Tape measure protein multiplex PCR successfully classifies Arthrobacteriophages

Stu Mair and Tamarah Adair

Department of Biology, Baylor University, Waco, Texas 76798

Introduction

Currently, when a new bacteriophage is discovered, full-genome analysis is used to sort it into a cluster, which represent groupings of genetically-similar phages. However, this method of clustering has a few drawbacks:

- It requires that the full genome be sequenced, which takes time and resources
- The DNA extraction for sequencing requires a high-titer lysate, which takes time in the lab and can be difficult to obtain for certain phages

In a 2013 article published in *BMC Genomics*, Smith et al. published an article detailing the use of a single gene, the Tape Measure Protein (TMP), to cluster Mycobacteriophages – a phage type associated with the host *Mycobacterium smegmatis* being used in the HHMI SEA-Phages program. The host range associated with the program has extended to include several other Actinobacteria hosts, with the use of *Arthrobacter sp.* 21022 extending to several university programs. At the time of publication, there were 552 phages isolated on the host, with 191 sequenced phages. Despite these numbers, there are some noticeable gaps in data among the 12 clusters, with just two phages in the AP cluster and four in the AT cluster.

As a way to increase the number of phages clustered and increase the ease with which researchers can classify their discovered phages, analysis of TMP sequences with a Polymerase Chain Reaction is being explored.

Objectives

TMP Single Gene Analysis

Use alignments of Tape Measure Protein sequences to differentiate and sort phages into clusters identical to those identified with full genomes

PCR Primer Design

Design PCR primer sets for each cluster using conserved regions of the Tape Measure Proteins that amplify to identifiable product sizes

PCR Testing and Gel Electrophoresis

Utilize PCR reactions and gel electrophoresis to verify the design of the primers and ability to cluster bacteriophages in vitro

TMP Single Gene Analysis

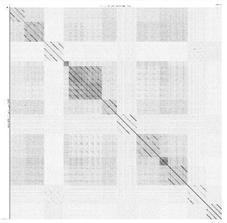


Figure 1a. Full Genome Gepard Dotplot

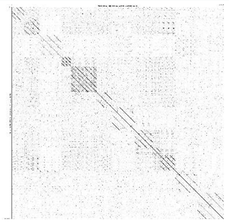


Figure 1b. TMP Sequence Gepard Dotplot

Figure 1. These Gepard dotplots offer a comparison of clustering for bacteriophages isolated on *Arthrobacter sp.* ATCC 21022 by whole genomes and by tape measure proteins (TMP). By combining the genome sequences of 3 bacteriophages from each cluster and plotting them against themselves, Figure 1a demonstrates the similarities evident in the genomes. As can be seen by the boxes along a diagonal, there are clear regions of high similarity. Figure 1b is the same process done with the TMP sequences. Though weaker, the clusters still appear, confirming the ability to cluster based on the TMP sequences.

PCR Primer Design

Methods: The primers were designed, in silico, by finding regions of high similarity on a Mega7 alignment of the TMP sequences and using the NCBI primer design tools to select for product sizes and appropriate melting temperatures for each cluster.

Species/Abbrv				*	*	*	*				*
1. Shrooms_AL	C	G	G	C	T	T	C	A	C	C	G
2. Laroye_AL	-	-	-	C	T	T	C	C	A	A	G
3. Salgado_AL	-	-	-	C	T	T	C	C	A	A	G

Figure 2. This is an example of the TMP sequence alignments. The asterisks represent full conservation, the blank headers represent mismatches and the dashes in the sequences represent gaps.

In designing a pair of primers, two sizable (20+ base pairs) regions were scanned for perfect and specific conservation among the cluster. The product sizes were then compared to ensure the differences could be resolved by gel electrophoresis. This process enabled consistent primer binding resulting in a product size that would enable cluster assignments (Table 1).

Cluster	Forward Primer	Reverse Primer	Product Size (bp)
AK	AACGCGGTATTCGGGTTCTTCT	GCACCTACAATCTTTCCTGCCG	590
AL	AAGGACTACACGGGCCTGAC	CCTCAAAGGTCTTCGTGGACTT	334
AM	GTTATGGCCGAAGAGGCTGCA	CCATCTTAGCGATGCCTTCGTCA	742
AN	GCCGTTTTGGCCGTTCGTAT	TTAGCTCGTTCCTGGGCACC	671
AO	TCGGCTCGAAGCTGAAGGGC	GGTTCGAGCGCGTCTTTCGTC	487
AP	ATCGCGGCCTTGAAGGACAT	GGCGAACTCGGTGAACATGC	201
AQ	CGGCTCTCGGCCAGATTCTC	AATCTTGCTGGACGCACCCT	899
AR	TTGACGGCTGGGGCATGCT	AACAGGGCCACGGCAATGGT	142
AS	GACCGTGGCATTCCGATTGTCC	GCGACCTGCACGACGAAAGC	827
AT	GAACAACAGCTTCAAGCGCACC	GCCAGTTCGGTCTTGGCGTC	276
AU	GTCTTCAAGCTCCTGAAGCGCA	TGTTGTAGAGCTTGCGGCCCT	418
AV	GGCTGTGGGAACTGGATTGCA	ACAACGCCGTCACCCTTACC	963

Table 1. These are the finalized primers for each cluster and the product size they are designed to produce when analyzed by gel electrophoresis analysis. All melting temperatures were held within 5°C of each other.

PCR Testing and Gel Electrophoresis

Methods: PCR Reactions were completed on known phage samples (either purified DNA or boiled lysate) from each cluster to verify that the primer sets properly bind to their template strands.

To generate the boiled lysate samples, a protocol from the SEA-PHAGES program was followed, boiling the lysate in phage buffer at 100°C. This yielded a sample that could be used for PCR reactions.

A standard PCR protocol was used - denaturation at 94°C; 30-35 cycles of denaturing at 94°C, annealing at 55.1°C, and extension at 72°C, followed by a final extension at 72°C.

The PCR reaction products were analyzed through gel electrophoresis on a 2% agarose gel.

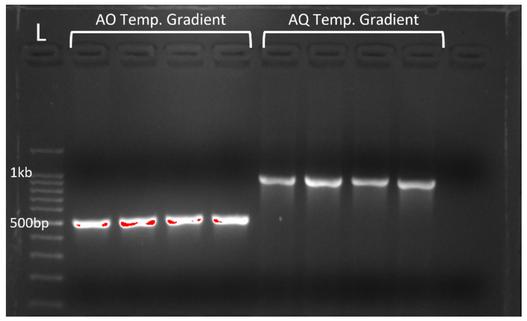


Figure 3. This gel electrophoresis both confirms the accuracy of two primer sets (AO and AQ) and demonstrates the gradient that the primers were tested across initially to find an optimal annealing temperature. The four temperatures used were 52.4°C, 53.7°C, 55.1°C, and 57.5°C. The most consistent results across all clusters were at 55.1°C, which was used as the annealing temperature for all PCR reactions without a gradient. The expected result for AO is 487bp and for AQ is 899bp.

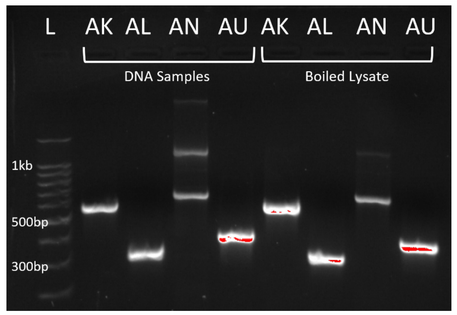


Figure 4. This PCR reaction verified that four of the primer sets (AK (590bp), AL (334 bp), AN (671 bp), and AU (418 bp)) functioned properly and also confirmed the ability to use boiled lysate as a source for the PCR template strand as an alternative to utilizing purified DNA extracted from a high-titer lysate.

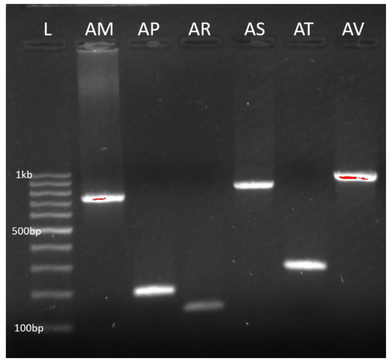
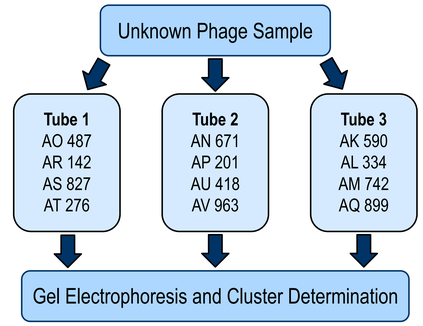


Figure 5. This PCR reaction was used to verify six of the primer sets (AM (742 bp), AP (201 bp), AR (142 bp) , AS (827 bp), AT (276 bp), and AV (963 bp)) and utilized boiled lysate for its template strands.

Multiplex PCR Reactions



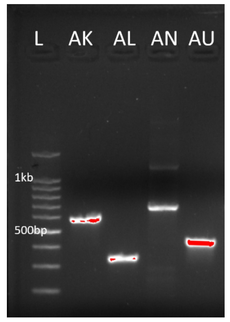


Figure 6. Each tube of this PCR reaction contained the primer sets for four clusters (AK, AL, AN, and AU) but only one of the DNA templates. As can be seen on the gel electrophoresis image, a multiplex PCR is an effective means by which to classify a phage sample. By grouping the primer sets by product size variations and extending the multiplex to three tubes, as shown in the experimental flowchart above, unknown samples from any known cluster can be identified.

Conclusions

Following the design of PCR primers corresponding to the existing clusters for bacteriophages isolated on *Arthrobacter sp.* ATCC 21022, the actual PCR reactions confirm that they can be used for single-primer screening and multiplex classification. The formulation of the gel could be adjusted in the future to get better product resolution between the larger products.

Moving forward, these primers can be used in a number of different ways. They can aid researchers in classifying low-titer lysates and screening soil extracts. In a related ecology method, researchers can use multiplex reactions to determine the arthrobacteriophage clusters present in different soil types and locations.

Finally, this research method, which will be published for public use, also stands to benefit students researching in the SEA-PHAGES program by introducing them to an additional genetic characterization technique.

Acknowledgements

My gratitude goes out to Tamarah Adair for her mentorship and effort that she has poured into helping me make this research project successful.

I would also like to thank Natalie Widdows and Niharika Koka for their valuable input during the initial planning for this project.

References

Smith, K. C., Castro-Nallar, E., Fisher, J. N., Breakwell, D. P., Grose, J. H., & Burnett, S. H. (2013). Phage cluster relationships identified through single gene analysis. *BMC Genomics*, *14*, 410.

SEA-PHAGES. Alternative to Purifying Bacteriophage DNA for PCR. *SEA-PHAGES Forums.*