehq.com/@cresawn-labs/pham-matrix>.

**Other**

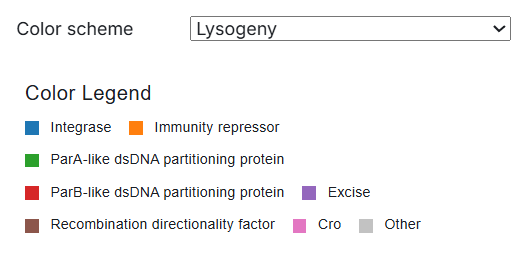
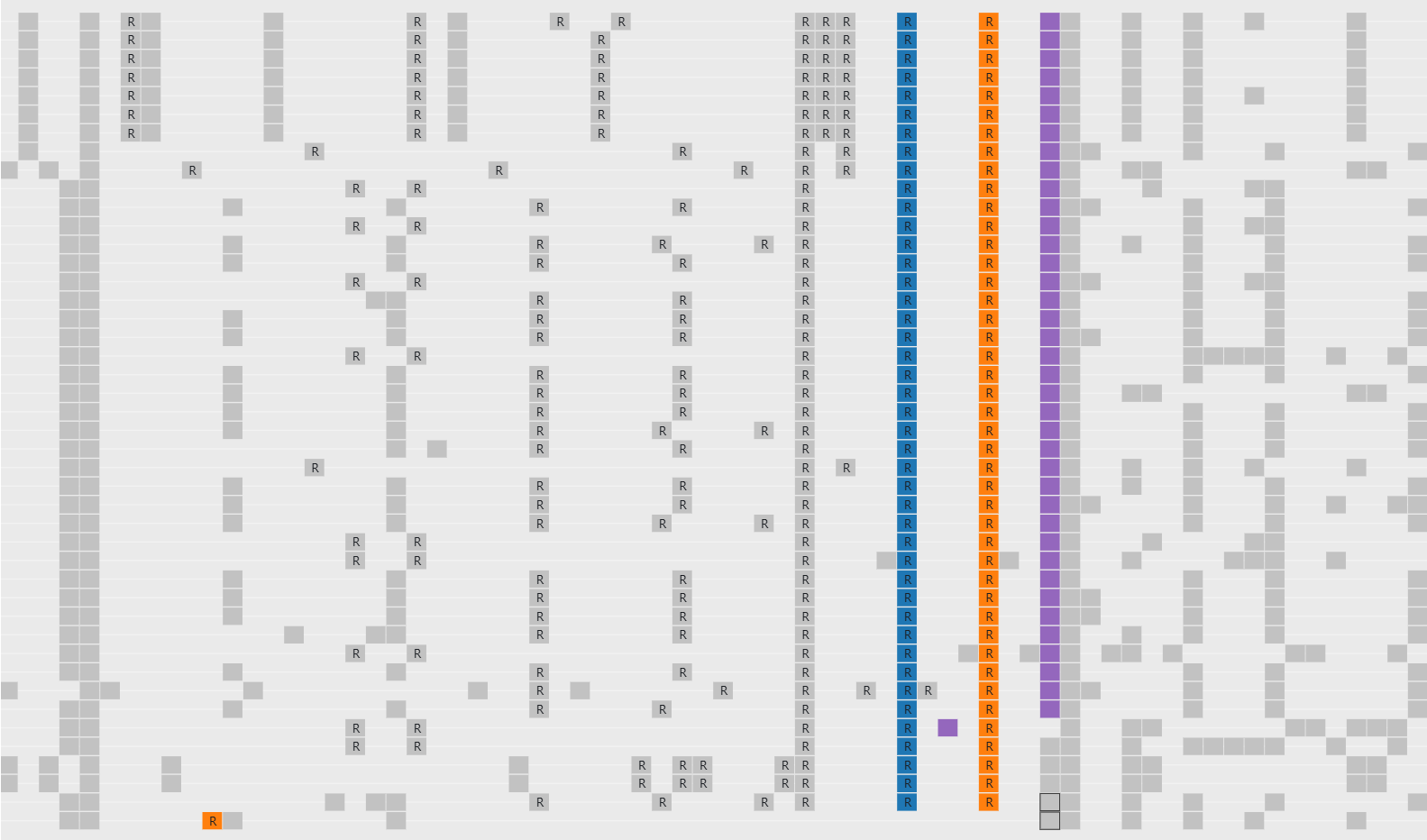
1. Pollenz RS, Bland J, Pope WH (2022) Bioinformatic characterization of endolysins and holin-like membrane proteins in the lysis cassette of phages that infect *Gordonia rubripertincta*. PLOS ONE 17(11): e0276603. <https://doi.org/10.1371/journal.pone.0276603>

**Appendix A – Genes Associated with Lysis in Cluster N**



Cobb et al., *Pham Matrices*. Accessed June 2, 2025. Observable. <https://observablehq.com/@cresawn-labs/pham-matrix>.

**Appendix B – Gene Associated with Lysogeny in Cluster N**



Cobb et al., *Pham Matrices*. Accessed June 2, 2025. Observable. <https://observablehq.com/@cresawn-labs/pham-matrix>.