Increasing the diversity of the mycobacteriophages you isolate

Though the collection of sequenced mycobacteriophages now numbers more than 1000, we are still occasionally finding new Singleton phages and have many clusters with very few members. How can you increase your chances of discovering these less-common types of phages?

Two-Pronged Approach

This topic was discussed at a session at the 2016 SEA-PHAGES Faculty Retreat and Workshop, and the approaches to discovering uncommon types of mycobacteriophages fell into two broad categories.

1. *Protocol tweaks to isolate uncommon phages from environmental samples.*
2. *Identifying which phages you've isolated are underrepresented and pursuing those, or excluding oversampled phages from further purification and/or sequencing.*

Background

First, the collection of mycobacteriophages as of June 2016 reveals phage clusters unevenly represented.

Table 1 – Numbers and percentages of sequenced mycobacteriophage genomes per cluster.

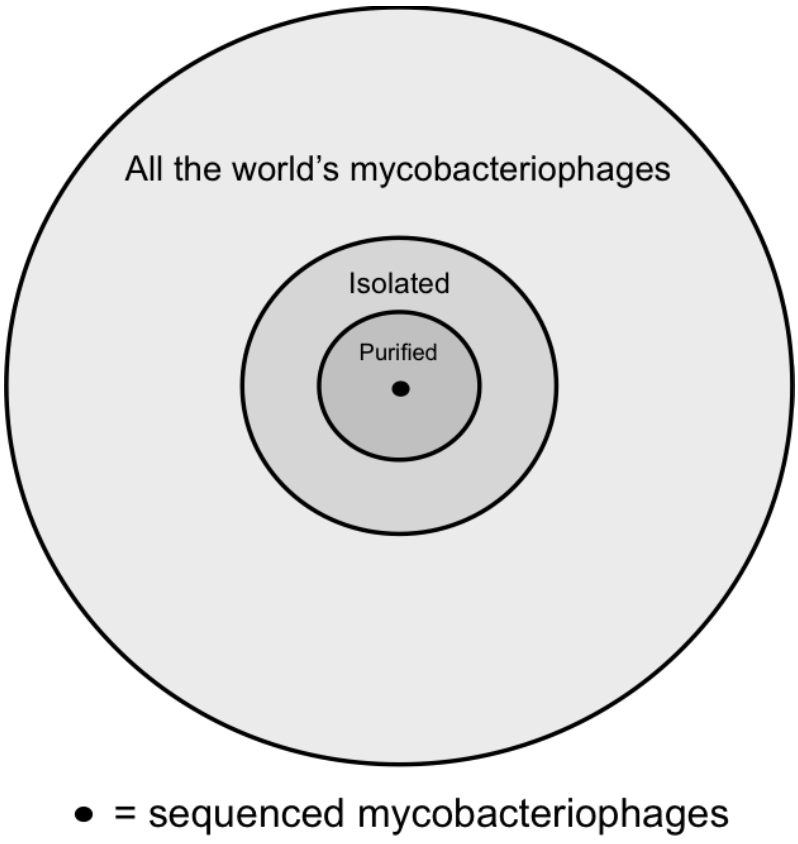
Cluster	# Phages	% of Mycos	Cluster	# Phages	% of Mycos
A	432	37.5	O	7	0.6
B	193	16.7	P	20	1.7
C	79	6.9	Q	5	0.4
D	11	1.0	R	5	0.4
E	75	6.5	S	3	0.3
F	109	9.5	T	4	0.3
G	33	2.9	U	2	0.2
H	5	0.4	V	3	0.3
I	5	0.4	W	3	0.3
J	23	2.0	X	2	0.2
K	75	6.5	Y	2	0.2
L	26	2.3	Z	2	0.2
M	9	0.8	Singletons	6	0.5
N	14	1.2			

Why does our current collection have this distribution? Does is accurately reflect the distribution of all mycobacteriophages found in nature or is it a biased collection? Perhaps it is a reflection of both.

Below is a diagram relating "all the world's" mycobacteriophages and the reduced numbers and diversity expected through 3 major stages ending with the selection of a phage(s) for genomic sequencing. The stages are listed below.

1. *Initial extraction and isolation of phage/plaques*
2. *Selection of a phage/plaque for purification*
3. *Selection of a phage genome(s) for sequencing*

Since the decision(s) made at each step determines the diversity of mycobacteriophages available for the next stage of work, it is helpful to have the mycobacteriophage diversity be as large as possible at each step along the way.



What follows is a list of modifications available at each step, and where relevant, short notes of their use by some institutions.

Modifications to Standard Procedure

Step 1: Initial Extraction and Isolation of Phages

Multiple factors can influence the size and diversity of phages extracted from your environmental samples.

a. Host organism
    - *M. smegmatis* MC^2155 is the standard host strain (BSL-1)
    - Consider other *Mycobacterium* species or *M. smegmatis* strains for phage isolation
        - If BSL-2 host is used but isolated phage also propagates on *M. smegmatis*, you can switch to the safer BSL-1 host for subsequent purification work
b. Soil sample – type, quality, geography
    - Mycobacteriophages are most easily isolated from fresh collections of rich compost soil; other types of environmental samples have lower success rates.
    - Use geographical information of previous sample sites (especially those that did not yield phages) to guide new site selection choice
    - Avoid professionally landscaped areas (e.g., mulch bedding sites throughout college campus may have similar mycobacteriophage populations)
    - Consider plant diversity found at environmental sample sites (iNaturalist app)
    - Consider previous soil collection data – temperature, depth, moisture content, etc.
c. Phage isolation method – Enrichment vs. direct plating
    - Direct plating is the original and preferred approach to phage isolation and, in theory, it provides a better chance of recovering "rare" phages; enrichment may select for and increase the proportional numbers of faster growing phages over more slowly growing phages
    - Below is a table showing the percentage of sequenced phage known to be isolated via enrichment or by direct plating. There are an additional nearly 300 sequenced phages for which the isolation method was not recorded. Most phage clusters have too few members to make any reliable statements regarding possible influence of the method of isolation. However, multiple clusters (A, B, C, J, K, Singletons; highlighted in bold lettering) may show biases based on the phage isolation method. Phages from cluster B and J appear more readily isolated from direct plating whereas the others show higher isolation numbers from enrichment, with the clusters J and K phages showing perhaps the most extreme degree of isolation method bias. Note that other variables (e.g., enrichment incubation temperature and time) may also differ.

Table 2 – Numbers and percentage of sequenced mycobacteriophages by isolation method.

	Enrichment		Direct plating		Total known	
Cluster	**#**	**%**	**#**	**%**	**#**	**%**
A	**286**	**40.68**	**43**	**26.88**	**329**	**38.12**
B	**99**	**14.08**	**39**	**24.38**	**138**	**15.99**
C	**39**	**5.55**	**15**	**9.38**	**54**	**6.26**
D	1	0.14	2	1.25	3	0.35
E	50	7.11	11	6.88	61	7.07
F	65	9.25	12	7.50	77	8.92
G	20	2.84	5	3.13	25	2.90
H	2	0.28	0	0.00	2	0.23
I	3	0.43	0	0.00	3	0.35
J	**5**	**0.71**	**12**	**7.50**	**17**	**1.97**
K	**62**	**8.82**	**1**	**0.63**	**63**	**7.30**
L	17	2.42	5	3.13	22	2.55
M	5	0.71	2	1.25	7	0.81
N	12	1.71	1	0.63	13	1.5
O	3	0.43	3	1.88	6	0.70
P	13	1.85	3	1.88	16	1.85
Q	3	0.43	0	0.00	3	0.35
R	2	0.28	2	1.25	4	0.46
S	0	0.00	3	1.88	3	0.35
T	3	0.43	1	0.63	4	0.46
U	1	0.14	0	0.00	1	0.12
V	1	0.14	0	0.00	1	0.12
W	1	0.14	0	0.00	1	0.12
X	2	0.28	0	0.00	2	0.23
Y	0	0.00	0	0.00	0	0.00
Z	1	0.14	0	0.00	1	0.12
Singleton	**7**	**1.00**	**0**	**0.00**	**7**	**0.81**
Total =	703	100.00	160	100.00	863	100.00

d. Enrichment conditions
   - Extraction, enrichment, and plate incubation times and temperature can all be modified and may influence the numbers and diversity of phages recovered.
     - Gonzaga and Hope College (and perhaps others) have run enrichment cultures incubated at ≤ 32°C and have recovered a higher proportion of Cluster K phages compared to results at 37°C.
     - We also heard from the Hatfull lab that the phage profile in an enrichment sample (during the enrichment incubation period) varies over time.

Note that a new 1-2 hour "phage extraction" step has also been added to the enrichment protocol prior to actual enrichment (2016). And the enrichment incubation time has been extended up to 2-5 days.

- Modifications of your base 7H9 media, such as changing the $CaCl_2$ concentration or possibly the cation used (e.g., $MgCl_2$), may also have an effect.
  - See Fullner and Hatfull (1997, Molecular Microbiology, 26(4):755-766) for $CaCl_2$ effects on mycobacteriophage infections
- Other – It has been suggested that adding glycerol to Phage Buffer (PB) may help stabilize some mycobacteriophages during extraction. The base 7H9 medium includes a small percentage of glycerol.

e. Direct plating processing
   - The volume of "extract" used per infection and plating can be increased when also employing concentrated TA.
     - For example, use 2-mL of filter-sterilized extract to infect 0.5-mL cells. After adding 2.5-mL 2X TA the content is poured onto standard L-agar base plate. This represents a 40-fold increase in extract volume tested per plate.

f. Concentrate large volumes of (direct plating PB) extract or enrichment prior to analysis.
   - Phage can be concentrated via simple microcentrifugation (see TEM preparation protocol).
   - Another suggestion: use Centricon concentrator (1 million kDa MW pore size). Suggest pre-filter through 0.45 μM sterilization filter first.

Step 2: Selection of Phages/Plaques for Purification

Multiple approaches exist to more carefully examine recovered phage for possible cluster identification and range of diversity.

a. Conditional growth testing – any growth condition that can be easily used to screen phages obtained from initial isolation would be helpful.
   - Growth temperature effects – Many/most cluster K phages may be temperature sensitive for growth at 42°C
     - Hope College has been testing their newly isolated and purified phages for the ability to propagate at 42°C; most/all presumptive K phages isolated from enrichments cultures incubated at ≤ 32°C appear uniquely sensitive for growth at 42°C.
   - $CaCl_2$ concentration effects
     - One year at Hope College, a one-time test of all student phages for ability to infect under different $CaCl_2$ concentrations (0 mM – 8 mM) revealed no obvious or possible effects. But see Fullner and Hatfull (1997) for description of $CaCl_2$ concentration effects of some mycobacteriophages.

- pH – other than offered as a "variable" condition that can be modified, no further information was shared on attempts to modify pH during isolation or screening
- Differential phage thermostability profiles?
- Differential phage sensitivity to other chemical treatments?

b. Plaque phenotypes – morphology and growth rate:
   - Examine the range of plaque morphologies observed in your class and encourage purification of a diverse set
   - Longer plate incubation times will allow slower-growing phages, seen as smaller plaques, to be observed and collected.
     - For example, the new non-cluster C myovirus, Phrappuccino, produces observable plaques after 48 hours of growth at 37°C.
   - Students can examine plates at 24 hours, record results, but then return plates for an additional 24 or more hours of incubation time.

c. Top agar concentration:
   - Reducing agar concentrations in TA will allow for faster plaque growth; higher agar concentrations in TA will slow plaque growth rates

d. Immunity test against known lysogens:
   - Test your newly isolated phages against known lysogens for some well-represented (sub)clusters (e.g., A1, K1); those unable to propagate on the lysogen can be identified as presumptive phage in that cluster
   - Caution: see recent findings with cluster N phages (cluster N lysogens immune to infection by non-cluster N phages)

e. PCR analysis using cluster-specific primers:
   - Use cluster-specific primers in PCR reactions to assay for and select/exclude a phage of a specific cluster(s)
   - There is some opportunity to test for multiple identifications in one PCR assay (Univ. of Wisconsin-Rive Falls, 2013 Grand Challenge)

Step 3: Selection of Phage Genome(s) for Sequencing

Unless you are sequencing the genomes of all purified phages prepared in your class/section this step reduces the number of "purified" phage genome(s) down to one or two for sequencing. Multiple approaches exist to more carefully examine purified phages for range of diversity.

a. Restriction digest analyses modifications:
   - Use web-based Phage Enzyme Tool (PET) software (developed by colleagues at Univ. Louisiana at Monroe, Gissendanner et al., 2014) to help identify possible cluster identification
   - Use second, cluster-distinguishing digest reactions, possibly selected from a dichotomous key, to test initial identification hypothesis based on PET analysis

b. PCR testing using cluster-specific primers:
   - See notes at end of previous section

c. TEM analysis of phage structure:

- Some phages of specific clusters can be presumptively identified following TEM image analysis.
- Caution: a new "unique" phage may have very (too) similar morphological characteristics to phages from a known, well-represented cluster

d. DOGEMS:
   - This opportunity technically occurs after the decision is made regarding which phage genome to sequence. It involves submitting a cocktail of phage genomes for sequencing.
   - If your phage genome cocktail includes a "unique" genome or two, those may be detected with the deconvoluting of assembled genomes.
   - Follow-up work to fully annotate any/all "unique" phage genomes is expected.

Other Experimental Procedure Considerations

a. Low DNA yield effect
   - Perhaps there are novel forms of mycobacteriophages that we have missed due to low DNA levels recovered using our standard genomic DNA isolation protocols
     - Easiest solution may be to perform additional genomic DNA isolation preps and pool samples (may require subsequent concentration step)

b. Non-dsDNA mycobacteriophage isolation
   - Most institutions have seen some students collect phage genomic DNA that runs as a smear on an agarose gel (QC or restriction digest) – could the genomes of these phages be a form other than dsDNA?

Engage the Students

Frame the "phage discovery" task for students to encourage search for less well-represented phages.

a. What are some explanations for why some phages have been found in much lower numbers (down to singletons)? Is there anything consistent about those cases (from soil collection through DNA isolation and analysis)?
b. For each possible explanation, how would those phages come through our procedures (consider plaque numbers, plaque features, plaque growth properties, etc.)?
c. For each possible explanation, is there some experimental procedure you can include (as a protocol modification) that would enhance the likelihood of you recovering that phage?