**Article title**

Complete genomes of five phietaviruses infecting *Staphylococcus aureus*

**Authors**

Taylor P. Andrews[a], J. Steen Hoyer[b], Nicole L. Fahrenfeld[c], Siobain Duffy[b#], Jeffrey M. Boyd[d#]

**Affiliations**

a. Microbial Biology graduate program, Rutgers University, New Brunswick, NJ

b. Department of Ecology, Evolution and Natural Resources, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ

c. Department of Civil and Environmental Engineering, School of Engineering, Rutgers University, Piscataway, NJ

d. Department of Biochemistry and Microbiology, School of Environmental and Biological Sciences, Rutgers University, New Brunswick, NJ

**Running title**

Phietaviruses infecting *S. aureus*

**Corresponding authors' email addresses**

#Address correspondence to Jeffrey Boyd, jeffboyd@sebs.rutgers.edu; Siobain Duffy duffy@sebs.rutgers.edu

**Abstract**

The annotated whole genome sequences of five cultured phietaviruses infecting *Staphylococcus aureus* are presented. They are closely related to prophages previously sequenced as part of *S. aureus* genomes. Three of these viruses were confirmed to be temperate in laboratory.

**Announcement**

*Staphylococcus aureus* is a human commensal bacterium that has the potential to cause life-threatening infection (1). Its interactions with bacteriophages are an increasingly studied part of microbiome studies (2). We present the annotated genomes of five, plaque-purified *S. aureus* temperate phages in genus *Phietavirus* (3). Four aliquots of municipal wastewater influent from a Mid-Atlantic, US, treatment plant were collected in March 2021. To enrich for *S. aureus* phages, five mL of each sample was co-cultured with *S. aureus* RN4220 (4) in tryptic soy broth containing 10 mM $CaCl_2$ (TSB, 5). Phages were isolated using centrifugation and 0.22 µm filtration before plating with *S. aureus* RN4220 using the pour-plate technique. Plaques underwent three rounds of subculturing through single plaques to yield purified phage stocks (6). Transduction analysis was conducted on the isolated phages using a chloramphenicol-resistant *S. aureus* donor strain (7, 8). Each of the phages were cocultured with the donor strain in TSB with 10 µg $mL^{-1}$ chloramphenicol. Phages were isolated by 0.22 µm filtration and cocultured with chloramphenicol-sensitive *S. aureus* RN4220 for one hour. The recipient cells were pelleted by centrifugation and free phage removed by decanting. The bacterial pellet was resuspended in 100 mM sodium citrate, and plated on selective on solid tryptic soy medium containing 10 µg $mL^{-1}$ chloramphenicol. Three phages with transducing ability (SAP1, SAP2 and SAP13, for *S. aureus* phage) and two with negative results (SAP3, SAP8) were chosen for further analysis. The DNA genomes of these five phages were extracted using QIAamp MinElute Virus Spin Kits.

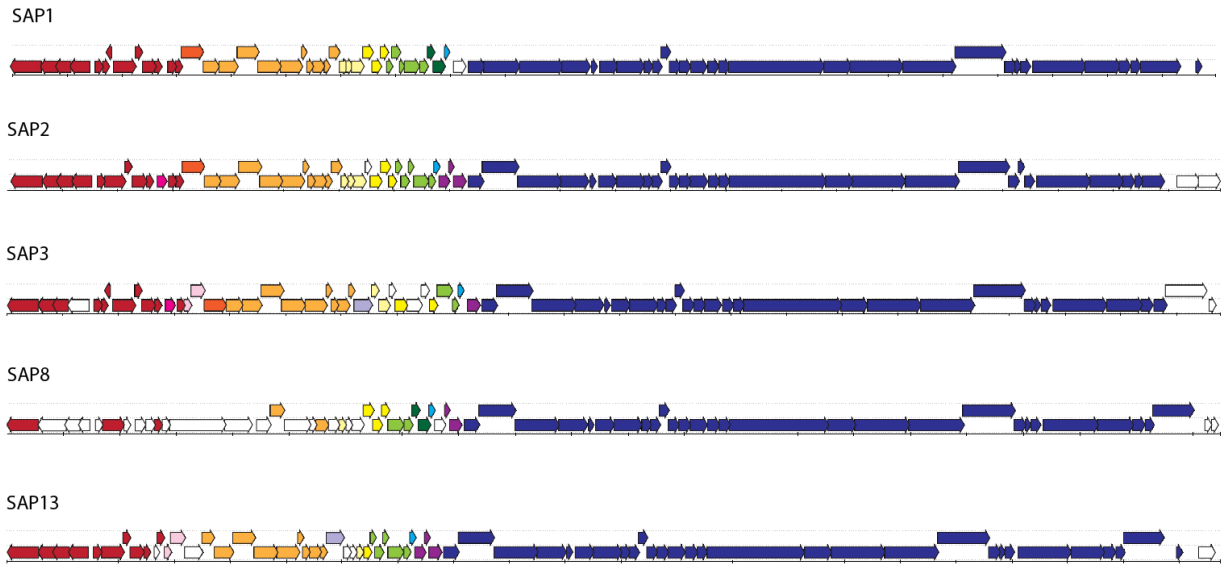
Paired end (2x150bp) Illumina sequencing was performed on NextSeq 2000 at MiGS (Microbial Genome Sequencing Center). Reads were analyzed using CPT Galaxy Phage genome assembler v2021.01 Workflow (9), which produced linear contigs with small overlaps at the end which suggested the genomes were circular. The overlaps were manually cut. Taxonomic assignment of five genomes with dsDNA phage genomes was performed with GRAViTy v1.1.0 (10), which showed they were phietaviruses (symmetrical Theil's U(Ref, Pred): 0.863) related to SAP26 (GU477322, which was arbitrarily linearized). The genomes were reoriented to reflect the termini of *Staphylococcus* prophages from a closely related genus (*e.g.* DQ530359). Genome annotation was performed as previously published (11, 12): ORFs were annotated using Prokka (parameters Genus: *Phietavirus*, Kingdom: Viruses) (v1.14.6, Galaxy) (13), further annotated for functionality with PHROGs v4 (14) database and Phyre2 v2.0 (15), and non-protein-coding features were identified including tRNAs (tRNAscan-SE v. 2.0) (16), terminators (ARNold v1.0) (17), ncRNAs (Rfam v14.8) (18), and promoters (Genome2D Prokaryote Promoter Prediction) (19). Sequence coverage was calculated using Map with BWA-MEM (v0.7.17.2, Galaxy) (20) and Samtools depth (v1.13, Galaxy) (21). Default parameters were used except where otherwise noted.

The five SAP genomes are ~43KB (Table 1) and portions of the genomes are very similar to one another (the most divergent pair, SAP1 and SAP8, are ≥94% identity by BLAST over 60% of the genome). There was significant synteny between the 63-69 ORFS of the genomes (Figure 1). The closest BLAST hits to these phage genomes in the NCBI nr database are all prophages within *S. aureus* genomes (e.g., SAP3 is 100% identical, 100% query cover by BLAST to CP051919).

**Table 1**: Summary of SAP Genomic Characteristics

| Phage | Genome Length (bp) | # of Predicted ORFs | # of Putative Promoters | # of Putative Rho-independent Terminators | Average Sequencing Coverage | GC Content | GenBank accession number |
|---|---|---|---|---|---|---|---|
| SAP1 | 43,962 | 68 | 10 | 22 | 9,518x | 34.3% | ON911714 |
| SAP2 | 43,863 | 69 | 6 | 23 | 9,069x | 34.0% | ON911715 |
| SAP3 | 43,586 | 66 | 11 | 18 | 11,412x | 34.6% | ON911716 |
| SAP8 | 42,981 | 63 | 8 | 20 | 11,997x | 34.1% | ON911717 |
| SAP13 | 43,478 | 67 | 10 | 25 | 11,145x | 34.6% | ON911718 |

**Figure 1:** Genomic maps of the five phage genomes. Colors indicate blocks of homology, ORFs without homology with other SAP genomes are depicted in white. All have integrase genes at the 5' end indicating they are likely capable of lysogeny. They share a large, syntenous block of genes towards the 3' end containing structural and hypothetical proteins.




## Data availability

Genomes are in GenBank: accession numbers ON911714 (SAP1), ON911715 (SAP2), ON911716 (SAP3), ON911717 (SAP8), ON911718 (SAP13). Illumina data are available in the NCBI SRA (PRJNA857681). The phages are available by request from the corresponding authors.

**Acknowledgements**

This work was supported by NIAID 1R01AI139100-01, NSF 1750624 and USDA MRF project NE−1028 to JMB, NSF 1453241 to SD, and Rutgers Center for COVID-19 Response and Pandemic Preparedness grant to NLF. TPA was supported by NIGMS NIH 1T32GM139804-01. We thank the utility partner (who wished to remain anonymous) for providing the wastewater influent. We thank the Center for Phage Technology for their open resources and Evelien Adriaenssens for her input on analysis.

**References**

1. Tong SYC, Davis JS, Eichenberger E, Holland TL, Fowler VG. 2015. Staphylococcus aureus Infections: Epidemiology, Pathophysiology, Clinical Manifestations, and Management. *Clinical Microbiology Reviews* **28:**603-661.
2. Federici S, Nobs SP, Elinav E. 2021. Phages and their potential to modulate the microbiome and immunity. *Cellular & Molecular Immunology* **18:**889-904.
3. Gutiérrez D, Adriaenssens EM, Martínez B, Rodríguez A, Lavigne R, Kropinski AM, García P. 2014. Three proposed new bacteriophage genera of staphylococcal phages: “3alikevirus”, “77likevirus” and “Phietalikevirus”. *Archives of Virology* **159:**389-398.
4. Kreiswirth BN, Löfdahl S, Betley MJ, O'Reilly M, Schlievert PM, Bergdoll MS, Novick RP. 1983. The toxic shock syndrome exotoxin structural gene is not detectably transmitted by a prophage. *Nature* **305:**709-712.
5. Price Erin E, Rudra P, Norambuena J, Román-Rodríguez F, Boyd Jeffrey M, Semrau Jeremy D. 2021. Tools, Strains, and Strategies To Effectively Conduct Anaerobic and Aerobic Transcriptional Reporter Screens and Assays in Staphylococcus aureus. *Applied and Environmental Microbiology* **87:**e01108-01121.
6. Hyman P. 2019. Phages for Phage Therapy: Isolation, Characterization, and Host Range Breadth. *Pharmaceuticals* **12:**35.
7. Mashruwala AA, Boyd JM. 2017. The Staphylococcus aureus SrrAB Regulatory System Modulates Hydrogen Peroxide Resistance Factors, Which Imparts Protection to Aconitase during Aerobic Growth. *PLOS ONE* **12:**e0170283.
8. Bezar IF, Mashruwala AA, Boyd JM, Stock AM. 2019. Drug-like Fragments Inhibit agr-Mediated Virulence Expression in Staphylococcus aureus. *Scientific Reports* **9**.
9. Ramsey J, Rasche H, Maughmer C, Criscione A, Mijalis E, Liu M, Hu JC, Young R, Gill JJ. 2020. Galaxy and Apollo as a biologist-friendly interface for high-quality cooperative phage genome annotation. *PLOS Computational Biology* **16:**e1008214.
10. Aiewsakun P, Simmonds P. 2018. The genomic underpinnings of eukaryotic virus taxonomy: creating a sequence-based framework for family-level virus classification. *Microbiome* **6**.
11. Turner D, Adriaenssens EM, Tolstoy I, Kropinski AM. 2021. Phage Annotation Guide: Guidelines for Assembly and High-Quality Annotation. *PHAGE* **2:**170-182.
12. Shen A, Millard A. 2021. Phage Genome Annotation: Where to Begin and End. *PHAGE* **2:**183-193.
13. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30:**2068-2069.
14. Terzian P, Eric, Galiez C, Lossouarn J, Rubén, Mom R, Toussaint A, Petit M-A, Enault F. 2021. PHROG: families of prokaryotic virus proteins clustered using remote homology. *NAR Genomics and Bioinformatics* **3**.

15. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. 2015. The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* **10:**845-858.
16. Chan PP, Lowe TM. 2019. tRNAscan-SE: Searching for tRNA Genes in Genomic Sequences, p 1-14 doi:10.1007/978-1-4939-9173-0_1. Springer New York.
17. Naville M, Ghuillot-Gaudeffroy A, Marchais A, Gautheret D. 2011. ARNold: A web tool for the prediction of Rho-independent transcription terminators. *RNA Biology* **8:**11-13.
18. Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, Rivas E, Eddy SR, Robert, Bateman A, Petrov AI. 2021. Rfam 14: expanded coverage of metagenomic, viral and microRNA families. *Nucleic Acids Research* **49:**D192-D200.
19. Baerends RJ, Smits W, De Jong A, Hamoen LW, Kok J, Kuipers OP. 2004. Genome2D: a visualization tool for the rapid analysis of bacterial transcriptome data. *Genome Biology* **5:**R37.
20. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv: Genomics*.
21. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25:**2078-2079.