Discovering Gene Functions in Mycobacteriophage Sbash Using a Genetic Screen

Erin Brouwer,* Daiki Jonouchi,* Byron Noordewier, Sara S. Tolsma
Department of Biology, Northwestern College, Orange City, IA 51041

*Contributed Equally

Abstract

Sbash is a temperate bacteriophage, which was isolated on the host, *Mycobacterium smegmatis* mc^2 155 from soil collected in South Africa. It is classified as cluster I and sub-cluster I2. Its genome consists of 55,832 base pairs and 89 protein-coding genes of which only 25 genes were assigned a function by bioinformatic analysis. We are using a genetic screen to uncover the functions of phage genes for which function is unknown. To begin to uncover the functions of the protein products of the genes in Sbash's genome, we cloned each gene into the pExTra plasmid and assayed each phage gene for two phenotypes: cytotoxicity, the ability to interfere with host cell growth, and defense, the ability to protect the host cell from infection by other phages. In total, we successfully cloned approximately half of the genes in Sbash's genome with sizes ranging from 90 bp to nearly 1500 bp. We identified two Sbash genes that defend host cells from infection by other mycobacteriophages. We identified six genes that reduced the growth of host cells when expressed. Here, we report our progress on this project. We have also analyzed genes in Mycobacteriophage Island3, a cluster I1 phage, for cytotoxicity and defense to complete the screen of this phage started by students in previous research groups.

Introduction

Bacteriophages (phages) are viruses that infect bacteria. Understanding bacteriophages contributes to our understanding of microbiology, molecular biology, and genetics in important ways including revealing the general character of viruses, the molecular mechanisms of genes, the mechanisms of evolution, and the understanding of ecosystems. Recently, bacteriophages have gotten attention as a new treatment for bacterial infections, especially for antibiotic-resistant bacteria (Strathdee, et al. 2023).

SEA-GENES (Science Education Alliance—Gene-function Elucidation by a Network of Emerging Scientists) is an educational project sponsored by HHMI (Howard Hughes Medical Institute) that promotes research experiences for undergraduate students. As part of SEA-PHAGES (Science Education Alliance—Phage Hunters Advancing Genomic and Evolutionary Science) and SEA-GENES, students find novel phages then annotate, and analyze their genomes. Recent data reveal that students have isolated over 20,400 novel bacteriophages, 4,000 of those phage genomes have been annotated and submitted to GenBank and phagesDB, where the genomic data can be viewed (Heller & Sivanathan, 2022; Russell & Hatfull, 2016). The SEA-GENES project is the third phase of SEA courses in which students investigate the function of genes using an in-depth genetic screen (Heller et al., 2022). We clone each gene in a phage genome such that the gene is under the control of a tet-inducible promoter and upstream of the mCherry reporter gene. Then we assay each gene in the host bacterium, *Mycobacterium smegmatis*, in two ways: interference with the growth of host cells (cytotoxicity) and protection of host cells from infection by a mycobacteriophage (defense).

Specifically, we focused our work on phage Sbash. Sbash is a temperate, cluster I2 phage isolated from soil collected in 2013 in Durban, South Africa. It infects the host, *M. smegmatis*. Sbash's genome is 55,832 bp in length and includes of 89 genes, only 25 of which have assigned functions by bioinformatics. Island3 is a temperate, cluster I1 phage isolated from soil collected in Oakland, PA in 2008. Island3 consists of 76 genes, all of which have been cloned. In the present study, we built off of previous work and screened some genes for cytotoxicity and defense in an effort to confirm findings and complete the genetic screen.

Materials and Methods

Molecular Cloning.
Each gene was PCR-amplified using gene-specific primers and the high fidelity Taq, Q5 (NEB, Ipswich, MA). The primers (IDT, Coralville, IA) were designed to add homology sequences to each end of the gene. Amplicon sizes were analyzed by gel electrophoresis and purified directly (Zymo, Irvine, CA) or from gels (Zymo, Irvine, CA). Clean amplicons were cloned into a pExTra backbone using isothermal assembly () and reaction products were transformed into chemically competent *E. coli* (NEB, Ipswich, MA). Select kanamycin-resistant colonies were assayed by PCR using universal primers and Go-Taq (NEB, Ipswich, MA). Amplicons were analyzed by gel electrophoresis to confirm that a gene of the expected size was present in the pExTra plasmid. Cells from one colony exhibiting an amplicon of the expected size were grown, frozen, and used to prepare purified plasmid (Zymo, Irvine, CA).

Phenotypic assays.
Plasmids containing phage genes were transformed into electrocompetent *M. smegmatis*. Three kan-resistant clones were picked and assayed for cytotoxicity by plating serial dilutions on plates containing kanamycin (Gold Biotech, Olivette, MO), kanamycin + 10 ng/ml aTc (to induce expression of the phage gene), and kanamycin + 100 ng/ml aTc (ThermoFisher, Waltham, MA). To assay for defense, two kan-resistant clones were grown to saturation in liquid culture and used to plate a top agar lawn on which serial dilutions of high titer phage lysates were spotted (https://seagenes.helpdocsonline.com/home).

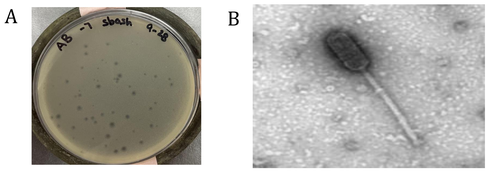
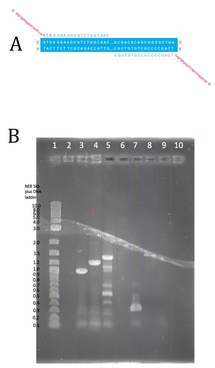


Figure 1. Sbash characteristics. A. Plaque assay of Sbash. Medium sized and round shaped plaques, slightly cloudy after 24 hours at 37°C. B. Transmission electron micrograph (TEM) of phage Sbash.

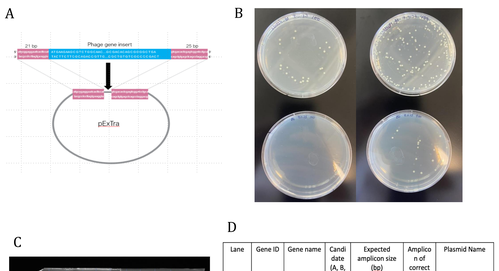
Phage	Genome size(bp)	Cluster	Genome Ends	%GC	Number of Annotated Genes	Isolation Location
Sbash	55,832	I2	10 bp, 3' sticky ends	65.6	89	Durban, South Africa

Table 1. Characteristics of Sbash genome.



Lane	Gene ID	Gene name	Expected amplicon size (bp)	Amplicon of correct size generated? (Y/N)
1	Ladder			
2				
3	000025	*Sbash_25*	*938+46=982*	*Y*
4	000026	*Sbash_26*	*1173+46=1219*	*Y*
5	000031	*Sbash_31*	*498+46=544*	*Y*
6				
7	000086	*Sbash_86*	*279+46=325*	*Y*
8	000087	*Sbash_87*	*1437+46=1483*	*N*
9	000089	*Sbash_89*	*369+46=415*	*N*
10				

Figure 2. PCR from high titer lysate using gene-specific primers. A. Primers include gene-specific sequences (blue) and homology sequences (pink) that add a total of 46 bp to the expected amplicon size for each gene. B. Agarose gel with representative PCR products using gene-specific primers with homology sequences. C. Data summary of PCR from high titer lysate.



Lane	Gene ID	Gene name	Candidate (A, B, C, etc.)	Expected amplicon size (bp)	Amplicon of correct size generated? (Y/N)	Plasmid Name
1	00086	*Sbash_86*	A	279+108=387	Y	pExTra_Sbash_86
2	00086	*Sbash_86*	B	279+108=387	N	
3	00086	*Sbash_86*	C	279+108=387	Y	
4	00087	*Sbash_87*	A	1437+108=1545	N	pExTra_Sbash_87
5	00087	*Sbash_87*	B	1437+108=1545	N	
6	00087	*Sbash_87*	C	1437+108=1545	N	
7	00089	*Sbash_89*	A	369+108=477	Y	pExTra_Sbash_89
8	00089	*Sbash_89*	B	369+108=477	Y	
9	00089	*Sbash_89*	C	369+108=477	Y	

Figure 3. Cloning into pExTra by isothermal assembly, chemical transformation of *E. coli*, and clone verification PCR. A. pExTra plasmid and amplicon contain homology sequences. B. Isothermal assembly reactions are chemically transformed into competent *E. coli* and plated on LB-kanamycin plates. 3 clones are assayed by PCR to confirm the insert size. C. Agarose gel with representative clone verification results using universal primers complementary to sequences just up and downstream of the phage gene produce amplicons that are the gene size plus 108 bp. D. Summary of clone verification PCR.

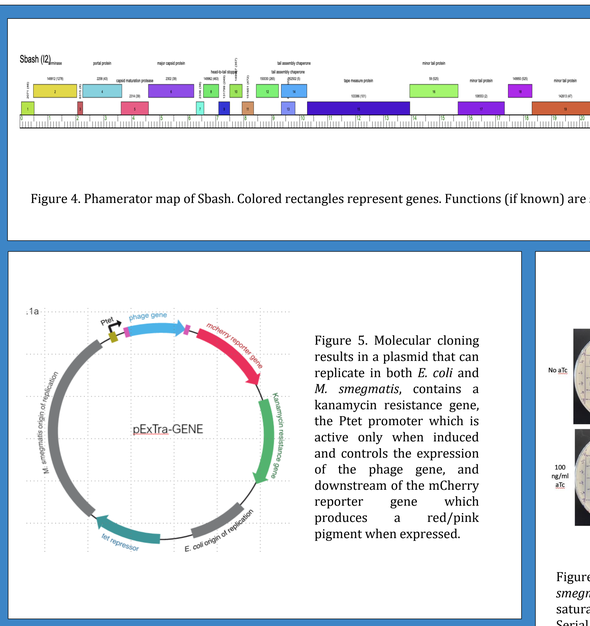
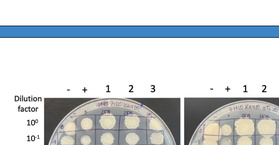


Figure 4. Phamerator map of Sbash. Colored rectangles represent genes. Functions (if known) are shown above or below the rectangles. Genes above tape are forward genes. Genes below tape are reverse genes.

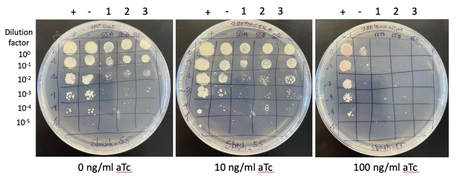
Figure 5. Molecular cloning results in a plasmid that can replicate in both *E. coli* and *M. smegmatis*, contains a kanamycin resistance gene, the Ptet promoter which is active only when induced and controls the expression of the phage gene, and downstream of the mCherry reporter gene which produces a red/pink pigment when expressed.



Lane	Gene ID	Plasmid name	Gene name	Toxic/Non-toxic	Colony color on 100 ng/ml aTc plate*
- Non-toxic control	--	pExTra03	*Fruitloop_52* mutant	Non-toxic	++
+ Toxic control	--	pExTra02	*Fruitloop_52*	Toxic	-
1	00025	pExTra_Sbash_25	*Sbash_25* replicate 1	Non-toxic	+
2	00025	pExTra_Sbash_25	*Sbash_25* replicate 2	Non-toxic	+
3	00025	pExTra_Sbash_25	*Sbash_25* replicate 3	Non-toxic	+

*Key: NG (no growth) - (no pink color) +(faint pink color) ++(obvious pink color) +++ (dark pink color)

Figure 6. Representative Cytotoxicity Assay. Serial dilutions of three *M. smegmatis* clones transformed with Sbash_25 plated on three 7H10 plates: no aTc, 10 ng/ml aTc, or 100 ng/ml aTc.
– represents Fruitloop gene 52 with an inactivating mutation and + represents a known cytotoxic gene from phage Fruitloop (gene 52). These data indicate that Sbash_25 is not cytotoxic.



Lane	Gene ID	Plasmid name	Gene name	Toxic/Non-toxic	Colony color on 100 ng/ml aTc plate*
+ Toxic control	--	pExTra02	*Fruitloop 52*	Toxic	++
- Non-toxic control	--	pExTra03	*Fruitloop 52* mutant	Non-toxic	-
1	00055	pExTra_Sbash_55	*Sbash_55* replicate 1	Toxic	NG
2	00055	pExTra_Sbash_55	*Sbash_55* replicate 2	Toxic	NG
3	00055	pExTra_Sbash_55	*Sbash_55* replicate 3	Toxic	NG

*Key: NG (no growth) - (no pink color) +(faint pink color) ++(obvious pink color) +++ (dark pink color)

Figure 7. Representative Cytotoxicity Assay. Serial dilutions of three *M. smegmatis* clones transformed with Sbash_55 plated on three 7H10 plates: no aTc, 10 ng/ml aTc, or 100 ng/ml aTc.
– represents Fruitloop gene 52 with an inactivating mutation and + represents a known cytotoxic gene from phage Fruitloop (gene 52). These data indicate that Sbash_55 is cytotoxic.

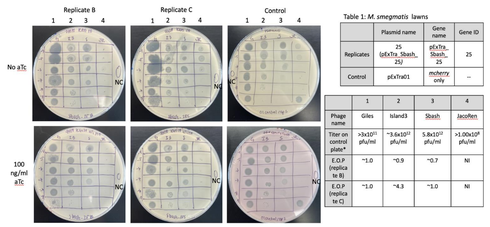


Table 1: *M. smegmatis* lawns

	Plasmid name	Gene name	Gene ID
Replicates	25 (pExTra_Sbash_25)	pExTra_Sbash_25	25
Control	pExTra01	*mcherry* only	--

	1	2	3	4
Phage name	Giles	Island3	Sbash	Jacoben
Titer on control plate*	$\times 3\times10^{11}$ pfu/ml	$\sim 3.6\times10^{12}$ pfu/ml	5.8×10^{12} pfu/ml	$\times 1.00\times10^{8}$ pfu/ml
E.O.P. (replicate B)	~1.0	~0.9	~0.7	NI
E.O.P. (replicate C)	~1.0	~4.3	~1.0	NI

Figure 8. Representative Defense Assay. Two replicates (B and C) of *M. smegmatis* transformed with Sbash gene 25 were grown in liquid culture to saturation and incorporated into a top agar lawn with or without 100 ng/ml aTc. Serial dilutions of four mycobacteriophages were spotted on the lawns. Similar numbers of plaques were seen on control plate (pExtra01) or on plates with *M. smegmatis* expressing Sbash gene 25 suggesting that gene 25 does not confer defense against infection by another mycobacteriophage.

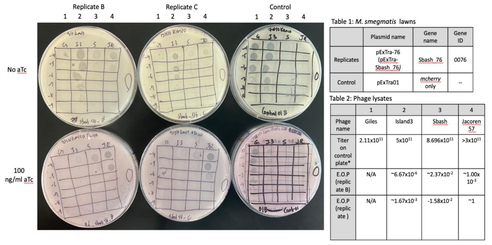


Table 1: *M. smegmatis* lawns

	Plasmid name	Gene name	Gene ID
Replicates	pExTra-76 (pExTra_Sbash_76)	Sbash_76	0076
Control	pExTra01	*mcherry* only	--

Table 2: Phage lysates

	1	2	3	4
Phage name	Giles	Island3	Sbash	Jacoren 57
Titer on control plate*	2.11×10^{11}	5×10^{11}	8.696×10^{11}	$>3\times10^{11}$
E.O.P. (replicate B)	N/A	$\sim 6.67\times10^{-3}$	$\sim 2.37\times10^{-2}$	$\sim 1.00\times10^{-3}$
E.O.P. (replicate)	N/A	$\sim 1.67\times10^{-3}$	$\sim 1.58\times10^{-2}$	~1

Figure 9. Representative Defense Assay. Two replicates (B and C) of *M. smegmatis* transformed with Sbash_76 were grown in liquid culture to saturation and incorporated into a top agar lawn with or without 100 ng/ml aTc. Serial dilutions of four mycobacteriophages were spotted on the lawns. More plaques were seen on control plate (pExtra01) than on plates with *M. smegmatis* expressing Sbash_76 suggesting that gene 76 does confer defense against infection by Island3 and Sbash.

Table 2. Summary of Results from Phage Sbash.	
Genes Cloned	41 of 89 genes
Cytotoxic Genes	5 genes
Non-toxic Genes	33 genes
Genes Conferring Defense	8 genes
Genes That Do Not Confer Defense	29 genes

Table 3. Annotated Functions for Cytotoxic Genes and Genes that Confer Defense

Cytotoxic Sbash Genes	Annotated function	**Genes that Confer Defense**	Annotated function
12	Unknown	56	Unknown
40	Tyrosine integrase	76	HNH endonuclease
55	Unknown		
59	Unknown		
75	Unknown		
76	HNH endonuclease		

Discussion and Future Directions

Understanding bacteriophages contributes to our understanding of topics in scientific study, the character of viruses and genes, and to the advancement of phage therapy (Strathdee et al., 2023). The SEA-PHAGES and SEA-GENES projects, sponsored by the HHMI, expand the phage database and known knowledge of the functions of phages and genes within their genomes (Heller et al., 2022). We are half of the way through the genetic screen of mycobacteriophage Sbash. In combination with previous work on the genome, 41 genes of Sbash's genome have been successfully cloned. Additionally, we have found 5 cytotoxic genes and 8 genes that confer defense in Sbash's genome in total. We are continuing to clone and analyze Sbash genes. We are carefully curating the Sbash data and plan to reclone and retest many genes, ensuring that any cytotoxic gene or gene that confers defense has been repeated at least twice. We will sequence all cloned genes to ensure that our cloning did not introduce any new mutations. We have started working on writing a paper in which we will communicate the results of our Sbash genetic screen that will be submitted to the journal *G3*. We also continue to finalize our genetic screen data for phage Island3 which includes recloning and reassaying some genes. We are also in the process of writing a *G3* paper to report these findings.

References

Heller et al. 2017. PLOS Genetics 13(9): e1007007. https://doi.org/10.1371/journal.pgen.1007007

Heller et al. 2022. G3 (Bethesda). 2022 Jul 29;12(8):jkac140. doi: 10.1093/g3journal/jkac140. PMID: 35727726; PMCID: PMC9339283.

Russell and Hatfull. 2017. Bioinformatics 33:784–786.

Strathdee, et al. 2023. Cell 186: 17-31.