Using a Genetic Screen to Discover Gene Functions in Mycobacteriophages Sbash and Island3

Abigail M. Bastian,* Sadie A. Gilmeister,* Keziah J. Knudson,* Erika J. McKenney,* Garrett Raymon,* Anneka E. Sterk,* Alaena L. Trevino,* April S. Van Tol,* Byron Noordewier, Sara S. Tolsma
Department of Biology, Northwestern College, Orange City, IA 51041

*Contributed Equally

Abstract

Sbash is a temperate bacteriophage that infects *Mycobacterium smegmatis*. It was assigned to cluster I2 based on gene-content similarity of 35% or higher to sequenced bacteriophages present in the Actinobacteriophage database, phagesDB. Its genome was annotated in 2014 and found to include 89 protein-coding genes, only 22 of which were assigned functions based on bioinformatic analysis. We are using a genetic screen to identify functions of phage genes for which no function is currently known. We cloned 40 of the genes in Sbash's genome with sizes ranging from 90 bp to 3,666 bp. We screened each gene for cytotoxicity and identified six genes that reduced growth of the host cells when expressed. We also screened for defense, the ability of each gene product to protect the host cell from infection by another phage. We identified eight Sbash gene products that defend host cells from infection by other mycobacteriophages. 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 were discovered more than 100 years ago and while we have learned a lot about them in the last century there is still much we don't know about these ubiquitous infectious particles (Koskella, et al. 2022). Understanding bacteriophages contributes to our understanding of evolution, ecosystems, phage-host interactions, and is an effective therapy for antibiotic resistant infections (Strathdee, et al. 2023). The HHMI-sponsored SEA-PHAGES (Science Education Alliance—Phage Hunters Advancing Genomic and Evolutionary Science) program enlists undergraduate students to discover, purify, and characterize new actinobacteriophages in the Discovery portion and to annotate sequenced phage genomes in the Bioinformatics portion of the program. So far, SEA-PHAGES students have discovered more than 22,000 novel actinobacteriophages and annotated nearly 3,000 of their genomes (Russell and Hatfull. 2017). These data reveal that approximately 75% of the genes in annotated phage genomes cannot be assigned a function by bioinformatics alone. SEA-GENES (SEA—Gene-function Elucidation by a Network of Emerging Scientists) is an extension of SEA-PHAGES in which students conduct in-depth gene function investigations for entire mycobacteriophage genomes using a 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, *M. smegmatis*, in two ways: interference with the growth of host cells (cytotoxicity) and protection of host cells from infection by a mycobacteriophage (defense).

We have been working with two mycobacteriophages: Island3 and Sbash. Island 3 is a temperate, cluster I1 phage isolated from soil collected in Oakland, PA in 2008. Sbash is a temperate, cluster I2 phage isolated from soil collected in 2013 in Durban, South Africa. Both phages infect the host, *Mycobacterium smegmatis*. Sbash's genome is 55,832 bp in length and includes of 89 genes, only 24 of which have assigned functions by bioinformatics. Island3's genome is 47,287 bp long and contains 76 genes, only 17 of which have assigned functions.

We have cloned all of Island3's 76 genes and have assayed 74 of them for cytotoxicity and defense. Fifteen of these genes are cytotoxic and three show signs of defense. We have cloned 40 of Sbash's 89 genes and have assayed 39 of them for cytotoxicity and defense. Six Sbash genes are cytotoxic and eight show signs of defense against infection by another mycobacteriophage.

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).

Sequencing.
Plasmids with universal primers (or gene-specific primers for large genes) were sequenced by Azenta (Burlington, MA) and analyzed using SnapGene Viewer.

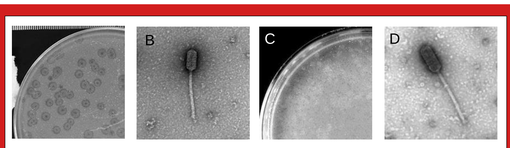


Figure 1. Phage characteristics. A. Plaque assay of phage Island3: 4mm/bulls-eye/24 hours/37C B. Transmission electron micrograph (TEM) of phage Island3 C. Plaque assay of phage Sbash: small sized and round shaped D. TEM of phage Sbash.

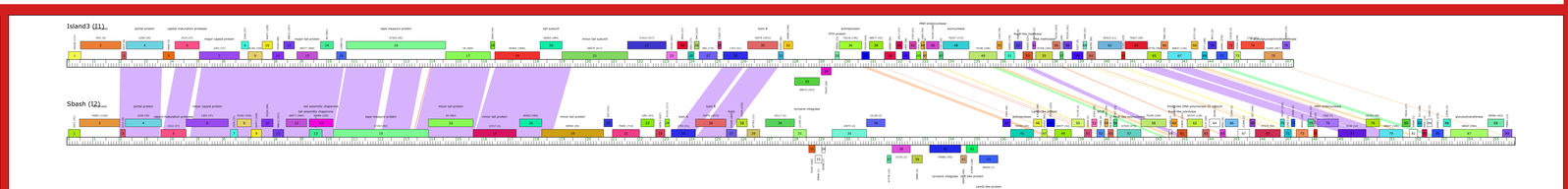


Figure 2. Phamerator map of Island3 and Sbash. Colored rectangles represent phage genes with functions (if known) printed above or below. Genes above tapeline are forward genes. Genes below tapeline are reverse genes. Purple regions indicate areas of high homology.

Table 1. Characteristics of Island3 and Sbash genomes.

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
Island3	47,287	I1	11 bp, 3' sticky ends	66.8	76	Oakland, PA

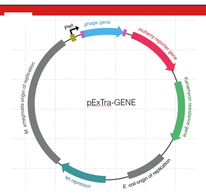


Figure 5. Cloning results in a plasmid that can replicate in both *E. coli* and *M. smegmatis*, containing a kanamycin resistance gene, the phage gene under control of the Ptet promoter which is active only when induced, and downstream of the mCherry reporter gene which produces a red/pink pigment when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

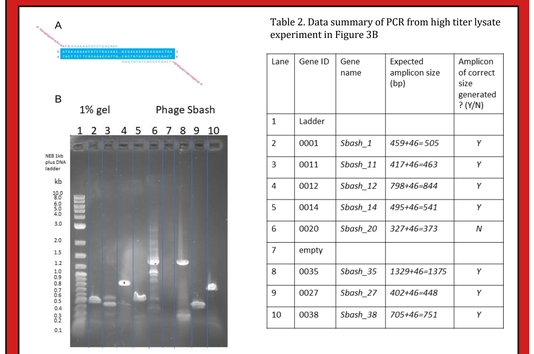


Table 2. Data summary of PCR from high titer lysate experiment in Figure 3B

Lane	Gene ID	Gene name	Expected amplicon size (bp)	Amplicon of correct size generated ? (Y/N)
1	Ladder			
2	0001	*Sbash_1*	*459+46=505*	*Y*
3	0011	*Sbash_11*	*417+46=463*	*Y*
4	0012	*Sbash_12*	*798+46=844*	*Y*
5	0014	*Sbash_14*	*495+46=541*	*Y*
6	0020	*Sbash_20*	*327+46=373*	*N*
7	empty			
8	0035	*Sbash_35*	*1329+46=1375*	*Y*
9	0027	*Sbash_27*	*402+46=448*	*Y*
10	0038	*Sbash_38*	*705+46=751*	*Y*

Figure 3. 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.

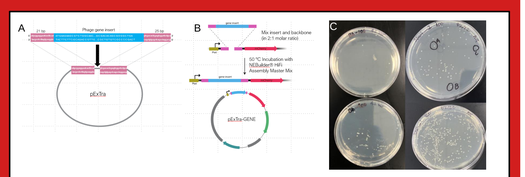
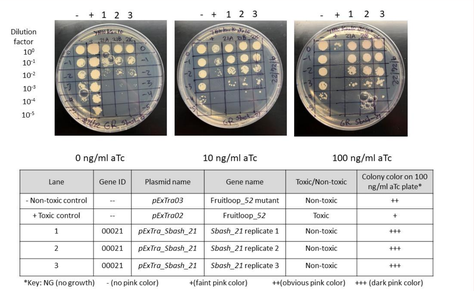


Table 3. Clone verification PCR summary data.

Lane	Gene ID	Gene name	Candidate (A, B, C, etc.)	Expected amplicon size (bp)	Amplicon of correct size generated? (Y/N)
L		Ladder			
1	00023	*Sbash_23*	*A*	*363+108=471*	*Y*
2	00023	*Sbash_23*	*B*	*363+108=471*	*Y*
3	00023	*Sbash_23*	*C*	*363+108=471*	*Y*
4	00049	*Sbash_49*	*A*	*648+108=756*	*Y*
5	00049	*Sbash_49*	*B*	*648+108=756*	*N*
6	00049	*Sbash_49*	*C*	*648+108=756*	*Y*
7	00076	*Sbash_76*	*A*	*816+108=924*	*Y*
8	00076	*Sbash_76*	*B*	*816+108=924*	*Y*
9	00076	*Sbash_76*	*C*	*816+108=924*	*Y*

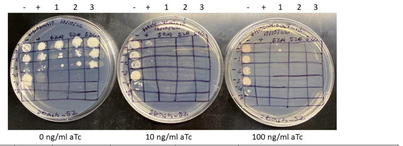
Figure 4. Cloning into pExTra by isothermal assembly, chemical transformation of *E. coli*, and clone verification PCR. A. pExTra plasmid and amplicon contain homology sequences. B. Thermolabile exonuclease creates 3' sticky overhangs, homology sequences hybridize, DNA polymerase fills gap, ligase covalently links the fragments. C. Isothermal assembly reactions are chemically transformed into competent *E. coli* and plated on LB-kanamycin plates. 3-4 clones are assayed by PCR to confirm the insert size. D. 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.



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	00021	*pExTra_Sbash_21*	*Sbash_21* replicate 1	Non-toxic	+++
2	00021	*pExTra_Sbash_21*	*Sbash_21* replicate 2	Non-toxic	+++
3	00021	*pExTra_Sbash_21*	*Sbash_21* 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 gene 21 plated on 7H10 plates with no aTc 10 ng/ml aTc or 100 ng/ml aTc. + represents a known cytotoxic gene from phage Fruitloop (gene 52) and – represents Fruitloop gene 52 with an inactivating mutation. These data indicate that Sbash_21 is not cytotoxic.



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	00052	*pExTra_Sbash_52*	*Sbash_52* replicate 1	Toxic	-
2	00052	*pExTra_Sbash_52*	*Sbash_52* replicate 2	Toxic	NG
3	00052	*pExTra_Sbash_52*	*Sbash_52* 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 gene 52 plated on 7H10 plates with no aTc 10 ng/ml aTc or 100 ng/ml aTc. + represents a known cytotoxic gene from phage Fruitloop (gene 52) and – represents Fruitloop gene 52 with an inactivating mutation. These data indicate that Sbash_52 is cytotoxic.

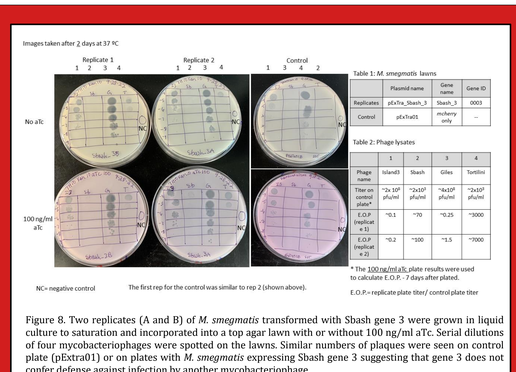


Table 1: *M. smegmatis* lawns

	Plasmid name	Gene name	Gene ID
Replicates	pExTra_Sbash_3	Sbash_3	0003
Control	pExTra01	mcherry only	--

Table 2: Phage lysates

	1	2	3	4
Phage name	Island3	Sbash	Giles	Tortilini
Titer on control plate*	~2 10⁸ pfu/ml	~2x10⁷ pfu/ml	~4x10⁶ pfu/ml	~2x10⁷ pfu/ml
E.O.P (replicate 1)	~0.1	~70	~0.25	~3000
E.O.P (replicate 2)	~0.2	~100	~1.5	~7000

* The 100 ng/ml aTc plate results were used to calculate E.O.P. - 7 days after plated.
E.O.P.= replicate plate titer/ control plate titer
NC= negative control
The first rep for the control was similar to rep 2 (shown above).

Figure 8. Two replicates (A and B) of *M. smegmatis* transformed with Sbash gene 3 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 3 suggesting that gene 3 does not confer defense against infection by another mycobacteriophage.

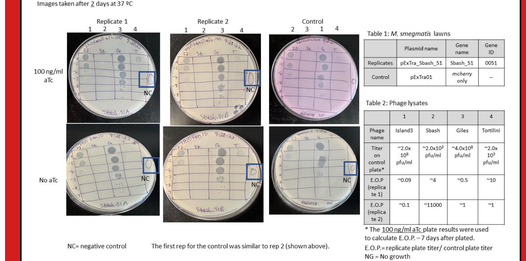


Table 1: *M. smegmatis* lawns

	Plasmid name	Gene name	Gene ID
Replicates	pExTra_Sbash_51	Sbash_51	0051
Control	pExTra01	mcherry only	--

Table 2: Phage lysates

	1	2	3	4
Phage name	Island3	Sbash	Giles	Tortilini
Titer on control plate*	~2.0x 10⁸ pfu/ml	~2.0x10⁸ pfu/ml	~4.0x10⁸ pfu/ml	~2.0x 10⁸ pfu/ml
E.O.P (replica te 1)	~0.09	~4	~0.5	~10
E.O.P (replica te 2)	~0.1	~11000	~1	~1

* The 100 ng/ml aTc plate results were used to calculate E.O.P. – 7 days after plated.
E.O.P.= replicate plate titer/ control plate titer
NG = No growth
NC= negative control
The first rep for the control was similar to rep 2 (shown above).

Figure 9. Two replicates (A and B) of *M. smegmatis* transformed with Sbash gene 51 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. Fewer plaques were seen on plates with *M. smegmatis* expressing Sbash gene 51 (aTc 100 ng/ml) compared to control plates (pExTra01) or plates without aTc suggesting that gene 3 confers defense against infection by Island3 (Cluster I1).

Table 4. Summary of Results from Phage Sbash

Genes cloned	40 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 5. Summary of Results from Phage Island3

Genes cloned	76 of 76 genes
Cytotoxic Genes	15 genes
Non-toxic Genes	59 genes
Genes Conferring Defense	3 genes
Genes That Do Not Confer Defense	69 genes

Table 6. Annotated Functions for Cytotoxic Genes and Genes that Confer Defense

Sbash gene	Annotated function	Island3 gene	Annotated function
	Cytotoxic Genes		**Cytotoxic Genes**
1	Unknown	6	Capsid maturation protease
37	Unknown	17	Unknown
40	Tyrosine integrase	18	Unknown
52	Unknown	32	Unknown
55	Unknown	35	HTH protein
66	Unknown	36	Antirepressor
75	Unknown	38	Unknown
	Confer Defense	39	Unknown
11	Unknown	41	Unknown
28	Unknown	53	RusA-like resolvase
32	Unknown	55	DNA methylase
51	Unknown	62	Unknown
54	Unknown	63	Unknown
55	Unknown	67	Unknown
66	Unknown	69	Unknown
75	Unknown		**Confer Defense**
		1	Unknown
		29	Unknown
		34	Unknown

Figure 10. Each pExTra plasmid containing a cloned Island3 or Sbash gene was/will be sequenced and compared to the original sequence of the gene in the phage genome to confirm that the gene sequence is what we expect it to be.

Discussion/Future Directions

We have nearly completed the genetic screen of mycobacteriophage Island3 and have made good progress on phage Sbash. We found 15 cytotoxic genes and 3 genes that confer defense in Island3's genome and 5 cytotoxic genes and 8 genes that confer defense in Sbash's genome. We are continuing to clone and analyze Sbash genes. We are carefully curating the Island3 data, ensuring that any gene for which we detected cytotoxicity or defense has been repeated at least twice and we have a few Island3 genes left to assay. We also need to complete the sequence analysis of the cloned Island3 genes. We have started working on writing a paper in which we will communicate the results of our Island3 genetic screen that we intend to submit to the journal *G3*. We are also working to reclone Island3 genes that display a cytotoxicity or defense into a2Hα so we can assay these genes in a bacterial two-hybrid (B2H) system (Heller, et al. 2017). Thus far, we have successfully cloned two genes, Island3_5 and Island3_39 into p2Hα. We hope to use the B2H system to identify binding partners in the host genome so we can further understand the functions of these phage genes.

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.

Koskella, et al. 2022. Annu. Rev. Virol. 9:57–78.

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

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