Summer 2023

Isolation and Identification of Phage-Host Protein Interactions as Possible Mediators of Mycobacteriophage Phayonce Gene Host Cytotoxicity

Matthew Defreitas
Coastal Carolina University, mjdefreit@coastal.edu

Isolation and Identification of Phage-Host Protein Interactions as Possible Mediators of Mycobacteriophage Phayonce Gene Host Cytotoxicity

By

Matthew Defreitas

Biology

Submitted in Partial Fulfillment of the
Requirements for the Degree of Bachelor of Science
In the HTC Honors College at
Coastal Carolina University

Spring 2023

Louis E. Keiner
Director of Honors
HTC Honors College

Daniel C. Williams
Associate Professor
Biology
Gupta College of Science

Abstract

Bacteriophages are ubiquitous viruses containing extremely diverse genomes. Many phage genomes have been bioinformatically annotated; however, many genes lack wet-bench functional characterization. Elucidating individual phage gene function via the isolation of phage-host protein interactions allows for the exploration of novel antibacterial therapies within the context of phage-host biology. Mycobacteriophage Phayonce, infecting the host *Mycobacterium smegmatis*, encodes two cytotoxic gene, 41 and 64, which lack an annotated function. Using a bacterial two-hybrid screen along with the construction of fusion-protein expression vectors, multiple punitive phage-host protein interactions were isolated for both genes. Once these interactions were verified and sequenced, multiple host *M. smegmatis* protein interacting partners were identified for gene 41 and one, unknown partner was isolated for gene 64. The host proteins identified using the screen establish a foundation for further investigation between the protein interactions between Phayonce and its host. Once such target identified for gene 41 was a prolyl oligopeptidase, which is predicted to degrade peptide bonds within the host. This may potentially be an explanation for gene 41's observed host cytotoxicity. Elucidating the phage-host protein interactions of Phayonce's cytotoxic genes will provide a greater understanding of phage biology, including the various ways that biology can be exploited in the development of novel antibacterial therapeutics.

Introduction

Rise of Antimicrobial Resistance in Mycobacteria

The rise and incidence of newly resistant strains of pathogenic bacteria is becoming an ever more increasing global health challenge (Lange 2018). Often the result of poorly administered and indiscriminate prescription of contemporary antibiotics for the treatment of bacterial infections, strains of previously sensitive bacteria, which are now resistant, have been selected for and allowed to proliferate (Lange 2018). Recently, multidrug resistant strains of mycobacteria, responsible for serious respiratory infections, have emerged across the globe and are posing a serious challenge to the healthcare infrastructure of numerous countries (Lange 2018). *Mycobacterium tuberculosis* and *Mycobacterium abscessus* are two such pathogens which have garnered resistance to antibiotics like rifampicin and isoniazid, of which they were previously sensitive to (Lange 2018). Consequently, antimicrobial research has been focused on devising new therapeutic strategies to combat this rising public health threat, including the testing of new antibiotic prescription strategies and investigation into novel treatments options other than traditional antibiotics (Lange 2018).

Mycobacteriophage Genome Diversity and Novelty: A Potential Solution

Mycobacteriophages, called phages, are ubiquitous viruses that selectively infect mycobacterial hosts. In the course of an infection, mycobacteriophages, like all phages, can bind to specific hosts and insert their genomes into the cell (Hatfull 2022). By hijacking the cellular replication, transcription, and translation machinery in the host, phages can replicate their genomes and synthesize the various proteins necessary to assemble new virions (Hatfull 2022). Once assembled, phages can undergo two different lifestyles (Hatfull 2022). Lytic phages can

lyse the host immediately, releasing new virions to infect other nearby bacteria (Hatfull 2022). In contrast, temperate phages can integrate their genomes with the hosts, replicating along with it until environmental factors trigger the phage to undergo lytic infection (Hatfull 2022). Either way, various structural and enzymatic proteins, encoded by the phage, must come together to ensure a successful infection (Hatfull 2022). Concurrently, phages have evolved alongside their hosts allowing mycobacteria to evolve various strategies to circumvent or prevent phage infection, such as efflux pumps and the CRISPR-Cas9 system (Hatfull 2022). In response, phages have devised mechanisms to overcome these bacterial defenses, contributing to phage diversity (Hatfull 2022). In particular, this coevolution and diversity in infection lifestyle has resulted in immense diversity with respect to the various genes encoded within a particular phage's genome, all in the pursuit of maximizing efficiency of a phage infection (Hatfull 2022). Continued diversification of phages has resulted in the generation of various clusters and subclusters, groups of closely related phages organized based on their nucleotide similarity, often determined through bioinformatic analysis and genome annotation (Hatfull 2022). Since phages occupy a large swath of the planet, numbering close to 10^{31} viral particles globally, very few phage genomes have been completely, bioinformatically annotated, let alone functionally characterized using wet-bench molecular techniques (Hatfull 2022). By gaining a better understanding of mycobacteriophage biology though genome characterization, novel strategies may be discovered which can be exploited in the struggle against multidrug resistant mycobacterial pathogens (Allué-Guardia 2021).

Characterization of Cytotoxic Genes 41 and 64 in Mycobacteriophage Phayonce

Phayonce is a temperate mycobacteriophage belonging to subcluster P5 (Pope 2015). Phayonce's genome is 49,203 bp long and encodes seventy-seven genes (Pope 2015). Some of

the genes have been previously annotated with predicted functions via bioinformatic analysis (Pope 2015). However, a majority of Phayonce's genes lack such characterization, instead, being annotated as having no known function (Pope 2015). Previous work, similar to that conducted in the study of the phage Waterfoul, concerning the functional characterization of Phayonce's genome has identified numerous genes, both with predicted and no known function annotations, exhibiting cytotoxicity when overexpressed in Phayonce's host, *Mycobacterium smegmatis* (Heller 2022). The aim of this study will be to identify the punitive phage-host protein interactions for two non-known function cytotoxic genes, Phayonce 41 and 64, using a bacterial two-hybrid screen developed by the SEA-GENES program of the Howard Hughes Medical Institute.

Methods

Generation and Cloning of p2Hα Constructs Containing Phayonce 41 and 64

To generate inducible expression vectors for the bacterial two-hybrid screen, Phayonce genes 41 and 64 were PCR amplified using pExTra_universal_F and pExTra_universal_F primers, Q5 polymerase, and pExTra constructs containing each gene of interest as the template DNA (Table 1). The pExTra_universal primers add 108 bp to the initial gene fragments. Resulting amplicons were resolved on a 1 % agarose gel to verify successful amplification of gene inserts (Table 1). Amplified gene inserts were then digested with NdeI and SbfI or BamHI restriction enzymes. Also, the p2Ha plasmid backbone was digested with NdeI and both SbfI or BamHI restriction enzymes and prepared with a Zymo DNA Clean and Concentrate Kit. Both the purified Phayonce 41 and 64 inserts were ligated into the purified and digested p2Hα backbones. Once the constructs were prepared, both were chemically transformed into chemically competent *E.coli* cells and incubated on LB agar supplemented with spectinomycin at 50 μg/ml. To verify

that transformants contained the constructs, two replicates of transformants underwent colony PCR for each gene using o-p2HαF & o-p2HαR primers and *Taq* polymerase (Table 2). The o-p2Hα primers add 269 bp to the initial gene fragments. Amplified inserts were resolved on an 1% agarose gel via electrophoresis to confirm successful transformation of each gene construct (Table 2). Transformed colonies were inoculated in LB liquid media supplemented with spectinomycin at 50 μg/ml. Using a Zymo plasmid mini-prep kit, p2Hα constructs were extracted and purified from inoculated cultures for downstream use in the bacterial two-hybrid screen.

Table 1. PCR protocol for initial amplification of Phayonce 41 and 64 gene inserts.

Step	Temperature (C)	Duration
1. Initial Denaturation	98	5 min
2. Denaturation	98	10 sec
3. Annealing	65	10 sec
4. Extension	72	25 sec
Repeat Steps 2-4 for 29 cycles		
5. Final Extension	72	5 min
6. Hold	4	Infinite

Table 2. Colony PCR Verification Protocol for Phayonce 41 and 64.

Step	Temperature (C)	Duration
1. Initial Denaturation	95	2 min
2. Denaturation	95	30 sec
3. Annealing	58	30 sec
4. Extension	72	35 sec
Repeat Steps 2-4 for 29 cycles		
5. Hold	12	Infinite

B2H Selection Screen

For both genes 41 and 64, purified p2Hα constructs were transformed into B2H-SELECT/pCl-SMEG Library electrocompetent cells via electroporation at 1.8 kV. Transformants were allowed to recover for two hours at 37 °C with two µl of IPTG (10 mM) being added halfway through the incubation period to induce expression of the p2Hα constructs and pCl library plasmids. For gene 41, recovered transformants were then plated, at 25 µl, 50 µl, and 100 µl amounts, on three selection LB agar plates containing carbenicillin (1500 µg/ml), kanamycin (50 µg/ml), chloramphenicol (25 µg/ml), spectinomycin (50 µg/ml), X-gal (40 µg/ml), and IPTG (20 µM). For gene 64, recovered transformants were plated on selection plates at 25 µl, 50 µl, and 100 µl followed by a subsequent plating using the same amounts but at a 4X concentration, for six total selection plates. The supplements in the selection plates allow for the selection of transformants containing the p2Hα and pCI vectors and for the selections of putative interactions expressing the reporter genes *bla* and *lacZ* present in the B2H library. For both genes, the remaining transformants recovered underwent a 10^{-7} serial dilution with 100 µl of the 10^{-3} and 10^{-4} dilution being plated on LB agar growth plate containing all the supplements used selection plates with carbenicillin excluded. All plates were incubated for 48 hours followed by a 2-hour refrigeration period to accentuate the blue colony color of isolation punitive interactions. The resulting colonies on growth plates were used to quantify transformation efficacy. The selection plates, with isolated putative interactions, were then subject to candidate verification.

Two-Hybrid Candidate Verification

To ensure a punitive interaction isolated from the two-hybrid screen is not a false positive, LB liquid cultures, supplemented with kanamycin (50 μg/ml), chloramphenicol (25 μg/ml) , and spectinomycin (50 μg/ml) were inoculated with select punitive colonies exhibiting reporter expression. Mixed plasmid preps, containing both the p2Hα and pCI expression vectors, were prepared for each selected candidate using the Zymo plasmid mini prep kit. Each candidate mixed prep was then chemically transformed into chemically competent B2H SELECT *E.coli* cells and plated on media containing kanamycin (50 μg/ml), chloramphenicol (25 μg/ml), and spectinomycin (50 μg/ml). Transformants were incubated at 30 °C for 24 hours.

For verification of punitive interactions, two colonies of each candidate transformant were propagated at 37 °C for an hour and were subjected to a 10^{-5} serial dilution in LB liquid culture media supplemented with kanamycin, chloramphenicol, spectinomycin, and IPTG. Candidate dilutions were then spotted on both growth and selection plates, prepared as described above, along with both a positive and negative control. Verification plates were incubated at 30 °C for 48 hours followed by a 2-hour refrigeration period, to reisolate punitive phage-host protein interactions.

Sequencing of Verified Interactions

For verified interactions, 10 μl of the corresponding mixed plasmid prep and 5 μl of a 5 pM pCI forward primer stock were used to prepare Sanger sequencing reaction mixes, which would be sent to a third-party sequencing service. The first 10 reactions sent for sequencing were prepared with undiluted pCI primer mix, yet still returned interpretable results for analysis. Subsequent chromatograms were subject to an NCBI blast search, within the *M. smegmatis* genome, to identify the host protein fragment, and corresponding ORF, present in the candidate mixed plasmid prep.

Results

Generation of Two-Hybrid Expression Vectors for Phayonce 41 and 64

For preparation of the p2Hα expression vectors for Phayonce 41 and 64, individual gene inserts were successfully amplified from Phayonce's genome utilizing PCR (Figure 1). The expected size of amplified bands is 387 bp and 255 bp for genes 41 and 64 respectively, since the pExTra_universal primers add 108 bp to the initial gene fragments at 279 bp for 41 and 147 bp for 64 (Figure 1). Once amplified, the inserts were ligated into the p2Hα expression vectors and transformed into the B2H-SELECT SMEG/pCI *E.coli* library cells. Using the o-p2HαF & o-p2HαR primers, the gene inserts were then reamplified successfully verifying successful ligation of vectors (Figure 2). Expected size of amplified bands are 548 bp and 416 bp for genes 41 and 64 respectively, since the o-p2hα primers add 269 bp to the initial gene fragments at 279 bp for 41 and 147 bp for 64 (Figure 2).

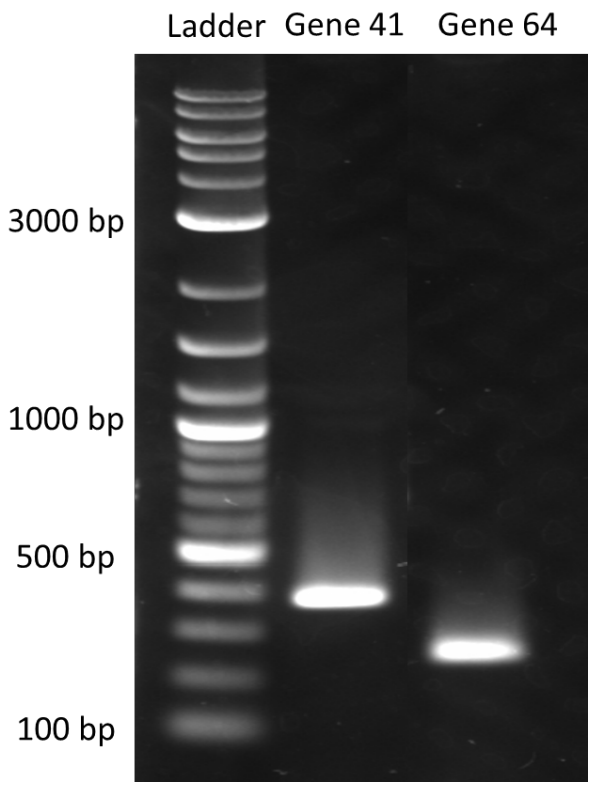

Figure 1. 1% Agarose gel of amplified gene inserts for Phayonce 41 and Phayonce 64. 1kb Plus DNA ladder was used. There are two bright bands measuring about 400 bp and 250 bp, respectively. The expected size of amplified genes is 387 bp for gene 41 and 255 bp for gene 64.

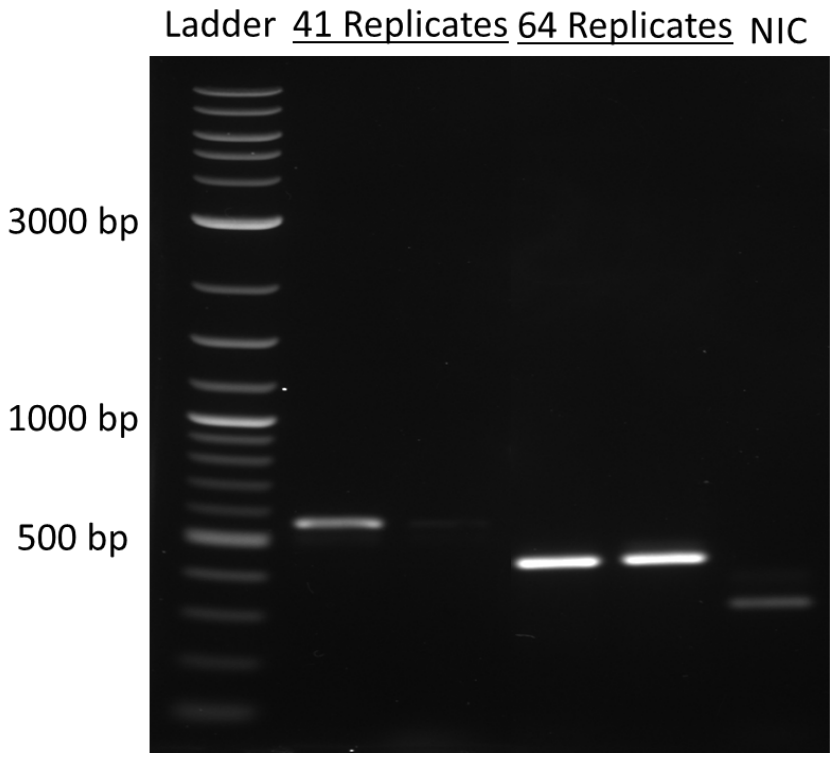

Figure 2. Colony PCR amplicons resolved on a 1% agarose gel along with a 1 kb Plus DNA ladder. For Phayonce 41, one of the replicates was amplified, measuring around 550 bp. For Phayonce 64, both replicates were successfully amplified with two, bright bands at around 450 bp visible. Expected sizes for gene 41 replicates is 548 bp and 416 bp for gene 64 replicates. A no insert control was included with a faint band appearing at around 300 bp.

Total Two-Hybrid Screen Results

Phayonce 41 hybrid transformants were plated on three selection plates of increasing volumes of 25, 50, and 100 µl (Figure 3). The greatest number of punitive interactions, 28, were isolated on the 100 µl selection plate with the least amount of punitive interactions, 8, were isolated on the 25 µl selection plate (Figure 3). Nine punitive interactions were isolated in the 50 µl selection plate (Figure 3). Since the total number of transformants plated failed to reach the 5 $x10^5$ to 5 x 10^6 CFU plated optimal range, indicating a near complete screen of the B2H library,

the Phayonce 64 screen utilized six selection plates with increased concentrations of transformants plated (Figure 3). For Phayonce 41, a total of 1.1 x 10^5 transformants of B2H library *E.coli* cells containing the p2Hα-41 expression vector were plated on selective media, specific to reporter gene expression (Table 3). A total of 45 colonies were isolated with 34 exhibiting reporter gene expression for both *bla* and *lacZ* genes (Table 3). Of colonies exhibiting expression for both reporter genes, 22 were verified and sequenced (Table 3).

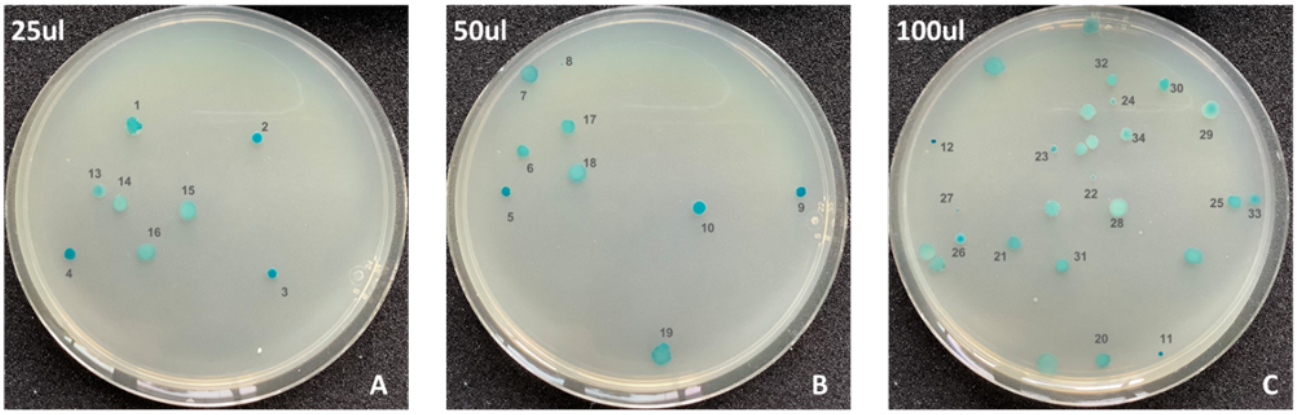

Figure 3. Two-Hybrid Screen for Phayonce 64. **A**. Out of 1.6 x 10^4 transformants plated, 8 colonies were isolated. **B.** Out of 3.2 x 10^4 transformants plated, 9 colonies were isolated. **C**. Out of 6.4 x 10^4 transformants plated, 28 colonies were isolated. Numbers assigned to colonies indicate the order from which they were selected for downstream verification.

Phayonce 64 hybrid transformants were plated following the same scheme outlined for 41 with the addition of three more selection plates containing 4X concentrations of cells. Despite the increase in the number of transformants plated, the optimal range of transformants screened was not achieved (Figure 4). However, more colonies were isolated compared to 41 with the greatest number of colonies isolated numbering 61 on a single plate (Figure 4). For Phayonce 64, a total of 4.8 x 10^5 transformants of B2H library *E.coli* cells containing the p2Hα-64 expression

vector were plated on selective media, specific to reporter gene expression (Table 3). A total of 125 colonies were isolated with 43 exhibiting reporter gene expression for both *bla* and *lacZ* genes (Table 3). Of colonies exhibiting expression for both reporter genes, 18 were verified and sequenced (Table 3).

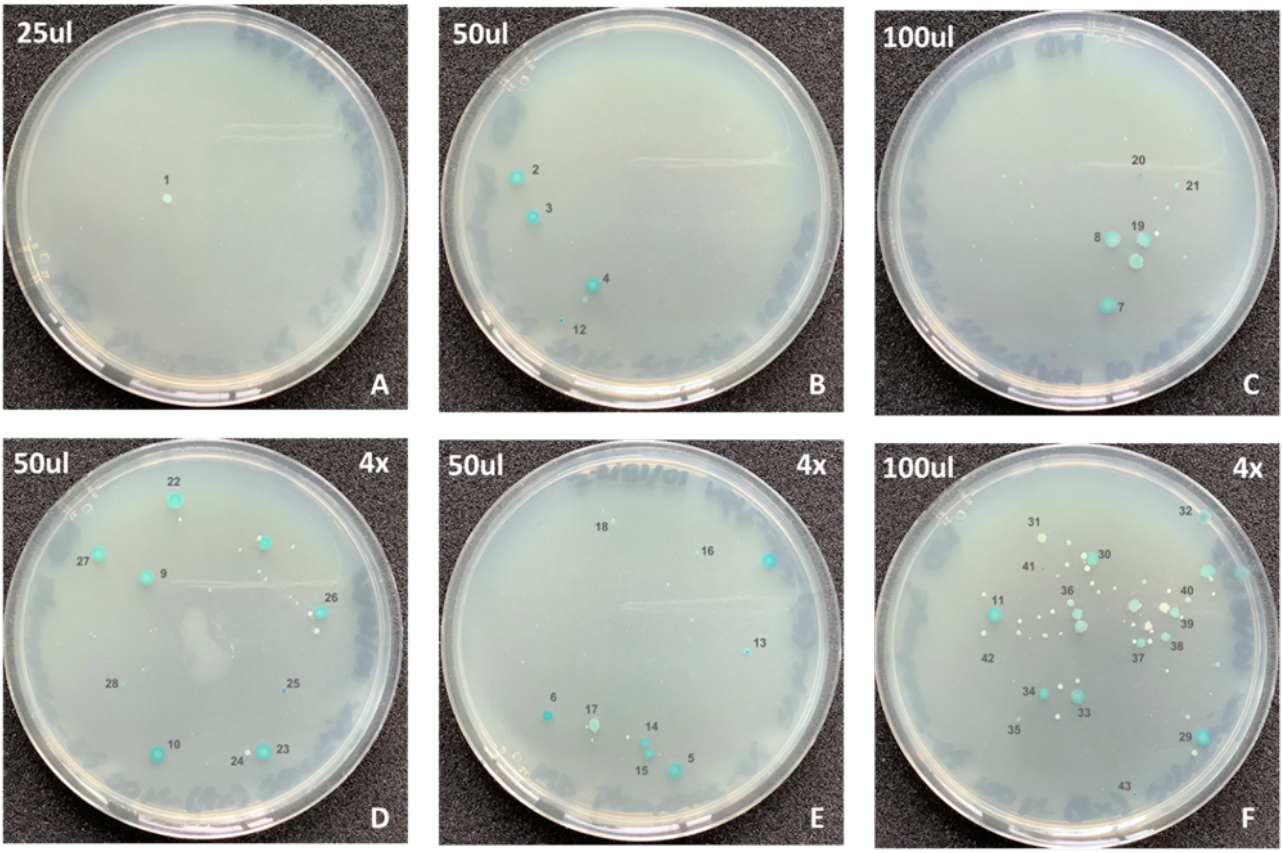

Figure 4. Two-Hybrid Screen for Phayonce 64. **A**. Out of 1.25 x 10^4 transformants plated, only one colony was isolated. **B**. Out of 2.5 x 10^4 transformants plated, 5 colonies were isolated. **C**. Out of 5 x 10^4 transformants plated, 15 colonies were isolated. **D**. Out of 1 x 10^5 transformants plated, 25 colonies were isolated. **E**. Out of 1 x10^5 transformants plated, 18 colonies were isolated. **F**. Out of 2 x10^5 transformants plated, 61 colonies were isolated. Numbers assigned to colonies indicate the order from which they were selected for downstream verification.

Table 3. Cumulative Two-Hybrid Screen and Verification Results for Phayonce 41 and 64.

Gene	Transformants Plated (CFU)	Colonies Isolated	Selected for Verification	Verified and Sequenced
41	1.1×10^5	45	34	22
64	4.8×10^5	125	43	18

Total Sequencing Hits

Following analysis of returned and interpretable sequencing chromatograms, nine separate host gene targets were identified for Phayonce 41 (Table 4). Two gene targets were annotated as having no known function (Table 4). Three gene targets for Phayonce 41, ORFs 1252, 3172, and 5523, contained multiple, independent hits corresponding to different segments of the gene, while the rest contained either one hit or multiple, identical hits (Table 4). ORF 5523, annotated for a prolyl oligopeptidase protein, contained a total of two copies of two overlapping, independent hits (Table 4). Comparatively, ORF 1252 had the same number of total hits, with three being unique, yet was annotated as having no known function (Table 4). Interestingly, fifteen identical hits were returned for a single, no known function gene, 1545, within the *M. smegmatis* genome (Table 4). The hits for this particular gene encompassed the entire length of the open reading frame for the gene.

Table 4. Sequencing hits within the *M. smegmatis* genome obtained from Phayonce 41 and 64

	M.smegmatis ORF	Total Screen Hits	Independent Hits	Protein Encoded
Phayonce 41	1252	4	3	NKF
	4832	1	1	acyl-CoA-dehydrogenase
	2492	1	1	D-lactate dehydrogenase
	3172	2	2	DNA polymerase IV 1
	4682	3	1	Na+/H+ Antiporter
	5523	4	2	prolyl oligopeptidase
	153	1	1	2-dehydropantoate 2-reductase
	6095	1	1	D-amino deaminase
	3783	1	1	acyl-CoA-dehydrogenase
	1247	1	1	NKF
Phayonce 64	1545	15	1	NKF

Discussion

Bacterial Two-Hybrid Efficiency

The bacterial two-hybrid screening system, developed by SEA-GENES, isolated numerous punitive phage-host protein interactions between select Phayonce gene protein products and multiple *M. smegmatis* proteins (Table 4). Ideally, the screen was optimized to examine between 5 x10^5 to 5 x 10^6 transformants, to identify all potential protein interactions possible within the B2H *E.coli* host library. However, in either cumulative screen for both Phayonce 41 and 64, the minimum threshold of 5 x 10^5 transformants screened was not achieved with only the screen for Phayonce 64 coming the closest at around 4.8 x 10^5 total transformants screened (Table 3). The electrocompetent cells used for the transformations had undergone several cycles of thawing and refreezing in the freezer in which they were stored. This fluctuation in temperature may have damaged the integrity of the cells prior to electroporation resulting in an overall low quality of transformation efficiency and total amount of transformants

available for screening. Consequently, there is the possibility that potential phage-host interactions were overlooked due to the failure in being able to confidently screen through the entire host library. Despite this, however, the two screens provided enough verifiable interactions leading to multiple, independent hits of host genes, and their encoded proteins, capable of interacting with either Phayonce 41 or 64 (Table 4). The amount of potential host interacting partners provides a foundation from which further molecular and biochemical investigation can be conducted to characterize the relationship between the phage and host proteins.

Phayonce 41 and Host Peptidase Interaction: A Possible Mechanism of Cytotoxicity

One of the potential interacting partners identified for Phayonce 41 includes a host prolyl oligopeptidase protein, which functions by enzymatically cleaving the bonds between small peptides (Table 4). Potentially, the interaction between the Phayonce 41 protein product and the host peptidase may provide an explanation for why the overexpression of Phayonce 41 is deleterious to host growth. The crystal structure of the prolyl oligopeptidase includes a beta-propeller domain covering a catalytic triad, responsible for the cleavage of peptide bonds (Polgár 2002). The beta-propeller motif creates a pore which allows for small peptides to enter and reach the catalytic triad, while preventing larger peptides from entering (Polgár 2002). Alteration of this function might result in the poor regulation of peptide levels within the host. Depending on whether the interaction of Phayonce 41 results in a gain of function or inhibition of the prolyl oligopeptidase, multiple models describing how this interaction can result in host toxicity can be explored. One of the key avenues moving forward will be to determine the essentiality of the host peptidase, through the generation of knockout host mutant. Additionally, a co-immunoprecipitation experiment to isolate the two proteins will provide an additional verification of the interaction.

Conclusion

In all, the bacterial-two hybrid screen has isolated and identified multiple punitive host interacting partners for two cytotoxic phage genes, Phayonce 41 and 64 (Table 4). The number of possible interacting partners identified provide a groundwork for further investigation of phage-host interactions between Phayonce and *M. smegmatis* to be investigated. By characterizing the relationship between phage and host proteins, as in the Phayonce 41 and prolyl oligopeptidase relationship, potential cytotoxic genes and susceptible host proteins targets can be identified for use in novel antibiotic therapeutics (Heller 2022). Therefore, the dynamics between mycobacteriophage and mycobacteria can potentially be exploited in therapeutics against multidrug resistant pathogens like *M. tuberculosis* and *M. abscessus*.

Acknowledgements

This research is supported by Howard Hughes Medical Institute's SEA-PHAGES program. This project was supported by grant P20GM103499-20 (SC INBRE) from the National Institute of General Medical Sciences, National Institutes of Health

This research was also made possible by the contributions of Dr. Williams, Kyla Thomas, Jaqueline Gould, Emma Hofseth, and all past SEA-PHAGES and SEA-GENES researchers.

References

Allué-Guardia, A., Saranathan, R., Chan, J., & Torrelles, J. B. (2021). Mycobacteriophages as potential therapeutic agents against drug-resistant tuberculosis. *International Journal of Molecular Sciences*, 22(2), 735. https://doi.org/10.3390/ijms22020735

Hatfull, G. F. (2022). Mycobacteriophages: From petri dish to patient. *PLOS Pathogens*, 18(7), e1010602. https://doi.org/10.1371/journal.ppat.1010602

Heller, D., Amaya, I., Mohamed, A., Ali, I., Mavrodi, D., Deighan, P., & Sivanathan, V. (2022). Systematic overexpression of genes encoded by mycobacteriophage Waterfoul reveals novel inhibitors of mycobacterial growth. *G3 Genes Genomes Genetics*, 12(8). https://doi.org/10.1093/g3journal/jkac140

Lange, C., Chesov, D., Heyckendorf, J., Leung, C. C., Udwadia, Z., & Dheda, K. (2018). Drug-resistant tuberculosis: An update on disease burden, diagnosis, and treatment: Drug-resistant tuberculosis. *Respirology*, 23(7), 656–673. https://doi.org/10.1111/resp.13304

Polgár, L. (2002). The prolyl oligopeptidase family. *Cellular and Molecular Life Sciences CMLS*, 59(2), 349–362. https://doi.org/10.1007/s00018-002-8427-5

Pope, W. H., Jacobetz, E., Johnson, C. A., Kihle, B. L., Sobeski, M. A., Werner, M. B., Adkins, N. L., Kramer, Z. J., Montgomery, M. T., Grubb, S. R., Warner, M. H., Bowman, C. A., Russell, D. A., & Hatfull, G. F. (2015). Genome sequence of mycobacteriophage Phayonce. *Genome Announcements*, 3(3), e00598-15. https://doi.org/10.1128/genomeA.00598-15