PLOS PATHOGENS

REVIEW

Mycobacteriophages: From Petri dish to patient

Graham F. Hatfull [ORCID] *

Department of Biological Sciences, University of Pittsburgh, Pittsburgh, United States of America

* gfh@pitt.edu

Abstract

Mycobacteriophages—bacteriophages infecting *Mycobacterium* hosts—contribute substantially to our understanding of viral diversity and evolution, provide resources for advancing *Mycobacterium* genetics, are the basis of high-impact science education programs, and show considerable therapeutic potential. Over 10,000 individual mycobacteriophages have been isolated by high school and undergraduate students using the model organism *Mycobacterium smegmatis* mc^{2}155 and 2,100 have been completely sequenced, giving a high-resolution view of the phages that infect a single common host strain. The phage genomes are revealed to be highly diverse and architecturally mosaic and are replete with genes of unknown function. Mycobacteriophages have provided many widely used tools for *Mycobacterium* genetics including integration-proficient vectors and recombineering systems, as well as systems for efficient delivery of reporter genes, transposons, and allelic exchange substrates. The genomic insights and engineering tools have facilitated exploration of phages for treatment of *Mycobacterium* infections, although their full therapeutic potential has yet to be realized.

Competing interests: I have read the journal's policy and have the following conflicts: I receive research support from Janssen Pharmaceuticals, and am a consultant for Tessera Inc. and Janssen.

Why study phages and mycobacteriophages in particular?

Bacteriophages were discovered a little over a hundred years ago and quickly developed into workhorses for molecular biology and genetics [1]. They provided suitably sized genomes for development of early DNA sequencing techniques, although phage genomics did not really take off until the early 1990s. Mycobacteriophages—the subject of this review—became a focus of interest for addressing 5 key priorities. First, is the deep exploration of phages that infect a single common host strain (*Mycobacterium smegmatis* mc^{2}155) for insights into viral diversity, evolution, and origins? Second, is the use of phage discovery and genomics as a platform for advancing science education? Third, is the investigation of all aspects of host–phage dynamics, including host range, life cycle regulation, viral defense, and counterdefense? Fourth, is the exploitation of mycobacteriophage genomes to develop tools for simplifying and advancing *Mycobacterium* genetics? Lastly, is the potential for clinical use of mycobacteriophages to control *Mycobacterium* diseases? I will provide introductory information for the general reader on the world of bacteriophages and will then discuss recent advances toward these 5 priorities.

Introduction to the phage world

Bacteriophages are viruses that infect bacteria. They are prevalent in the environment and throughout the human body, and there are an estimated 10^{31} phage particles in total in the biosphere, making them the majority of all biological entities [2–4]. The phage population is old —plausibly originating over 3 billion years ago—and is highly dynamic. Phages are constantly infecting bacterial hosts, replicating, and replenishing the population, and there is estimated to be about 10^{23} phage infections per second, with the phage population replacing itself every few days [4].

Given these parameters, it is not surprising that phages are genomically highly diverse [5]. Phage lytic infections—those resulting in replication and production of new phage particles— kill the bacterial host, and there is strong selection for development of resistance and survival. Such mechanisms are widespread and include receptor variation, restriction-modification systems, and CRISPR-Cas arrays among a large variety of newly discovered systems [6]. Phages would not exist if they could not circumvent these defenses and resistance mechanisms, and they actively coevolve along with their hosts [7]. They do so by a variety of mechanisms including the acquisition of counterdefenses or switching to a new bacterial host that lacks the defenses. These dynamics have prominently shaped microbial evolution.

Bacteriophages can use DNA or RNA as their genetic material, which is typically encapsulated in a protein shell, sometimes with lipid components [8]. A variety of morphologies can be seen by electron microscopy, although the predominant types are those with a tail attached to a capsid (or phage head) containing double-stranded DNA (dsDNA) [9]. These dsDNA tailed phages (the *Caudovirales*) can be divided into 3 families depending on their tails: the *Myoviridae* with contractile tails; the *Siphoviridae* with long, flexible, noncontractile tails; and the *Podoviridae* with short stubby tails. The capsids are usually isometric (soccer ball-like) or prolate (elongated, rod-like) and vary in size according to the length of the DNA packaged within them. The virion genome is typically linear but may genetically be either linear or circular depending on the DNA packaging mechanism [10]. Phages using *cos*-type packaging signals have genomes with defined physical ends such that all phage particles have the same genome termini. In contrast, phages using headful DNA packaging systems have genomes that are circularly permuted and terminally redundant, with the physical genome ends different in each particle.

A key phage property is host specificity. Phage host ranges vary considerably, but it is uncommon for any given phage to infect bacteria in more than one genus. In the few instances where host range spans different genera, the bacteria are usually phylogenetically closely related [11,12]. Sometimes host range is restricted to a few strains within a species, and some phages infect only a single known bacterial strain.

Finally, phages typically can be ascribed to having one of 2 lifestyles, either lytic or temperate. The key distinction is that temperate phages can establish lysogeny within a host cell, whereas lytic phages do not. Both lytic and temperate phages can undergo cycles of lytic growth in which the phage replicates, the cell dies, and progeny phage particles are released. However, this is the sole outcome for a lytic phage (a canonical example is phage T4 of *Escherichia coli*). In contrast, temperate phages can enter one of 2 life cycles upon infection: lytic growth as described above, or lysogeny in which the lytic genes are switched off (a canonical example is phage lambda of *E. coli*); a prophage is established either by chromosomal integration or plasmidial extrachromosomal replication, and the lysogenic cell continues to grow. The frequency of lysogeny varies enormously depending on various parameters (e.g., growth state, multiplicity of infection), and can be as low as a few percent, or greater than 50% of infections. The prevalence of temperate phages varies depending on the host, and this is reflected in

the richness of resident prophages in sequenced bacterial genomes [13]. It should be noted that some types of temperate phages do not lyse the host cell when replicating, including filamentous phages (inoviruses) such as *Pseudomonas* phage Pf [14].

Mycobacteriophages are viruses of mycobacteria

Mycobacteriophages are phages that infect *Mycobacterium* hosts. The genus *Mycobacterium* includes several important human pathogens, including *Mycobacterium tuberculosis*, the causative agent of tuberculosis, and *Mycobacterium avium* complex (MAC) and *Mycobacterium abscessus* complex (MAB) that are notable opportunistic pathogens. Mycobacteria characteristically are acid-fast staining and have unusual cell walls containing mycolic acids, glycolipids, and glycopeptidolipids (GPLs) surrounding the cytoplasmic membrane and peptidoglycan layer. This unusual cell wall presents a unique challenge to bacteriophages that must bind specifically, navigate their tail tips to the membrane, and inject their DNA. At the conclusion of lytic growth, they must also compromise these walls to release progeny phage particles from the cell.

The first mycobacteriophages were isolated in the 1950s—using *M. smegmatis* as a host—primarily for phage-typing mycobacterial infections, taking advantage of mycobacteriophage host range variations [15]. Some of these typing phages have been recovered, although others are lost [16,17]. The first mycobacteriophage genome to be fully sequenced was that of L5 [18], which was originally isolated in the 1950s in Japan [19]. This was followed by reports of genome sequences of phages D29, TM4, and Bxb1 [20–22], but further exploration of mycobacteriophage diversity would require additional phage isolation, most prominently using *M. smegmatis* mc^{2}155 as a host [23].

Phage discovery and genomics as integrated research-education programs

With just a handful of sequenced mycobacteriophages available, in the early 2000s it was clear that gaining a fuller understanding of viral diversity would need a deeper dive into phage isolation and genomic analyses [24]. An alternative approach—metagenomic analysis of enriched viral samples—is also attractive, but generally provides informative data sets rather than individual phages that can be propagated, engineered, investigated mechanistically, and potentially used translationally. Envisaging the need to isolate phages at substantial scale, we developed integrated research-education programs in which phage discovery and genomics is the primary goal, providing research opportunities to young scientists including high school students and first year undergraduate students [25]. The Phage Hunters Integrating Research and Education (PHIRE) starting in 2002 provided research experiences locally in Pittsburgh [26,27] with the platform expanding into the Science Education Alliance Phage Hunters Advancing Genomics and Evolutionary Science (SEA-PHAGES) in 2008 [28,29]; both programs are supported by the Howard Hughes Medical Institute (HHMI). The SEA-PHAGES program has grown with addition of institutions from the United States and beyond, and now has over 170 participating schools with over 5,500 students each year. SEA-PHAGES is an example of an Inclusive Research Education Community (iREC) that encompasses institutions ranging from community colleges to Research I universities and involves all students without pre-selection based on generally irrelevant criteria [29,30].

These programs have produced a massive archive of individual bacteriophages and a rich data set of sequenced and manually annotated genomes focusing on carefully chosen bacterial hosts. The most widely used host is *M. smegmatis* mc^{2}155, a natural extension from earlier studies; over 10,000 individual phages have been isolated, by far the deepest investigation of phages using a single bacterial strain. Over 2,000 have been sequenced, and sampling of the

remainder suggests that no substantial component of the diversity has been missed. An additional 10,000 phages have been isolated using other bacterial genera within the phylum *Actinobacteria*, including, *Arthrobacter*, *Corynebacterium*, *Gordonia*, *Microbacterium*, *Rhodoccocus*, and *Streptomyces* [31]. Information about the entire phage collection is available at http://phagesdb.org. A genome map of mycobacteriophage Tweety [32] is shown as an example in Fig 1 and illustrates several of the features described below.

Mycobacteriophage genomics, diversity, and architecture

What have we learned about mycobacteriophage genomes? Several key themes have emerged. First, the overall diversity is impressive, although complex in its structure. Comparative genomics clearly shows that some phages are closely related, and a system was established early on to accommodate these relationships. Phages that are closely related to each other are grouped
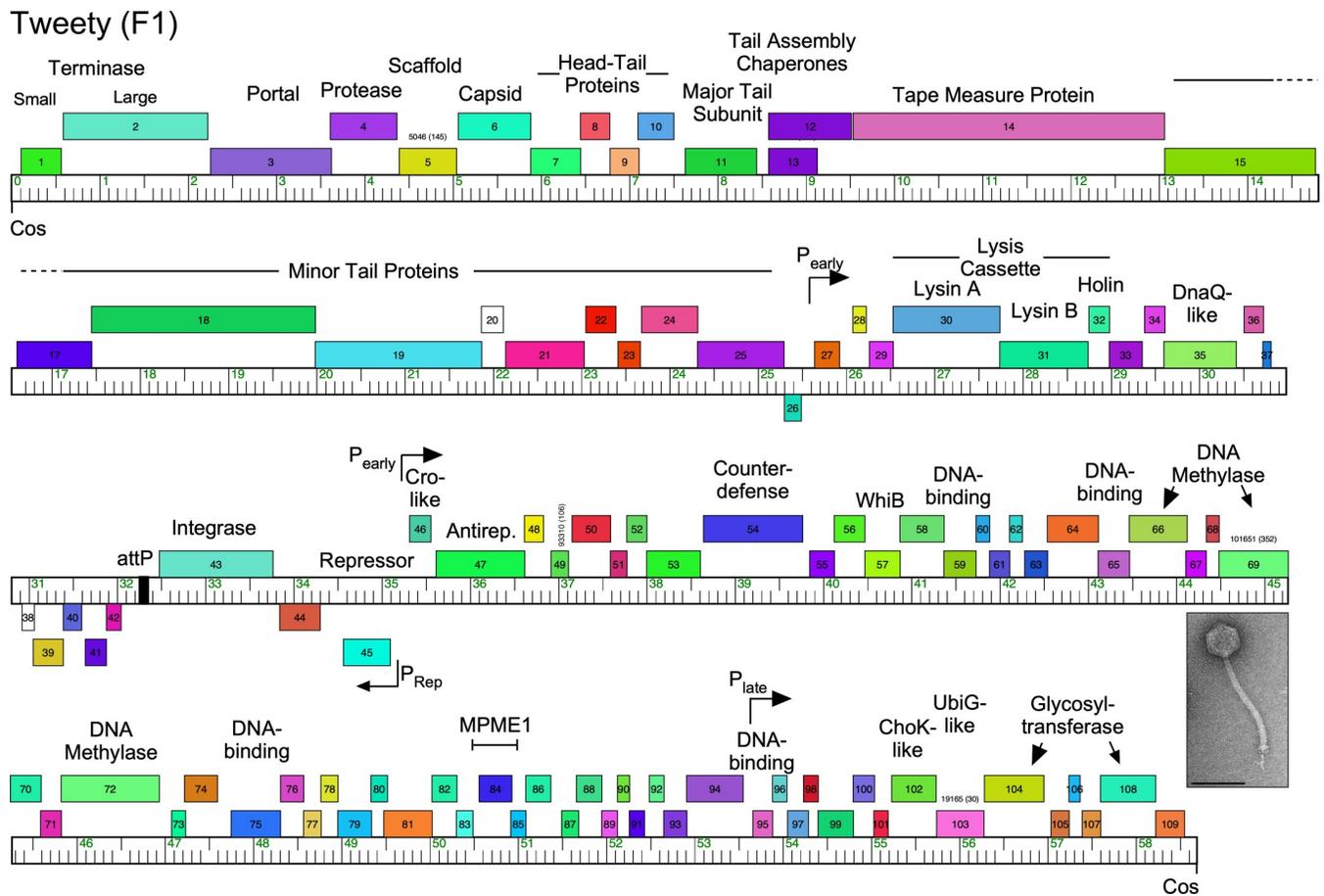

Fig 1. Genome organization of Mycobacteriophage Tweety. Phage Tweety is a member of Subcluster F1, is temperate, has a siphoviral morphology (inset, electron micrograph; scale marker, 100 nm), and its genome organization illustrates several common features. The genome is represented as a bar with markers every 1 kbp, and the predicted genes are shown as colored boxes above or below the genome, indicating rightwards- and leftwards-transcription, respectively. Gene numbers are shown within the boxes and predicted functions are shown above. The virion structure and assembly genes are arranged in a rightwards-transcribed operon at the left end of the genome (genes 1 to 25), followed by the lysis cassette. The repressor (*45*) and putative Cro-like (*46*) genes are divergently transcribed, and the integration cassette (*attP* and *43*) are located nearby, with *attP* defining a "left arm" from *cos–attP* and a "right arm" from *attP–cos*. Note that the left arm genes—predominantly the virion structure and lysis genes—are relatively large and have known functions, whereas the right arm genes are relatively small, and most have unknown functions. Shown are the positions of the MPME mobile element and the counterdefense gene, *54*. The positions of putative early and later lytic promoters as determined by RNAseq are indicated. Note that genes are arranged in long operons such that there are relatively few transcriptional changes (tdc), such as at the *25/26* and *45/46* junctions.

https://doi.org/10.1371/journal.ppat.1010602.g001

into "clusters," with different phages in different clusters (e.g., Cluster A, Cluster B, Cluster C). Many clusters have recognizable divisions, as seen through average nucleotide identity (ANI) comparisons, and are thus divided into "subclusters" (e.g., Subcluster A1, Subcluster A2, Subcluster A3). Phages for which there are no close relatives are recognized as "singletons". The initial parameters used for cluster grouping required recognizable nucleotide sequence similarity spanning greater than 50% of the genome lengths, and with only a few dozen genomes all mycobacteriophages could be grouped without any evident challenge to these thresholds [33–35].

As the mycobacteriophage collection grew to several hundred genomes, these "boundaries" of similarity between clusters became fuzzy, with the discovery of phages that only just meet or just miss the thresholds for inclusion in a cluster, or are close to the threshold for inclusion in more than one cluster [36]. This suggested that the cluster boundaries are artificial and do not reflect hard biological barriers to recombination and exchange of information between phage genomes. Phages therefore appear to form a continuum of diversity but with unequal representation and heterogenous sampling [36]. This heterogeneity is reflected in the abundance of more than 700 Cluster A phages and more than 350 Cluster B phages, compared to the 7 singletons and 10 clusters with 5 or fewer phage members. As the phage collection continued to expand, the thresholds were redefined with clearer and more quantifiable parameters, and we currently use a threshold of 35% shared gene content (SGC) for cluster inclusion [37]. This is determined by assorting all of the gene products into groups of related proteins ("phamilies," or "phams") using the program Phamerator [38] and a pipeline based on MMseqs2. This can then be used to determine the number of shared phams, and a tool is available at https:// phagesdb.org/ for simple pairwise SGC comparisons. The overall genomic relationships among the mycobacteriophages can be viewed as a network phylogeny using Splitstree [39], simplified by using a single representative of each cluster and subcluster together with the 7 singletons (Fig 2A). There are numerous examples in which 2 genomes from different clusters (e.g., Bxb1 from Subcluster A1 and Bipper from Cluster Y) have no shared genes at all. There are also examples of phages in different clusters sharing considerable number of genes (e.g., 25% between Fruitloop and Omega in Clusters F and J, respectively). Not surprisingly, there is also considerable intracluster diversity illustrated by the example of the Cluster F phages (Fig 2B). There are currently 5 F subclusters (F1 to F5), with by far the vast majority (95%) grouped in Subcluster F1 (Fig 2B). Pairwise comparisons show that genomes within Subcluster F1 can share as few as 40% of their genes. In general, the cluster/subcluster/singleton "taxonomy" is of pragmatic use—because of other shared properties—but it is a crude representation of phage diversity.

The second key finding is that phage genomes are architecturally mosaic. Each genome is essentially a compilation of modules or mosaic "tiles" that can be assembled in vast numbers of different combinations (see Fig 1) [23]. These modules are often single genes, and amino acid sequence comparisons reveal many instances of genomes having a shared gene but in different genomic contexts, such that the flanking genes are different. Most commonly, there is a lack of nucleotide sequence similarity and only weak amino acid sequence identity, reflecting relatively old recombination events in phage evolution. The mechanisms giving rise to genome mosaicism are not completely clear, but likely involve non-sequence-directed illegitimate recombination events between genomes [3,23]. This is supported by observations of mosaic boundaries between genomes where there is extensive sequence similarity, and the finding that these usually occur at gene—and sometimes domain—boundaries [3,4]. Mobile elements such as transposons and homing endonucleases are likely also involved, as well as integrase-mediated site-specific recombination events [40,41]. We note that several transposons have been noted in the mycobacteriophage genomes, including the novel small mycobacteriophage
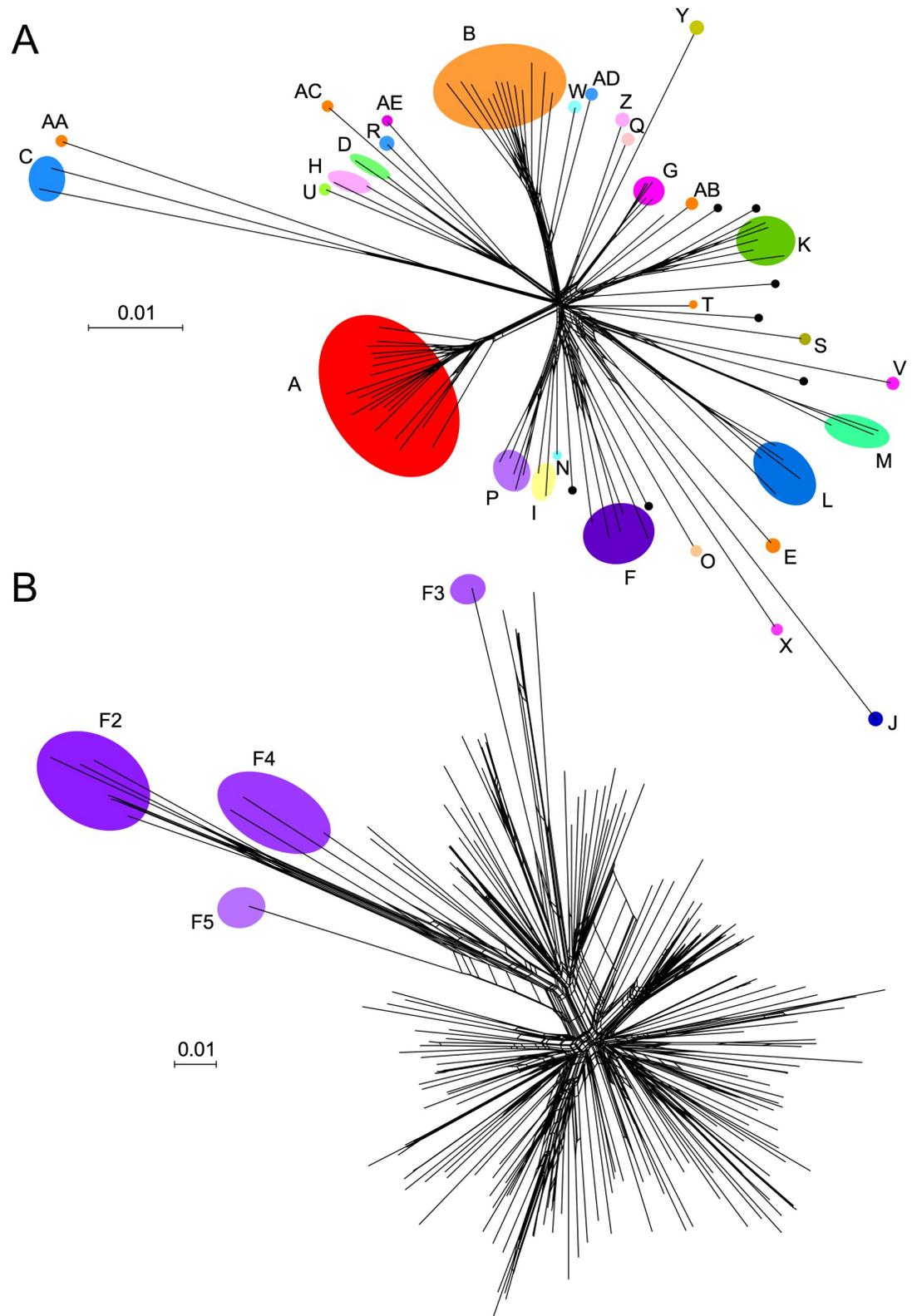
Fig 2. Diversity of mycobacteriophage genomes displayed as network phylogenies based on shared gene content. (A) Relationships among representative members of clusters, subclusters, and singleton genomes. One member of each mycobacteriophage cluster and subcluster together with the 7 singletons were compared using Splitstree [39] with a nexus file

recording the numbers of shared genes. Clusters are illustrated with colored shading; note that some clusters (e.g., Cluster A) contain several subclusters indicated as nodes, whereas other clusters are not subdivided. Singletons are shown as unlabeled black circles. (**B**) Diversity of Cluster F mycobacteriophages. All currently sequenced Cluster F mycobacteriophages ($n = 188$) are displayed as nodes in a network phylogeny using Splitstree. Colored circles show the positions of the Subclusters F2 to F5 genomes; all of the others ($n = 177$) are grouped in Subcluster F1. This illustrates the substantial intracluster diversity, and pairwise comparisons of Subcluster F1 phages show they may share as few as 40% of their genes.

https://doi.org/10.1371/journal.ppat.1010602.g002

mobile elements (MPMEs) first described in Cluster G genomes [42] (see Fig 1). Phage-encoded homologous recombination systems that can act on short segments of imperfect homology are also implicated in events that can give rise to mosaicism [43].

The third key finding is that there are clearly architectural constraints on phage genomes. In general, phage average gene size is smaller (approximately 0.6 kbp) than in bacterial genomes (approximately 1 kbp) [33], and the genes are assembled into longer operons than in bacteria, so that there are fewer changes in the direction of transcription (tdc) (see Fig 1). The vast majority of mycobacteriophages (in total, and as defined by clusters/subclusters/singletons) have siphoviral morphologies, and these all contain a characteristic operon (approximately typically 20 to 25 kbp) of virion structure and assembly genes (Fig 1). These genes are on average more similar in length to bacterial genes. In contrast, the nonstructural genes in the rest of the genome are much smaller on average, so that there is a bimodal distribution of phage gene length (Fig 1). Genome length varies substantially, and mycobacteriophages with siphoviral morphologies vary from approximately 40 kbp to approximately 110 kbp. Because the genome requirements for virion structure and assembly genes don't vary greatly, it is the number of nonstructural genes that varies as genome length changes. Some of these genes have predicted functions such as in DNA replication and metabolism, but most are of unknown function (Fig 1). The total number of genes of unknown function among the 2,000 sequenced mycobacteriophages is impressive (>100,000) and suggests that the phage population harbors the greatest reservoir of unexplored gene diversity and function.

Mycobacteriophage lifestyles

M. smegmatis mc^{2}155 is prophage-free, but a majority of the mycobacteriophages isolated on this strain are temperate. Specifically, of the 31 clusters and 7 singletons, 21 are temperate or have temperate representatives (Clusters A, E, F, G, I, J, K, L, M, N, P, Q, T, X, Y, Z, and singletons IdentityCrisis, Kumao, LilSpotty, MalagasyRose, and Sparky). It is noteworthy that it is relatively common to isolate phages that have clear plaque morphologies and appear obligatorily lytic but are members of a cluster where most other members are temperate. An early example of this was phage D29 [20], which is grouped in Subcluster A2 along with temperate phages such as L5 [18], but has a deletion that removes the repressor gene [20,44]. There are many similar examples in other clusters.

The temperate phages typically encode a repressor that is required for the maintenance of lysogeny (analogous to lambda cI), and an integrase that mediates site-specific integration into the bacterial genome to form a prophage. It is noteworthy though that several phages—all within Cluster A—form extrachromosomal prophage replicons (i.e., plasmidial) and code for *parAB* partitioning systems instead of an integrase [45]. For many mycobacteriophages (with the notable exception of Cluster A), the repressor gene is centrally located and transcribed divergently from the early lytic genes, with the repressor's most proximal gene being Cro-like (i.e., analogous to phage lambda Cro). Typically, both repressor and Cro-like proteins contain predicted DNA-binding domains and their genes are separated by an intergenic region of 100 to 300 bp containing promoters for early lytic and repressor gene expression, and an operator
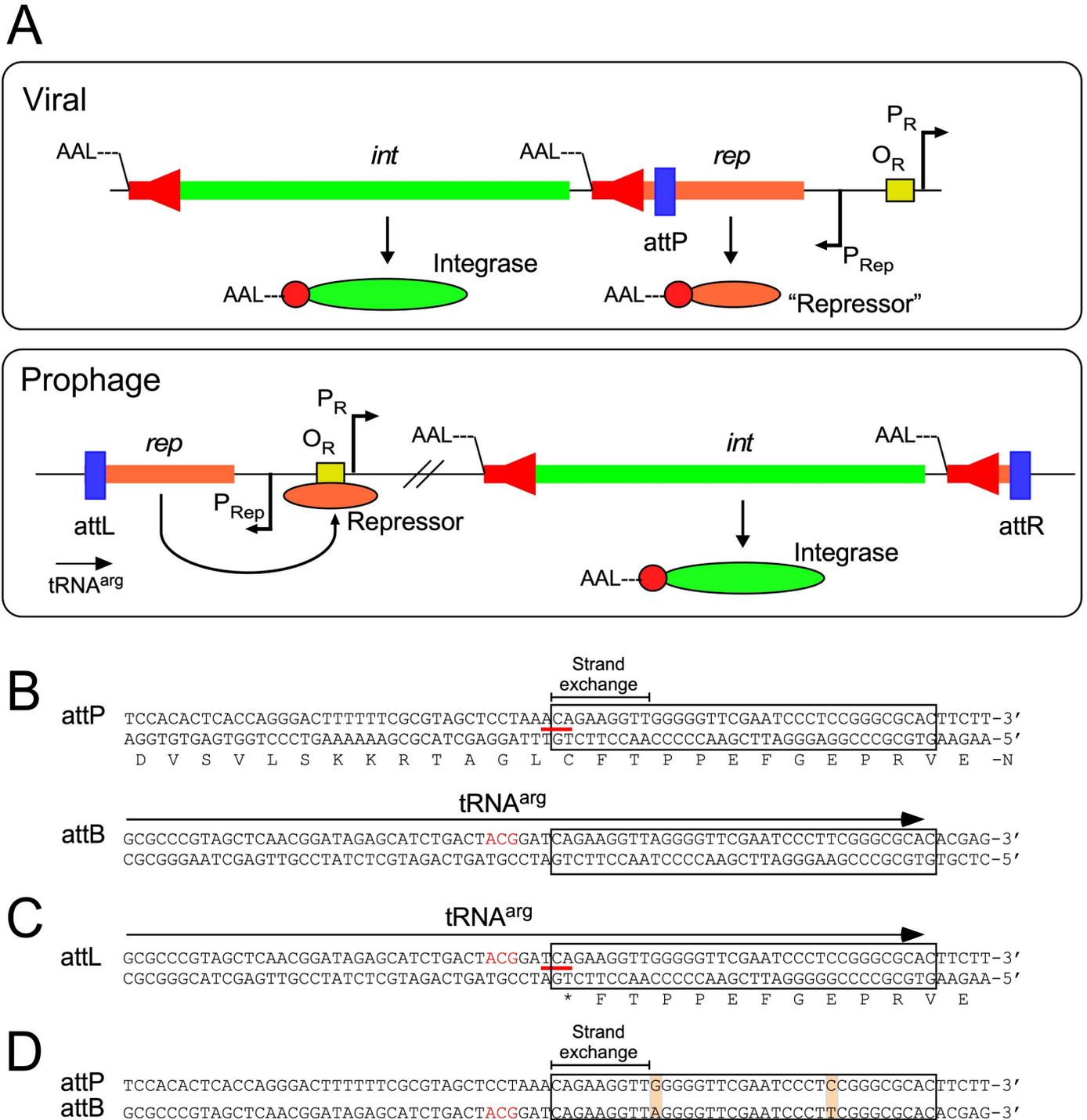
site for repressor binding [46,47] (see Fig 1). However, few of these repressor or Cro-like proteins have been studied in detail. Two unusual immunity systems warrant further discussion.

Integration-dependent immunity: An atypical system for lysogenic establishment

First, several mycobacteriophages (including those in Clusters G, N, and P) use an unusual integration-dependent immunity system (Fig 3) [46]. In the canonical phage lambda system, establishment of superinfection immunity—prophage repressor expression that prevents superinfection—is distinct from chromosomal integration [48]. Both are under regulatory control by the cII protein but are separable molecular events. That is, establishment of superinfection immunity (protection from reinfection by particles of the same phage) does not require integration, and integration is not required to establish immunity. In contrast, the integration-dependent immunity systems fundamentally operate differently, and there are several distinguishing and unusual features [46]. First, the repressor and integrase genes are cotranscribed in an operon, and the attachment site (*attP*) used for integrase-mediated site-specific recombination is within the repressor gene itself (Fig 3A). This is seemingly paradoxical because integration and prophage formation leads to loss of the 3′ end of the repressor gene, and yet it is the prophage-expressed form of the repressor that needs to be active to confer immunity (Fig 3A). The resolution of this conundrum is that the virally encoded form of the repressor contains a C-terminal ssrA-like tag that targets it for proteolytic degradation. Integration leads to a C-terminally truncated repressor protein that is stable because of loss of the proteolytic tag, and thus active for superinfection immunity (Fig 3A). However, this leads to another difficulty. Lysogeny is usually established in only a small proportion of infections (5% to 20%, depending on conditions), and if integration always results in immunity, what prevents all infections from entering lysogeny? The answer is that the integrase protein itself is also under proteolytic control, and similarly carries a C-terminal ssrA-like tag (Fig 3A). The frequency of lysogeny is therefore likely determined by the activity of host proteases, analogous to the way *E. coli* FtsH determines the activity of lambda cII [49].

Like many phages that use integrases of the tyrosine family of site-specific recombinases (i.e., "tyrosine integrases"), all of the phages using integration-dependent immunity systems use an *attB* site overlapping a host tRNA gene (Fig 3B). The common core sequence—a 35-bp region shared by *attP* and *attB*—must therefore contain not only the sequences for integrase binding and strand cleavage, but also the 3′ end of the tRNA so that tRNA function is maintained in the lysogen (Fig 3B). Amazingly, the sequences for BPs integration at *attL* not only result in loss of the C-terminal ssrA-like tag from the repressor, but also the bacterium genome sequence base pair immediately adjacent to the common core changes the third position of a cysteine codon in the repressor gene and introduces a translation termination codon (Fig 3C). In some instances, such as with phage BPs, there are mismatches between the common cores in *attP* and *attB*, although these do not impair normal base paring in the tRNA (Fig 3D).

Finally, these integration-dependent immunity systems do not evidently code for a recombination directionality factor (RDF) that normally controls the ratio of integrative and excisive integrase-mediated recombination events [50]. As such, the frequency of these may simply be determined by the amount of stable integrase in the cell. A consequence of this is that the attachment sites (*attP, attL, attR*) do not require arm-type integrase binding sites, and only use the common core sequences around the scissile bonds for integrase binding (Fig 3). It is easy to imagine how this evolved, because *attP* lies within the repressor gene, and an *attP* site the size of that for lambda (250 bp), would impose severe constraints on the amino acid sequence of the repressor (Fig 3). These types of immunity systems are not only found in

A

Viral

AAL--- *int* AAL--- *rep* P_R O_R

↓ ↓

AAL-- Integrase AAL-- "Repressor"

attP P_{Rep}

Prophage

rep P_R O_R AAL--- *int* AAL---

attL P_{Rep} Repressor attR

tRNA^{arg}

AAL-- Integrase

B

Strand exchange

attP
```
TCCACACTCACCAGGGACTTTTTTCGCGTAGCTCCTAAACAGAAGGTTGGGGGTTCGAATCCCTCCGGGCGCACTTCTT–3'
AGGTGTGAGTGGTCCCTGAAAAAAGCGCATCGAGGATTTGTCTTCCAACCCCCAAGCTTAGGGAGGCCCGCGTGAAGAA–5'
 D  V  S  V  L  S  K  K  R  T  A  G  L  C  F  T  P  P  E  F  G  E  P  R  V  E  –N
```

tRNA^{arg}

attB
```
GCGCCCGTAGCTCAACGGATAGAGCATCTGACTACGGATCAGAAGGTTAGGGGTTCGAATCCCTTCGGGCGCACACGAG–3'
CGCGGGAATCGAGTTGCCTATCTCGTAGACTGATGCCTAGTCTTCCAATCCCCAAGCTTAGGGAAGCCCGCGTGTGCTC–5'
```

C

tRNA^{arg}

attL
```
GCGCCCGTAGCTCAACGGATAGAGCATCTGACTACGGATCAGAAGGTTGGGGGTTCGAATCCCTCCGGGCGCACTTCTT–3'
CGCGGGCATCGAGTTGCCTATCTCGTAGACTGATGCCTAGTCTTCCAACCCCCAAGCTTAGGGGGCCCGCGTGAAGAA–5'
                                        *  F  T  P  P  E  F  G  E  P  R  V  E
```

D

Strand exchange

attP
```
TCCACACTCACCAGGGACTTTTTTCGCGTAGCTCCTAAACAGAAGGTTGGGGGTTCGAATCCCTCCGGGCGCACTTCTT–3'
```
attB
```
GCGCCCGTAGCTCAACGGATAGAGCATCTGACTACGGATCAGAAGGTTAGGGGTTCGAATCCCTTCGGGCGCACACGAG–3'
```

Fig 3. Integration-dependent superinfection immunity systems. (A) Typical organization of the immunity regions of phages encoding integration-dependent immunity systems. In the viral genome, the repressor (*rep*) and integrase (*int*) genes are cotranscribed from the P_{Rep} promoter, and the phage attachment site (*attP*, blue box) is located within the repressor gene. The virally encoded Integrase and Repressor proteins both carry C-terminal ssrA-like tags (——LAA, or similar) targeting the proteins for proteolytic degradation (red arrows in the int and rep genes; red circles in the proteins; the N-termini are indicated). The virally encoded form of the Repressor is not active in conferring superinfection immunity. The establishment of lysogeny requires integrase-mediated site-specific recombination between the phage *attP* site and a chromosomal *attB* site (which overlaps a host tRNA gene) to form a prophage. Integrative site-specific recombination removes the ssrA-like tag from the repressor and the stable, active form of the repressor binds to the operator (O_R, yellow box) to shut down the early lytic promoter (P_R) and confer superinfection immunity. (**B**) Organization of the *attP* and *attB* sites of phage BPs and *M. smegmatis*. Both DNA strands are shown, and the amino acid sequence of part of the leftwards-transcribed repressor is shown; DNA and protein polarities are indicated. The 35-bp common core sequence (conserved in *attP*,

attB, *attL*, and *attR*) is boxed, and the region within which strand exchange for recombination must occur is indicated. The codon spanning the left side of the common core is indicated by a red line, and the third position base is shown in bold pink type. The location of the tRNA^{arg} gene at *attB* is shown by an arrow; the anticodon is shown in red type. (**C**) Organization of the BPs *attL* site. Conservation of the common sequence between *attP* and *attB* results in construction of an active tRNA^{arg} gene when the prophage is established. However, at *attL* the repressor gene encounters a termination codon (TGA) formed at the junction of the bacterial and phage sequences, changing the third base of a cysteine codon (red line) to a nonsense codon; the third position base is shown in bold pink type. (**D**) Alignment of *attP* and *attB* showing the common core (boxed), and the positions of 2 mismatches in the common core, neither of which introduced mispairing in the tRNA^{arg} product. Only the top strand of each site is shown.

mycobacteriophages, but also in a variety of the diverse prophages resident in sequenced *Mycobacterium* genomes [51].

The repressor–stoperator system of Cluster A mycobacteriophages

The second system of interest is the repressor–stoperator systems of Cluster A phages. Bioinformatic prediction of repressor genes can be challenging, and the repressors of Cluster A genomes (e.g., L5, Bxb1) were identified using genetic and biochemical methods [52,53]. Moreover, the Cluster A repressor genes are situated toward the right end of the genome within a large set of leftwards-transcribed genes and are not closely linked to the centrally located integrase gene. The Cluster A genomes unusually possess many asymmetric 13 to 14 bp sequence motifs located throughout the genome to which the repressor binds [54,55]. One of these is a true operator site and overlaps the early lytic promoter, P$_{left}$. The others are referred to as "stoperators," as they are proposed to be sites where repressor binding results in termination of transcription, rather than inhibition of transcription initiation [55]. The sites are located primarily within short intergenic regions, are in 1 orientation with respect to the direction of transcription, and confer repressor-dependent orientation-specific down-regulation of a reporter gene [55]. The presumed function is to silence phage genes in lysogeny and prevent leaky expression that could impose a growth disadvantage to the lysogen [55]. The immune specificities of the numerous Cluster A phages are highly complex but reveal insights into how these specificities have evolved [54].

RNAseq approaches show that many mycobacteriophages display 2 main patterns of lytic gene expression, occurring early (approximately 30 min) and late (approximately 120 min after infection [44,56–58]). Early lytic promoters are often SigA-like, with recognizable −35 and −10 hexamers recognized by the host RNA polymerase and operator sites for repressor binding [46,55,59]. Few late promoters have been characterized, but transcription starts upstream of the virion structure and assembly genes sometimes upstream of the right *cos*-site (Fig 1). Further investigation of the phage late promoters and their regulation is warranted as they are highly active and potentially useful for heterologous gene expression.

Phage integration systems

Most of the temperate mycobacteriophages encode integration systems similar to those described above for the integration-dependent immunity systems. The integrases can be either lambda-like tyrosine integrases (Int-Y) or large serine-integrases (Int-S), although the former are more common (see Fig 1). At least 14 different *attB* sites have been identified (Fig 4A), distributed around the *M. smegmatis* genome. Interestingly, the Int-Y genomes exclusively use *attB* sites overlapping host tRNA genes, and the common core facilitates reconstruction of the tRNA upon integration. Because these common cores are 25 bp or longer, the *attB* site for most Int-Y phages can be predicted using homology searches. This is not so for Int-S phages, for which the common core can be very small (as few as 3 bp), and they do not integrate into tRNA genes [60–63]. Their *attB* sites therefore have to be determined experimentally, and Bxb1 and Bxz2 (Subcluster A1 and A3, integrating into attB-7 and attB-4, respectively; Fig 4A)
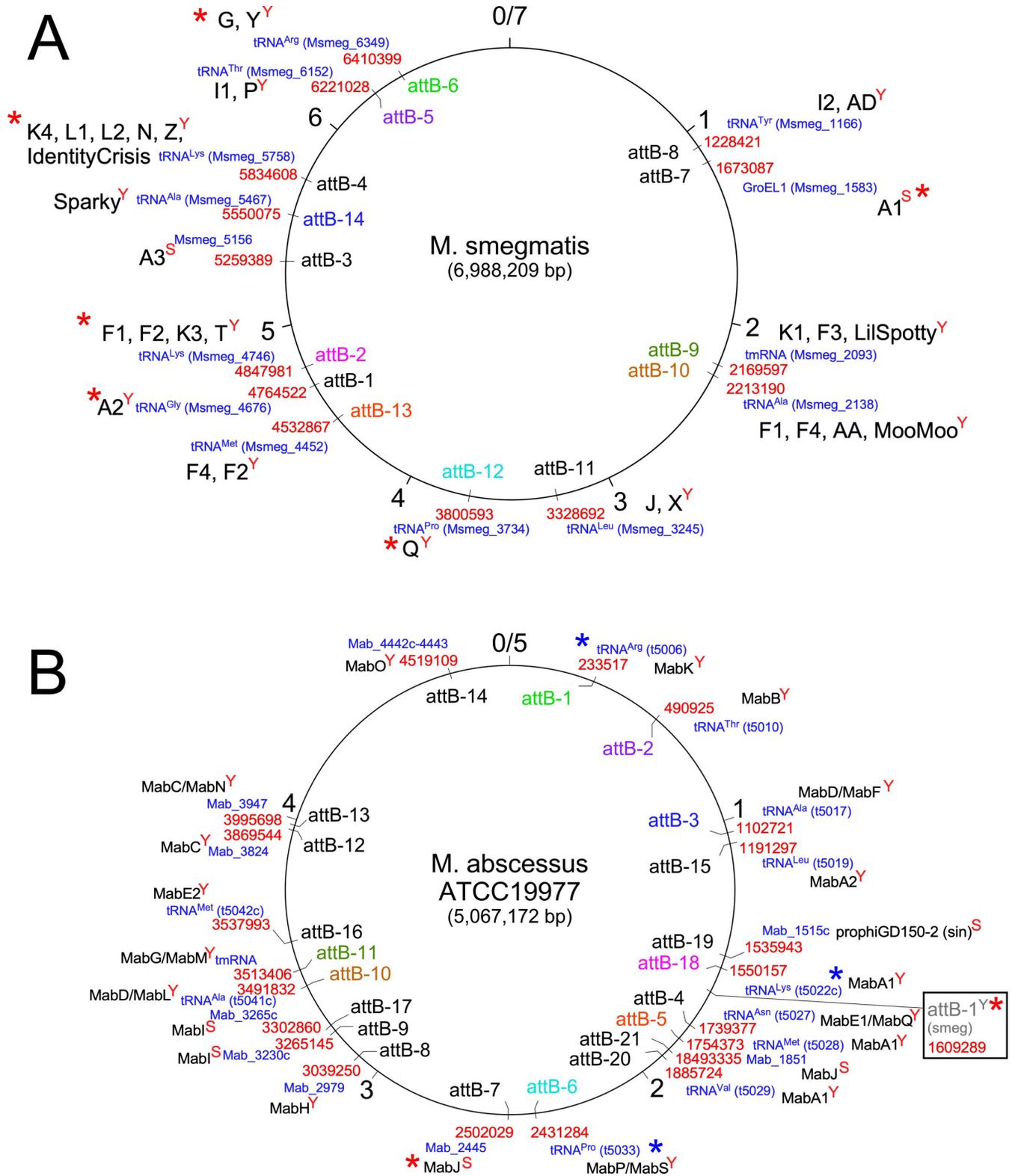
Fig 4. Chromosomal *attB* attachment sites used by temperate phages in *M. smegmatis* and *M. abscessus.* (**A**) *attB* sites in the *M. smegmatis* mc²155 genome. The 7-Mbp *M. smegmatis* genome is represented as a circle with markers at each kbp indicated. The *attB* sites (e.g., attB-1, attB-2, etc.) are shown, together with the genome coordinates in red type; the genes are shown in blue type. The font color of the *M. smegmatis attB* sites (i.e., attB-x^smeg) label is coordinated with homologous sites in *M. abscessus* (i.e., attB-x^Mab) shown in panel B. The names of clusters or subclusters (or the phage name if it is a singleton) within which

one or more phages use that site for integration are shown in large black type. "Y" or "S" superscripts on the cluster names denote whether the site is used by tyrosine- (Y) or serine (S) -integrases. *attB* sites for which integration-proficient vectors have been developed are indicated with a red asterisk. (**B**) *attB* sites in the *M. abscessus* ATCC19977 genome mapped by identification of integrated prophages. The font color of the *attB* sites (i.e., attB-x^Mab) is coordinated with their homologues in *M. smegmatis* (i.e., attB-x^smeg) shown in panel A. Genome coordinates are shown in red type; genes are shown in blue type. The clusters/ subclusters of prophages for which members are found integrated in those sites are shown in black type. "Y" or "S" superscripts on the cluster names denote whether the site is used by tyrosine- (Y) or serine (S) -integrases. The integration vectors based on phage L5 (attB-1^smeg) integrate at a homologous site in *M. abscessus* (shown in box) although no prophages have been identified there. This and the attB-7^Mab site are only ones for which integration vectors have been shown to work (marked with red asterisks), but several vectors developed for *M. smegmatis* are predicted to also work in *M. abscessus* (blue asterisks).

have been shown to integrate within the protein-coding genes for *groEL1* (Msmeg_1583) and a putative DNA glycosylase (Msmeg_5156), respectively [61,64]. Interestingly, Bxb1 integration inactivates GroEL1 and confers a defect in biofilm formation [65].

Interestingly, although *M. tuberculosis* genomes are devoid of full-length prophages, prophages are abundant and diverse in *M. abscessus* genomes [51,66], and the *attB* sites can be easily determined (Fig 4B). Of the 21 *attB* sites, 16 are used by Int-Y and 5 by Int-S (Fig 4B). Interestingly, of the 16 Int-Y *attB* sites, only 12 overlap a host tRNA gene and 4 are either within reading frames or in regulatory sequences. Eight of the *attB* sites used by Int-Y are conserved between *M. abscessus* and *M. smegmatis* (Fig 4). Ten of the *attB* sites identified in *M. smegmatis* are also present in *M. tuberculosis* H37Rv [67].

Host range

The bacterium-phage billion-year-old battle for supremacy has profoundly influenced microbial evolution, sharply honed the variation among bacterial and phage strains, and molded the specificity of phages for their hosts. The expectation is thus that phages isolated on one host may infect other hosts, but that there will be substantial variations in host range. Furthermore, host range is malleable, and if an appropriate nonpermissive strain is tested, host range mutants (HRMs) of the phage can be readily isolated that either expand or switch the phage preferences [68].

Currently, our understanding of the host range of mycobacteriophages is at low resolution, with relatively few *Mycobacterium* strains tested [68–70]. There is a great deal more to be learned by expanding the scale of experimental analyses, and determination of the phage preferences for hundreds of *Mycobacterium* species and strains. However, important insights have been gleaned by determining the phage infection profiles on *M. tuberculosis* strains, *M. abscessus* clinical isolates, and other *M. smegmatis* strains [68,70,71]. A notable finding is that there is a close correlation between the ability of phages to infect both *M. smegmatis* mc²155 and *M. tuberculosis* H37Rv and their cluster/subcluster/singleton designation [68]. The correlation is imperfect in that not all phages within a cluster/subcluster necessarily infect both strains, but there are many groups for which none of the phages infect both strains. Interestingly, a phage isolated in the 1960s (DS6A) only infects strains in the *M. tuberculosis* complex [72].

Early host range studies with *M. tuberculosis* used the common H37Rv lab strain [68]. In general, phages isolated on *M. smegmatis* that efficiently infect *M. tuberculosis* H37Rv also infect other *M. tuberculosis* strains [71]. There are some exceptions though, and phage Muddy efficiently infects only about one-third of the *M. tuberculosis* reference strains tested [71]. However, HRMs of Muddy can be isolated on the "nonpermissive" strains with a single amino acid substitution in the putative tail spike protein that not only efficiently infect the strain from which it was recovered, but also all *M. tuberculosis* (lineage 1 to 4) strains [71]. Interestingly, phage BPs does not efficiently infect *M. tuberculosis* H37Rv, but HRMs can be readily isolated that do, and these also have substitutions in a putative tail spike protein [42,68].

However, none of the HRMs tested infect any of the other *M. tuberculosis* reference strains [71]. Presumably, different HRMs could be isolated on other nonpermissive *M. tuberculosis* isolates.

M. abscessus clinical isolates are at the opposite end of the spectrum to *M. tuberculosis* in their variation in phage infection profiles [70]. In general, *M. abscessus* isolates have either smooth or rough colony morphotypes, with the smooth strains having abundant surface GPLs; relatively few of the *M. smegmatis* phages infect and kill any of the smooth isolates [70]. The phages that infect *M. abscessus* rough colony strains are generally within clusters/subclusters that also infect *M. tuberculosis* [70]. However, there is great variation among *M. abscessus* clinical isolates in their phage infection profiles, even among those that overall are very closely related genetically [70]. It would thus be misleading to test just a single isolate of *M. abscessus* and assume that its phage infection profile was representative of all isolates.

What determines phage host range? There are a multitude of players, but they can be generally grouped into those acting at the cell surface (e.g., receptors) and those that act after the DNA has entered the cell (e.g., restriction). Interestingly, many *Mycobacterium* strains—including *M. abscessus*—are CRISPR-free that eliminates a major potential contributor to host range, but they likely have a variety of restriction–modification (R–M) systems [73–76]. They may also vary in their surface structures that are recognized as receptors by the phages, although few phage receptors in *Mycobacterium* strains have been described. Curiously, GPLs are implicated as potential receptors for phage I3 infection of *M. smegmatis*, and yet their abundance in smooth *M. abscessus* strains is implicated in reduced phage infection [77]. As noted above, several phage HRMs have been described with altered tail spike proteins and these presumably alter receptor recognition by the phages, even though the receptors are not well defined. But some HRMs map in other genes including in portal proteins likely responding to defense mechanisms operating at post-injection phage growth [78].

Prophage-mediated viral defense systems

Interestingly, prophages can be a major source of phage defense systems [57,79,80]. It is no surprise that lysogens are immune to superinfection by the same (and closely related) phages (i.e., homotypic defense), and the mechanisms of repressor-mediated immunity have been long studied. However, prophages can play more complex roles by expressing genes that influence the infection of other, unrelated phages (i.e., heterotypic defense) [79,81]. In some instances, these heterotypic defenses can be very specific for one or only a small number of superinfecting phages, while having no effect on genetic siblings or cousins of those phages. Because *M. abscessus* genomes are so richly endowed with prophages (see Fig 4B), these are prime candidates for influencing the highly varied phage infection profiles [51].

Many different prophage-mediated defense systems have been described, some of which likely act by influencing adsorption or DNA injection for the superinfecting phage, and others that prevent phage infection through abortive infection systems. One group of phages prominently implicated in these systems are the Cluster N mycobacteriophages, which are genomically diverse in a central portion of the genomes—adjacent to the immunity and integration region—and each phage has 3 to 9 genes that are expressed lysogenically [79]. In one example, an *M. smegmatis* lysogen carrying a Charlie prophage expresses a protein, gp32, a predicted membrane protein, and defends against infection by phage Che9c (Cluster I1). Charlie gp32 not only impairs plaque formation but also lysogeny, and gp32 likely acts in an exclusion-like process, preventing DNA injection into the cytoplasm [79]. Che9c is the only known phage to which Charlie gp32 provides defense. In a second example, phage Phrann codes for a putative (p)ppGpp synthetase (gp29) and gp30 which when expressed together from a prophage confer

defense against several unrelated phages including phage Tweety [79]. Expression of Phrann gp29/gp30 strongly inhibits Tweety plaque formation but not Tweety lysogeny. In this case, we propose that Tweety DNA injection results in activation of the (p)ppGpp synthetase, growth arrest, and abortive infection [79].

A counterdefense system in mycobacteriophage Tweety

Interestingly, Tweety encodes for an intriguing counterdefense system, which is not active against the Phrann gp29/gp30 system but is activatable or "tunable" mutationally [79]. Specifically, Tweety defense escape mutants (DEMs) can be readily isolated that overcome the defense (i.e., form plaques efficiently on a strain expressing Phrann gp29/gp30) and have mutations that map in Tweety gene *54* (Fig 1). Tweety gene *54* is unusual in that it codes for a gp54 protein with unique C- and N-terminal regions flanking 40 to 48 copies of a tetrapeptide repeat of which the first 2 amino acids are alanine [79]. The DEMs vary by the number of these tetrapeptide repeats—either more or fewer—suggesting that specific juxtaposition of the C- and N-terminal "domains" are required for it to be tuned to inactivate the Phrann gp29/gp30 system. We propose a model in which Phrann gp29 normally binds to gp30 to prevent enzymatic activity, and that activation of the system occurs by dissociation of gp29 and gp30 in response to some—as yet unknown—aspect of Tweety infection [79]. However, if Tweety gp54 is tuned to Phrann gp29/gp30 (as in the DEMs), Tweety gp54 may lock in the Phrann gp29/gp30 complex, prevent dissociation, and thus prevent defense.

Mycobacteriophage genes toxic to bacterial growth

It is likely that phage genes play important roles not just in host range determination, but in other aspects of viral dynamics, including out-competing other phages in lytic growth. Several studies have shown that a substantial proportion (approximately 20%) of phage genes are toxic to the host cell when expressed alone [82,83]. This toxicity itself is likely immaterial to lytic phage growth, as the cell is going to die anyway, and we suggest it more commonly derives from inactivation of cell processes that other phages depend on to lytically infect the bacterium. Thus, a phage undergoing lytic growth can prevent viral "gate-crashers" by shutting down systems that the interlopers are dependent on. One example is Fruitloop gp52, a small (93 amino acids) protein expressed in phage Fruitloop early lytic growth that interacts directly with the host DivIVA protein and inactivates it [82]. An unrelated phage, Rosebush, requires DivIVA to infect, and is therefore excluded from Fruitloop-infected cells by this mechanism [82]. It seems likely that other genes identified as toxic are involved in similar processes [83].

Phage-based tools for *Mycobacterium* genetics

Mycobacteriophages have facilitated development of several widely used tools for *Mycobacterium* genetics. The slow growth rate and pathogenicity confounded genetic manipulation of *M. tuberculosis* for many years, and the need to exploit phage-based systems to overcome these obstacles was recognized early on by Jacobs and colleagues [84]. The extant tools fall into 2 main groups: those that use phage derivatives to take advantage of efficient DNA injection and those that utilize phage genes for tool development. The first category includes phage-based systems for delivery of transposons, allelic exchange modules, and reporter genes and are useful for many *Mycobacterium* hosts, including *M. smegmatis*, *M. abscessus*, and *M. tuberculosis* [85–92]. An obvious advantage of these systems is that phages introduce DNA into bacterial cells at high frequencies, much greater than is typically achieved by transformation or conjugation. Thus, reporter genes can be introduced into clinical samples for rapid determination of drug susceptibilities [88,89,93], transposons can be delivered for the construction of high-

density mutant libraries [91], and allelic exchange substrates can be delivered for constructing knockout mutants [85,87]. All of these approaches have been facilitated by the construction and manipulation of shuttle phasmids, which are chimeras that replicate as phages in *Mycobacterium*, and as large plasmids in *E. coli* [84,94–96].

Phages have also been used for generalized transduction, as described for phage I3 [97,98], and subsequently for Bxz1 [99] (both in Cluster C), although primarily in *M. smegmatis*. Interestingly, all of the phages known to infect *M. tuberculosis* have *cos*-packaging systems and are unlikely to mediate generalized transduction. A generalized transducing phage for *M. tuberculosis* thus remains a desirable but as yet unachieved goal.

Secondly, several generally applicable tools have been developed from mycobacteriophage genomes. Integration-proficient plasmid vectors have found wide usage, as they facilitate the construction of single-copy recombinants in a wide variety of strains. They simply contain a plasmid vector backbone, a selectable marker, and the integrase gene and *attP* site from a mycobacteriophage; typically, they exclude the phage-encoded RDF, which otherwise would promote plasmid excision. The first such vector was derived from phage L5, which integrates into the *attB*-L5 site (attB-1^{smeg}), a short (approximately 40 bp) sequence overlapping a host tRNA^{gly} gene (Fig 3). Transformation with an L5 integration vector is efficient, the host tRNA is "reconstructed" following integration, and recombinants generally are genetically stable [100]; however, they will be lost in the absence of selection if recombinant gene impairs bacterial growth. Excision can be facilitated if desired by introduction of the phage RDF gene [101]. Integration vectors have been described using at least 6 different *attB* sites, including from phages Bxb1 [64], Tweety [32], and Giles [102] (Fig 4A); a multitude of additional vectors could be constructed to expand the suite of useful and compatible plasmids. Because many of the *attB* sites are conserved between *Mycobacterium* species, most of the vectors can be used in both *M. tuberculosis* and *M. abscessus*, as well as *M. smegmatis* (Fig 4). The vectors can also be used with selectable markers based on phage immunity genes rather than antibiotic resistance, using lytic versions of the phages to select for transformants [47,52].

Recombineering systems are also widely used for *Mycobacterium* genetics [103–106]. These mimic the recombineering systems developed from phage lambda and related systems for *E. coli* [107,108], which do not work efficiently in *Mycobacterium* strains. Plasmids for *Mycobacterium* recombineering use mycobacteriophage-derived recombinases, most notably from phage Che9c [103–106]. Strains can be readily constructed that express Che9c genes *60* and *61*, encoding an exonuclease and a DNA-pairing protein, respectively, that promote recombination of dsDNA substrates introduced by electroporation [103]. These can be used for constructing gene knockouts and other mutants in both slow- and fast-growing *Mycobacterium* strains. Alternatively, single-stranded DNA (ssDNA) recombineering can be achieved using oligonucleotides, which requires only the DNA-pairing protein (Che9c gp60) [104]. Several important adaptations of these systems are noteworthy, including the ORBIT system that combines ssDNA recombineering with the Bxb1 integration system [109], and its application for the engineering of the phages themselves [110]. This latter system, dubbed Bacteriophage Recombineering of Electroporated DNA (BRED) is a useful approach for constructing phage mutants and recombinants [110–112], which can be made more efficient by coupling with CRISPR-mediated counterselection in the CRISPY-BRED configuration [113].

Therapeutic applications of mycobacteriophages

The potential therapeutic use of bacteriophages has been postulated for over 100 years, with numerous reports of successful use, especially in eastern Europe and the former Soviet Union [114]. The notion of antimicrobial action by lytic phages is not surprising, especially for topical

applications in which access to the pathogen is anticipated, but prospects for lung infections, skin diseases, and disseminated infections such as those caused by pathogenic *Mycobacteria* are much less certain. The first report of phages for treating a *Mycobacterium* infection was for a pediatric cystic fibrosis (CF) patient with a disseminated *M. abscessus* infection following a bilateral lung transplant and immunosuppressive therapy to support the new lungs [78]. Resolution of the infection was neither fast nor complete, but there was substantial improvement after about 6 weeks of treatment, with improvement in a large infected node in the liver, reduction in skin nodules, and closure of the sternal transplant wound [78]. There are several important lessons from this case study. First, a considerable search was needed to identify phages that efficiently infect and kill the specific *M. abscessus* isolate (designated GD01). Very few phages have been identified using any strain of *M. abscessus* as a host for isolation, and de novo phage discovery using *M. abscessus* GD01 was mostly unproductive [78]. However, a screen of the large collection of *M. smegmatis* phages identified 2 that did, and a third was developed by isolating an HRM of phage BPs. Second, 2 of the 3 phages are temperate, which do not kill bacteria efficiently (see above), and it was necessary to construct lytic derivatives using BRED engineering to remove all or part of the repressor gene [56,78]. Third, the 3-phage cocktail was used with a view to avoid treatment failure due to the emergence of phage resistance. The cross-resistance profiles for the phages were not known, but the 3 phages are genomically distinct (from 3 different clusters), optimizing the prospects for using different infection mechanisms. Indeed, no resistance was observed in any post-treatment isolate. Fourth, the phages were administered intravenously twice daily (10^9 plaque forming units/ dose) without severe adverse reactions. Fifth, no neutralizing antibody response to the phages was observed, although immunosuppressive drugs were being administered [78].

Broadening the use of phages for treating *Mycobacterium* infections must overcome several limitations. Perhaps most importantly, there is substantial variation in the phage infection profiles among clinical isolates of *M. abscessus*, and thus the phages useful in the first case are not generally useful for other infections [70]. A screen of 82 clinical isolates illustrates this variation, and also shows that the repertoire of potentially therapeutically useful phages is quite limited (6 to 8) [70]. Moreover, these isolates can be categorized as having either smooth or rough colony morphotypes, and relatively few phages infect and kill any of the smooth strains efficiently. Thus, identifying therapeutically useful phages for *M. abscessus* smooth strains is a key priority. Even among rough strains, the phage infection profiles vary considerably among closely related strains [70], and the plasmids and prophages are likely major contributors to the phage infection profiles [51]. About 75% of the rough strains are infected and killed by one or more phages and expanding the repertoire of therapeutically useful phages is also a high priority. Using a broader variety of *M. abscessus* isolates for de novo phage isolation may be helpful, as well as exploring the utility of phages isolated on alternative *Mycobacterium* strains. Because *M. abscessus* genomes are replete with diverse prophages [51], these may be a good source new phages, following spontaneous induction, growth on a permissive host, and engineering to be obligatorily lytic [70].

Screening of over 200 *M. abscessus* clinical isolates facilitated therapeutic interventions in at least 20 additional compassionate use cases of highly drug resistant *Mycobacterium* strains [115]. These revealed a further set of key insights. First, favorable clinical or microbiological outcomes were observed in at least 11 of these cases and are incomplete or inconclusive in 5 others. However, little or no change in clinical outcomes were observed in 4 cases. All of these are confounded by complex circumstances involving other infections and clinical conditions; in any one case the linkage between phage administration and clinical outcome may seem compelling, but the roles of other factors cannot be excluded. For 1 CF patient, phage administration over the course of a year was associated with conversion to sputum culture negative,

re-listing for lung transplantation, and an eventual bilateral lung transplant [116]. In another patient, a highly refractory severe *M. chelonae* skin infection was resolved after IV administration without any other changes in antibiotic treatment [117]. Interestingly, phage resistance leading to treatment failure was not observed in any of these cases, even though 11 were treated with only a single phage. Although it is premature of conclude that phage resistance is not an impediment to therapy of *Mycobacterium* infections, it may be less of a concern than for some other bacterial pathogens. Finally, emergence of a neutralizing antibody response to the phages was not uncommon. In one case, this correlated with loss of sustained clinical improvement [118,119], but in others, it does not appear to have interrupted clinical improvement. Overall, these compassionate use interventions provide helpful findings for constructing robust clinical trials aimed at elucidating safety, efficacy, dosage, routes of administration, and pharmacokinetics.

Several other *Mycobacterium* diseases might be targets for therapeutic phage treatments including Buruli ulcers caused by *M. ulcerans* infections [69,120] and tuberculosis caused by *M. tuberculosis* [71,121–124]. Tuberculosis remains an interesting target but presents a different set of challenges than *M. abscessus* infections. Importantly, there is much more limited strain variation in the phage infection profiles, reflecting the lack of plasmids and prophages in *M. tuberculosis* clinical isolates, such that a small cocktail of phages should be useful for a wide variety of clinical isolates [71]. However, gaining phage access to intracellular bacteria may be a challenge requiring modified delivery systems such as liposomes or "Trojan horse" strategies [125–127], although phage virions can be phagocytosed by macrophages, and transcytosis has been demonstrated [128,129]. However, there could be potential benefits to improving disease and limiting transmission by killing extracellular bacteria, in conjunction with ongoing antibiotic treatments [123]. The prevalence of highly drug resistant strains of *M. tuberculosis*, and the threat of resistance to newly developed anti-TB drugs warrants continued investigation into possible roles for phages in controlling this important human disease.

Conclusions

Our understanding of mycobacteriophages has expanded enormously since they were first described over 70 years ago. However, there is still so much more to learn. We have a high-resolution view of the genomics of phages that infect *M. smegmatis*, but there are hundreds of other *Mycobacterium* species and strains that could be used to discover entirely new sets of phages unrelated to these. We note for example, that most *M. abscessus* prophages are unrelated to *M. smegmatis* phages, and their diversity is at least as great [51]. Preliminary analyses suggest that this is also true for prophages of other *Mycobacterium* strains. We also lack any detailed understanding of the host receptors used for specific recognition by the phages, or more generally what determines host range. It seems likely that there are a multitude of systems contributing to this, including both core bacterial functions and those provided by prophages and plasmids. There are additional genetic tools that could be developed from the phages including in vivo and in vitro packaging systems, vector systems including those using plasmidial prophage origins of replication [45], and phage-based vaccine systems [130]. The therapeutic potential of mycobacteriophages is also in its infancy, and it remains to be seen whether they will prove useful only as treatments of last resort, or whether they are more generally useful. Overall, mycobacteriophages have a rich and promising future.

Acknowledgments

References

1. Rohwer F, Youle M, Maughan H, Hisakawa N. Life in our phage world: A centenial field guide to the earth's most diverse inhabitants. San Diego, CA: Wholon; 2014.

2. Mushegian AR. Are There 10(31) Virus Particles on Earth, or More, or Fewer? J Bacteriol. 2020; 202 (9). Epub 2020/02/20. https://doi.org/10.1128/JB.00052-20 PMID: 32071093; PubMed Central PMCID: PMC7148134.

3. Hendrix RW, Smith MC, Burns RN, Ford ME, Hatfull GF. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. Proc Natl Acad Sci U S A. 1999; 96(5):2192–7. Epub 1999/03/03. https://doi.org/10.1073/pnas.96.5.2192 PMID: 10051617; PubMed Central PMCID: PMC26759.

4. Hendrix RW. Bacteriophages: evolution of the majority. Theor Popul Biol. 2002; 61 (4):471–80. https://doi.org/10.1006/tpbi.2002.1590 PMID: 12167366.

5. Hatfull GF, Hendrix RW. Bacteriophages and their Genomes. Curr Opin Virol. 2011; 1:298–303. https://doi.org/10.1016/j.coviro.2011.06.009 PMID: 22034588

6. Bernheim A, Sorek R. The pan-immune system of bacteria: antiviral defence as a community resource. Nat Rev Microbiol. 2020; 18(2):113–9. Epub 2019/11/07. https://doi.org/10.1038/s41579-019-0278-2 PMID: 31695182.

7. Pawluk A, Davidson AR, Maxwell KL. Anti-CRISPR: discovery, mechanism and function. Nat Rev Microbiol. 2018; 16(1):12–7. Epub 2017/10/25. https://doi.org/10.1038/nrmicro.2017.120 PMID: 29062071.

8. Hendrix RW, Bamford D, Casjens S, Christie G, Duda B, Grimes SN, et al. Bacteriophages. Fields Virology. 6th ed. (Vol. 1). Wolters Kluwer Health Adis (ESP); 2013.

9. Ackermann HW. 5500 Phages examined in the electron microscope. Arch Virol. 2007; 152(2):227–43. https://doi.org/10.1007/s00705-006-0849-1 PMID: 17051420.

10. Russell DA. Sequencing, Assembling, and Finishing Complete Bacteriophage Genomes. Methods Mol Biol. 2018; 1681:109–25. Epub 2017/11/15. https://doi.org/10.1007/978-1-4939-7343-9_9 PMID: 29134591.

11. Dyson ZA, Tucci J, Seviour RJ, Petrovski S. Lysis to Kill: Evaluation of the Lytic Abilities, and Genomics of Nine Bacteriophages Infective for Gordonia spp. and Their Potential Use in Activated Sludge Foam Biocontrol. PLoS ONE. 2015; 10(8):e0134512. https://doi.org/10.1371/journal.pone.0134512 PMID: 26241321; PubMed Central PMCID: PMC4524720.

12. Petrovski S, Seviour RJ, Tillett D. Characterization of the genome of the polyvalent lytic bacteriophage GTE2, which has potential for biocontrol of Gordonia-, Rhodococcus-, and Nocardia-stabilized foams in activated sludge plants. Appl Environ Microbiol. 2011; 77(12):3923–9. https://doi.org/10.1128/AEM.00025-11 PMID: 21498753; PubMed Central PMCID: PMC3131622.

13. Gauthier CH, Abad L, Venbakkam AK, Malnak J, Russell DA, Hatfull GF. DEPhT: a novel approach for efficient prophage discovery and precise extraction. Nucleic Acids Res. 2022. Epub 2022/04/23. https://doi.org/10.1093/nar/gkac273 PMID: 35451479

14. Secor PR, Burgener EB, Kinnersley M, Jennings LK, Roman-Cruz V, Popescu M, et al. Pf Bacteriophage and Their Impact on Pseudomonas Virulence, Mammalian Immunity, and Chronic Infections. Front Immunol. 2020; 11:244. Epub 2020/03/11. https://doi.org/10.3389/fimmu.2020.00244 PMID: 32153575; PubMed Central PMCID: PMC7047154.

15. Engel HW. Mycobacteriophages and phage typing. Ann Microbiol (Paris). 1978; 129 (1):75–90. PMID: 418716

16. Uchiyama J, Mizukami K, Yahara K, Kato SI, Murakami H, Nasukawa T, et al. Genome Sequences of 12 Mycobacteriophages Recovered from Archival Stocks in Japan. Genome Announc. 2018; 6(25). Epub 2018/06/23. https://doi.org/10.1128/genomeA.00472-18 PMID: 29930032; PubMed Central PMCID: PMC6013612.

17. Ujihara T, Uchiyama J, Nasukawa T, Ando H, Murakami H, Ohara N, et al. Recovery of mycobacteriophages from archival stocks stored for approximately 50 years in Japan. Arch Virol. 2018; 163 (7):1915–9. Epub 2018/03/08. https://doi.org/10.1007/s00705-018-3788-8 PMID: 29511830.

18. Hatfull GF, Sarkis GJ. DNA sequence, structure and gene expression of mycobacteriophage L5: a phage system for mycobacterial genetics. Mol Microbiol. 1993; 7(3):395–405. Epub 1993/02/01. https://doi.org/10.1111/j.1365-2958.1993.tb01131.x PMID: 8459766.

19. Doke S. Studies on mycobacteriophages and lysogenic mycobacteria. J Kumamoto Med Soc. 1960; 34:1360–73.

20. Ford ME, Sarkis GJ, Belanger AE, Hendrix RW, Hatfull GF. Genome structure of mycobacteriophage D29: implications for phage evolution. J Mol Biol. 1998; 279(1):143–64. Epub 1998/06/24. S0022-2836(97)91610-7 [pii] https://doi.org/10.1006/jmbi.1997.1610 PMID: 9636706.

21. Ford ME, Stenstrom C, Hendrix RW, Hatfull GF. Mycobacteriophage TM4: genome structure and gene expression. Tuber Lung Dis. 1998; 79(2):63–73. Epub 2000/01/25. https://doi.org/10.1054/tuld.1998.0007 S0962-8479(98)90007-7 [pii]. PMID: 10645443.

22. Mediavilla J, Jain S, Kriakov J, Ford ME, Duda RL, Jacobs WR Jr, et al. Genome organization and characterization of mycobacteriophage Bxb1. Mol Microbiol. 2000; 38(5):955–70. Epub 2000/12/21. mmi2183 [pii]. https://doi.org/10.1046/j.1365-2958.2000.02183.x PMID: 11123671.

23. Pedulla ML, Ford ME, Houtz JM, Karthikeyan T, Wadsworth C, Lewis JA, et al. Origins of highly mosaic mycobacteriophage genomes. Cell. 2003; 113(2):171–82. https://doi.org/10.1016/s0092-8674(03)00233-2 PMID: 12705866.

24. Hatfull GF. Molecular Genetics of Mycobacteriophages. In: Hatfull GF, Jacobs WR Jr, editors. Molecular Genetics of the Mycobacteria. Washington, DC: ASM Press; 2000. p. 37–54.

25. Hatfull GF. Wildy Prize Lecture, 2020–2021: Who wouldn't want to discover a new virus? Microbiology (Reading). 2021; 167(9). Epub 2021/09/02. https://doi.org/10.1099/mic.0.001094 PMID: 34468308.

26. Hatfull GF, Racaniello V. PHIRE and TWiV: Experiences in Bringing Virology to New Audiences. Annu Rev Virol. 2014; 1:37–53. https://doi.org/10.1146/annurev-virology-031413-085449 PMID: 26958714

27. Hanauer DI, Jacobs-Sera D, Pedulla ML, Cresawn SG, Hendrix RW, Hatfull GF. Inquiry learning. Teaching scientific inquiry. Science. 2006; 314(5807):1880–1. https://doi.org/10.1126/science.1136796 PMID: 17185586.

28. Jordan TC, Burnett SH, Carson S, Caruso SM, Clase K, DeJong RJ, et al. A broadly implementable research course in phage discovery and genomics for first-year undergraduate students. MBio. 2014; 5(1):e01051–13. https://doi.org/10.1128/mBio.01051-13 PMID: 24496795.

29. Hanauer DI, Graham MJ, Sea P, Betancur L, Bobrownicki A, Cresawn SG, et al. An inclusive Research Education Community (iREC): Impact of the SEA-PHAGES program on research outcomes and student learning. Proc Natl Acad Sci U S A. 2017; 114(51):13531–6. Epub 2017/12/07. https://doi.org/10.1073/pnas.1718188115 PMID: 29208718; PubMed Central PMCID: PMC5754813.

30. Hanauer DI, Graham MJ, Arnold RJ, Ayuk MA, Balish MF, Beyer AR, et al. Instructional Models for Course-Based Research Experience (CRE) Teaching. CBE Life Sci Educ. 2022; 21(1):ar8. Epub 2022/01/04. https://doi.org/10.1187/cbe.21-03-0057 PMID: 34978921.

31. Hatfull GF. Actinobacteriophages: Genomics, Dynamics, and Applications. Annu Rev Virol. 2020; 7 (1):37–61. Epub 2020/09/30. https://doi.org/10.1146/annurev-virology-122019-070009 PMID: 32991269.

32. Pham TT, Jacobs-Sera D, Pedulla ML, Hendrix RW, Hatfull GF. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. Microbiology. 2007; 153(Pt 8):2711–23. https://doi.org/10.1099/mic.0.2007/008904-0 PMID: 17660435.

33. Hatfull GF, Jacobs-Sera D, Lawrence JG, Pope WH, Russell DA, Ko CC, et al. Comparative Genomic Analysis of 60 Mycobacteriophage Genomes: Genome Clustering, Gene Acquisition, and Gene Size. J Mol Biol. 2010; 397(1):119–43. Epub 2010/01/13. https://doi.org/10.1016/j.jmb.2010.01.011 PMID: 20064525.

34. Hatfull GF, Pedulla ML, Jacobs-Sera D, Cichon PM, Foley A, Ford ME, et al. Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. PLoS Genet. 2006; 2(6):e92. https://doi.org/10.1371/journal.pgen.0020092 PMID: 16789831.

35. Pope WH, Jacobs-Sera D, Russell DA, Peebles CL, Al-Atrache Z, Alcoser TA, et al. Expanding the Diversity of Mycobacteriophages: Insights into Genome Architecture and Evolution. PLoS ONE. 2011; 6(1):e16329. https://doi.org/10.1371/journal.pone.0016329 PMID: 21298013

36. Pope WH, Bowman CA, Russell DA, Jacobs-Sera D, Asai DJ, Cresawn SG, et al. Whole genome comparison of a large collection of mycobacteriophages reveals a continuum of phage genetic diversity. Elife. 2015; 4:e06416. https://doi.org/10.7554/eLife.06416 PMID: 25919952; PubMed Central PMCID: PMC4408529.

37. Pope WH, Mavrich TN, Garlena RA, Guerrero-Bustamante CA, Jacobs-Sera D, Montgomery MT, et al. Bacteriophages of Gordonia spp. Display a Spectrum of Diversity and Genetic Relationships. mBio.

2017;8(4). Epub 2017/08/16. https://doi.org/10.1128/mBio.01069-17 PMID: 28811342; PubMed Central PMCID: PMCPMC5559632

38. Cresawn SG, Bogel M, Day N, Jacobs-Sera D, Hendrix RW, Hatfull GF. Phamerator: a bioinformatic tool for comparative bacteriophage genomics. BMC Bioinformatics. 2011; 12:395. Epub 2011/10/14. https://doi.org/10.1186/1471-2105-12-395 PMID: 21991981; PubMed Central PMCID: PMC3233612.

39. Huson DH. SplitsTree: analyzing and visualizing evolutionary data. Bioinformatics. 1998; 14(1):68–73. https://doi.org/10.1093/bioinformatics/14.1.68 PMID: 9520503.

40. Mageeney C, Pope WH, Harrison M, Moran D, Cross T, Jacobs-Sera D, et al. Mycobacteriophage Marvin: a new singleton phage with an unusual genome organization. J Virol. 2012; 86(9):4762–75. https://doi.org/10.1128/JVI.00075-12 PMID: 22357284; PubMed Central PMCID: PMC3347389.

41. Kelley DS, Lennon CW, Sea P, Belfort M, Novikova O. Mycobacteriophages as Incubators for Intein Dissemination and Evolution. MBio. 2016; 7(5). https://doi.org/10.1128/mBio.01537-16 PMID: 27703073; PubMed Central PMCID: PMC5050341.

42. Sampson T, Broussard GW, Marinelli LJ, Jacobs-Sera D, Ray M, Ko CC, et al. Mycobacteriophages BPs, Angel and Halo: comparative genomics reveals a novel class of ultra-small mobile genetic elements. Microbiology. 2009; 155(Pt 9):2962–77. https://doi.org/10.1099/mic.0.030486-0 PMID: 19556295.

43. Martinsohn JT, Radman M, Petit MA. The lambda red proteins promote efficient recombination between diverged sequences: implications for bacteriophage genome mosaicism. PLoS Genet. 2008; 4(5):e1000065. https://doi.org/10.1371/journal.pgen.1000065 PMID: 18451987.

44. Dedrick RM, Mavrich TN, Ng WL, Hatfull GF. Expression and evolutionary patterns of mycobacteriophage D29 and its temperate close relatives. BMC Microbiol. 2017; 17(1):225. Epub 2017/12/05. https://doi.org/10.1186/s12866-017-1131-2 PMID: 29197343; PubMed Central PMCID: PMC5712189.

45. Wetzel KS, Aull HG, Zack KM, Garlena RA, Hatfull GF. Protein-Mediated and RNA-Based Origins of Replication of Extrachromosomal Mycobacterial Prophages. MBio. 2020; 11(2). Epub 2020/03/27. https://doi.org/10.1128/mBio.00385-20 PMID: 32209683; PubMed Central PMCID: PMC7157519.

46. Broussard GW, Oldfield LM, Villanueva VM, Lunt BL, Shine EE, Hatfull GF. Integration-dependent bacteriophage immunity provides insights into the evolution of genetic switches. Mol Cell. 2013; 49 (2):237–48. https://doi.org/10.1016/j.molcel.2012.11.012 PMID: 23246436; PubMed Central PMCID: PMC3557535.

47. Petrova ZO, Broussard GW, Hatfull GF. Mycobacteriophage-repressor-mediated immunity as a selectable genetic marker: Adephagia and BPs repressor selection. Microbiology. 2015; 161(8):1539–51. https://doi.org/10.1099/mic.0.000120 PMID: 26066798; PubMed Central PMCID: PMC4681040.

48. Court DL, Oppenheim AB, Adhya SL. A new look at bacteriophage lambda genetic networks. J Bacteriol. 2007; 189(2):298–304. https://doi.org/10.1128/JB.01215-06 PMID: 17085553.

49. Little JW. Evolution of complex gene regulatory circuits by addition of refinements. Curr Biol. 2010; 20 (17):R724–34. Epub 2010/09/14. https://doi.org/10.1016/j.cub.2010.06.028 PMID: 20833317.

50. Lewis JA, Hatfull GF. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. Nucleic Acids Res. 2001; 29(11):2205–16. Epub 2001/05/29. https://doi.org/10.1093/nar/29.11.2205 PMID: 11376138; PubMed Central PMCID: PMC55702.

51. Dedrick RM, Aull HG, Jacobs-Sera D, Garlena RA, Russell DA, Smith BE, et al. The Prophage and Plasmid Mobilome as a Likely Driver of Mycobacterium abscessus Diversity. MBio. 2021; 12(2). Epub 2021/04/01. https://doi.org/10.1128/mBio.03441-20 PMID: 33785627; PubMed Central PMCID: PMC8092301.

52. Donnelly-Wu MK, Jacobs WR Jr, Hatfull GF. Superinfection immunity of mycobacteriophage L5: applications for genetic transformation of mycobacteria. Mol Microbiol. 1993; 7(3):407–17. https://doi.org/10.1111/j.1365-2958.1993.tb01132.x PMID: 8459767.

53. Jain S, Hatfull GF. Transcriptional regulation and immunity in mycobacteriophage Bxb1. Mol Microbiol. 2000; 38(5):971–85. https://doi.org/10.1046/j.1365-2958.2000.02184.x PMID: 11123672.

54. Mavrich TN, Hatfull GF. Evolution of Superinfection Immunity in Cluster A Mycobacteriophages. MBio. 2019; 10(3). Epub 2019/06/06. https://doi.org/10.1128/mBio.00971-19 PMID: 31164468; PubMed Central PMCID: PMC6550527.

55. Brown KL, Sarkis GJ, Wadsworth C, Hatfull GF. Transcriptional silencing by the mycobacteriophage L5 repressor. EMBO J. 1997; 16(19):5914–21. Epub 1997/10/06. https://doi.org/10.1093/emboj/16.19.5914 PubMed Central PMCID: PMC1170222. PMID: 9312049

56. Dedrick RM, Guerrero Bustamante CA, Garlena RA, Pinches RS, Cornely K, Hatfull GF. Mycobacteriophage ZoeJ: A broad host-range close relative of mycobacteriophage TM4. Tuberculosis (Edinb).

2019; 115:14–23. Epub 2019/04/06. https://doi.org/10.1016/j.tube.2019.01.002 PMID: 30948168; PubMed Central PMCID: PMC6452893.

57. Gentile GM, Wetzel KS, Dedrick RM, Montgomery MT, Garlena RA, Jacobs-Sera D, et al. More Evidence of Collusion: a New Prophage-Mediated Viral Defense System Encoded by Mycobacteriophage Sbash. MBio. 2019; 10(2). Epub 2019/03/21. https://doi.org/10.1128/mBio.00196-19 PMID: 30890613.

58. Dedrick RM, Marinelli LJ, Newton GL, Pogliano K, Pogliano J, Hatfull GF. Functional requirements for bacteriophage growth: gene essentiality and expression in mycobacteriophage Giles. Mol Microbiol. 2013; 88(3):577–89. https://doi.org/10.1111/mmi.12210 PMID: 23560716; PubMed Central PMCID: PMC3641587.

59. Nesbit CE, Levin ME, Donnelly-Wu MK, Hatfull GF. Transcriptional regulation of repressor synthesis in mycobacteriophage L5. Mol Microbiol. 1995; 17(6):1045–56. Epub 1995/09/01. https://doi.org/10.1111/j.1365-2958.1995.mmi_17061045.x PMID: 8594325.

60. Ghosh P, Kim AI, Hatfull GF. The orientation of mycobacteriophage Bxb1 integration is solely dependent on the central dinucleotide of attP and attB. Mol Cell. 2003; 12(5):1101–11. https://doi.org/10.1016/s1097-2765(03)00444-1 PMID: 14636570.

61. Singh S, Rockenbach K, Dedrick RM, Vandemark AP, Hatfull GF. Cross-talk between Diverse Serine Integrases. J Mol Biol. 2014; 426(2):318–31. https://doi.org/10.1016/j.jmb.2013.10.013 PMID: 24161951.

62. Bibb LA, Hancox MI, Hatfull GF. Integration and excision by the large serine recombinase phiRv1 integrase. Mol Microbiol. 2005; 55(6):1896–910. https://doi.org/10.1111/j.1365-2958.2005.04517.x PMID: 15752208.

63. Bibb LA, Hatfull GF. Integration and excision of the Mycobacterium tuberculosis prophage-like element, phiRv1. Mol Microbiol. 2002; 45(6):1515–26. https://doi.org/10.1046/j.1365-2958.2002.03130.x PMID: 12354222.

64. Kim AI, Ghosh P, Aaron MA, Bibb LA, Jain S, Hatfull GF. Mycobacteriophage Bxb1 integrates into the Mycobacterium smegmatis groEL1 gene. Mol Microbiol. 2003; 50(2):463–73. https://doi.org/10.1046/j.1365-2958.2003.03723.x PMID: 14617171.

65. Ojha A, Anand M, Bhatt A, Kremer L, Jacobs WR Jr, Hatfull GF. GroEL1: a dedicated chaperone involved in mycolic acid biosynthesis during biofilm formation in mycobacteria. Cell. 2005; 123(5):861–73. https://doi.org/10.1016/j.cell.2005.09.012 PMID: 16325580.

66. Glickman C, Kammlade SM, Hasan NA, Epperson LE, Davidson RM, Strong M. Characterization of integrated prophages within diverse species of clinical nontuberculous mycobacteria. Virol J. 2020; 17 (1):124. Epub 2020/08/19. https://doi.org/10.1186/s12985-020-01394-y PMID: 32807206; PubMed Central PMCID: PMC7433156.

67. Hatfull GF. The secret lives of mycobacteriophages. Adv Virus Res. 2012; 82:179–288. Epub 2012/03/17. B978-0-12-394621-8.00015–7 [pii] https://doi.org/10.1016/B978-0-12-394621-8.00015-7 PMID: 22420855.

68. Jacobs-Sera D, Marinelli LJ, Bowman C, Broussard GW, Guerrero Bustamante C, Boyle MM, et al. On the nature of mycobacteriophage diversity and host preference. Virology. 2012; 434(2):187–201. https://doi.org/10.1016/j.virol.2012.09.026 PMID: 23084079; PubMed Central PMCID: PMC3518647.

69. Rybniker J, Kramme S, Small PL. Host range of 14 mycobacteriophages in Mycobacterium ulcerans and seven other mycobacteria including Mycobacterium tuberculosis—application for identification and susceptibility testing. J Med Microbiol. 2006; 55(Pt 1):37–42. https://doi.org/10.1099/jmm.0.46238-0 PMID: 16388028.

70. Dedrick RM, Smith BE, Garlena RA, Russell DA, Aull HG, Mahalingam V, et al. Mycobacterium abscessus Strain Morphotype Determines Phage Susceptibility, the Repertoire of Therapeutically Useful Phages, and Phage Resistance. MBio. 2021; 12(2). Epub 2021/04/01. https://doi.org/10.1128/mBio.03431-20 PMID: 33785625; PubMed Central PMCID: PMC8092298.

71. Guerrero-Bustamante CA, Dedrick RM, Garlena RA, Russell DA, Hatfull GF. Toward a Phage Cocktail for Tuberculosis: Susceptibility and Tuberculocidal Action of Mycobacteriophages against Diverse Mycobacterium tuberculosis Strains. MBio. 2021; 12(3). Epub 2021/05/22. https://doi.org/10.1128/mBio.00973-21 PMID: 34016711; PubMed Central PMCID: PMC8263002.

72. Redmond WB, Carter JC. A bacteriophage specific to Mycobacterium tuberculosis varieties hominis and bovis. Am Rv Respir Dis. 1960; 82:781–6. https://doi.org/10.1164/arrd.1960.82.6.781 PMID: 13740230

73. Shankar S, Tyagi AK. MhaAI, a novel isoschizomer of PstI from Mycobacterium habana recognizing 5'-CTGCA/G-3'. Nucleic Acids Res. 1992; 20 (11):2891. https://doi.org/10.1093/nar/20.11.2891 PMID: 1614878

74. Shankar S, Tyagi AK. MfoAI, a novel isoschizomer of HaeIII from Mycobacterium fortuitum recognizing 5'-GG/CC-3'. Nucleic Acids Res. 1992; 20(11):2890. https://doi.org/10.1093/nar/20.11.2890 PMID: 1614877

75. Shankar S, Tyagi AK. MchAI and MchAII, two class-II restriction endonucleases from Mycobacterium chelonei. Gene. 1993; 132(1):119–23. https://doi.org/10.1016/0378-1119(93)90523-6 PMID: 8406034

76. Shankar S, Tyagi AK. Purification and characterization of restriction endonuclease MgoI from Mycobacterium gordonae. Gene. 1993; 131 (1):153–4. https://doi.org/10.1016/0378-1119(93)90686-w PMID: 8370536

77. Chen J, Kriakov J, Singh A, Jacobs WR Jr, Besra GS, Bhatt A. Defects in glycopeptidolipid biosynthesis confer phage I3 resistance in Mycobacterium smegmatis. Microbiology. 2009; 155(Pt 12):4050–7. https://doi.org/10.1099/mic.0.033209-0 PMID: 19744987.

78. Dedrick R, Guerrero Bustamante C, Garlena RA, Russell DA, Ford K, Harris K, et al. Engineered bacteriophages for treatment of a patient with a disseminated drug-resistant *Mycobacterium abscessus*. Nat Med. 2019; 25:730–3. https://doi.org/10.1038/s41591-019-0437-z PMID: 31068712

79. Dedrick RM, Jacobs-Sera D, Bustamante CA, Garlena RA, Mavrich TN, Pope WH, et al. Prophage-mediated defence against viral attack and viral counter-defence. Nat Microbiol. 2017; 2:16251. https://doi.org/10.1038/nmicrobiol.2016.251 PMID: 28067906; PubMed Central PMCID: PMC5508108.

80. Montgomery MT, Guerrero Bustamante CA, Dedrick RM, Jacobs-Sera D, Hatfull GF. Yet More Evidence of Collusion: a New Viral Defense System Encoded by Gordonia Phage CarolAnn. MBio. 2019; 10(2). Epub 2019/03/21. https://doi.org/10.1128/mBio.02417-18 PMID: 30890601; PubMed Central PMCID: PMC6426606.

81. Bondy-Denomy J, Qian J, Westra ER, Buckling A, Guttman DS, Davidson AR, et al. Prophages mediate defense against phage infection through diverse mechanisms. ISME J. 2016; 10(12):2854–66. https://doi.org/10.1038/ismej.2016.79 PMID: 27258950; PubMed Central PMCID: PMC5148200.

82. Ko CC, Hatfull GF. Mycobacteriophage Fruitloop gp52 inactivates Wag31 (DivIVA) to prevent heterotypic superinfection. Mol Microbiol. 2018. Epub 2018/03/01. https://doi.org/10.1111/mmi.13946 PMID: 29488662.

83. Ko CC, Hatfull GF. Identification of mycobacteriophage toxic genes reveals new features of mycobacterial physiology and morphology. Sci Rep. 2020; 10(1):14670. Epub 2020/09/06. https://doi.org/10.1038/s41598-020-71588-5 PMID: 32887931; PubMed Central PMCID: PMC7474061.

84. Jacobs WR Jr, Tuckman M, Bloom BR. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. Nature. 1987; 327(6122):532–5. https://doi.org/10.1038/327532a0 PMID: 3473289

85. Bardarov S, Bardarov S Jr, Pavelka MS, Sambandamurthy V, Larsen M, Tufariello J, et al. Specialized transduction: an efficient method for generating marked and unmarked targeted gene disruptions in Mycobacterium tuberculosis, M. bovis BCG and M. smegmatis. Microbiology. 2002; 148(Pt 10):3007–17. https://doi.org/10.1099/00221287-148-10-3007 PMID: 12368434.

86. Jain P, Hsu T, Arai M, Biermann K, Thaler DS, Nguyen A, et al. Specialized transduction designed for precise high-throughput unmarked deletions in Mycobacterium tuberculosis. MBio. 2014; 5(3): e01245–14. Epub 2014/06/05. https://doi.org/10.1128/mBio.01245-14 PMID: 24895308; PubMed Central PMCID: PMC4049104.

87. Tufariello JM, Malek AA, Vilcheze C, Cole LE, Ratner HK, Gonzalez PA, et al. Enhanced specialized transduction using recombineering in Mycobacterium tuberculosis. MBio. 2014; 5(3):e01179–14. Epub 2014/05/29. https://doi.org/10.1128/mBio.01179-14 PMID: 24865558; PubMed Central PMCID: PMC4045075.

88. Jacobs WR Jr, Barletta RG, Udani R, Chan J, Kalkut G, Sosne G, et al. Rapid assessment of drug susceptibilities of Mycobacterium tuberculosis by means of luciferase reporter phages. Science. 1993; 260(5109):819–22. Epub 1993/05/07. https://doi.org/10.1126/science.8484123 PMID: 8484123.

89. Jain P, Hartman TE, Eisenberg N, O'Donnell MR, Kriakov J, Govender K, et al. phi(2)GFP10, a high-intensity fluorophage, enables detection and rapid drug susceptibility testing of Mycobacterium tuberculosis directly from sputum samples. J Clin Microbiol. 2012; 50(4):1362–9. Epub 2012/01/27. JCM.06192-11 [pii] https://doi.org/10.1128/JCM.06192-11 PMID: 22278833; PubMed Central PMCID: PMC3318544.

90. Piuri M, Jacobs WR Jr, Hatfull GF. Fluoromycobacteriophages for rapid, specific, and sensitive antibiotic susceptibility testing of Mycobacterium tuberculosis. PLoS ONE. 2009; 4(3):e4870. https://doi.org/10.1371/journal.pone.0004870 PMID: 19300517.

91. Bardarov S, Kriakov J, Carriere C, Yu S, Vaamonde C, McAdam RA, et al. Conditionally replicating mycobacteriophages: a system for transposon delivery to Mycobacterium tuberculosis. Proc Natl Acad Sci U S A. 1997; 94(20):10961–6. Epub 1997/10/06. https://doi.org/10.1073/pnas.94.20.10961 PMID: 9380742; PubMed Central PMCID: PMC23545.

92. Sassetti CM, Boyd DH, Rubin EJ. Genes required for mycobacterial growth defined by high density mutagenesis. Mol Microbiol. 2003; 48(1):77–84. https://doi.org/10.1046/j.1365-2958.2003.03425.x PMID: 12657046.

93. O'Donnell MR, Pym A, Jain P, Munsamy V, Wolf A, Karim F, et al. A Novel Reporter Phage To Detect Tuberculosis and Rifampin Resistance in a High-HIV-Burden Population. J Clin Microbiol. 2015; 53 (7):2188–94. Epub 2015/05/01. https://doi.org/10.1128/JCM.03530-14 PMID: 25926493; PubMed Central PMCID: PMC4473227.

94. Jacobs WR Jr. *Mycobacterium tuberculosis*: A Once Genetically Intractable Organism. In: Hatfull GF, Jacobs WR Jr, editors. Molecular Genetics of the Mycobacteria. Washington, DC: ASM Press; 2000. p. 1–16.

95. Jacobs WR Jr, Snapper SB, Lugosi L, Jekkel A, Melton RE, Kieser T, et al. Development of genetic systems for the mycobacteria. Acta Leprol. 1989; 7(Suppl 1):203–7. PMID: 2503991

96. Jacobs WR Jr, Snapper SB, Tuckman M, Bloom BR. Mycobacteriophage vector systems. Rev Infect Dis. 1989; 11(Suppl 2):S404–10. https://doi.org/10.1093/clinids/11.supplement_2.s404 PMID: 2652256

97. Ramakrishnan T, Shaila MS. Interfamilial transfer of amber suppressor gene for the isolation of amber mutants of Mycobacteriophage I3. Arch Microbiol. 1979; 120(3):301–2. https://doi.org/10.1007/BF00423080 PMID: 443992

98. Raj CV, Ramakrishnan T. Transduction in Mycobacterium smegmatis. Nature. 1970; 228(268):280–1. https://doi.org/10.1038/228280b0 PMID: 5479524

99. Lee S, Kriakov J, Vilcheze C, Dai Z, Hatfull GF, Jacobs WR Jr. Bxz1, a new generalized transducing phage for mycobacteria. FEMS Microbiol Lett. 2004; 241(2):271–6. https://doi.org/10.1016/j.femsle.2004.10.032 PMID: 15598543.

100. Lee MH, Pascopella L, Jacobs WR Jr, Hatfull GF. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for Mycobacterium smegmatis, Mycobacterium tuberculosis, and bacille Calmette-Guerin. Proc Natl Acad Sci U S A. 1991; 88(8):3111–5. Epub 1991/04/25. https://doi.org/10.1073/pnas.88.8.3111 PMID: 1901654; PubMed Central PMCID: PMC51395.

101. Parish T, Lewis J, Stoker NG. Use of the mycobacteriophage L5 excisionase in *Mycobacterium tuberculosis* to demonstrate gene essentiality. Tuberculosis (Edinb). 2001; 81(5–6):359–64. https://doi.org/10.1054/tube.2001.0312 PMID: 11800587.

102. Morris P, Marinelli LJ, Jacobs-Sera D, Hendrix RW, Hatfull GF. Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. J Bacteriol. 2008; 190(6):2172–82. https://doi.org/10.1128/JB.01657-07 PMID: 18178732.

103. van Kessel JC, Hatfull GF. Recombineering in *Mycobacterium tuberculosis*. Nat Methods. 2007; 4 (2):147–52. https://doi.org/10.1038/nmeth996 PMID: 17179933.

104. van Kessel JC, Hatfull GF. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. Mol Microbiol. 2008; 67(5):1094–107. https://doi.org/10.1111/j.1365-2958.2008.06109.x PMID: 18221264.

105. van Kessel JC, Hatfull GF. Mycobacterial recombineering. Methods Mol Biol. 2008; 435:203–15. https://doi.org/10.1007/978-1-59745-232-8_15 PMID: 18370078.

106. van Kessel JC, Marinelli LJ, Hatfull GF. Recombineering mycobacteria and their phages. Nat Rev Microbiol. 2008; 6(11):851–7. https://doi.org/10.1038/nrmicro2014 PMID: 18923412.

107. Court DL, Sawitzke JA, Thomason LC. Genetic engineering using homologous recombination. Annu Rev Genet. 2002; 36:361–88. https://doi.org/10.1146/annurev.genet.36.061102.093104 PMID: 12429697.

108. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000; 97(12):6640–5. https://doi.org/10.1073/pnas.120163297 PMID: 10829079.

109. Murphy KC, Nelson SJ, Nambi S, Papavinasasundaram K, Baer CE, Sassetti CM. ORBIT: a New Paradigm for Genetic Engineering of Mycobacterial Chromosomes. MBio. 2018; 9(6). Epub 2018/12/13. https://doi.org/10.1128/mBio.01467-18 PMID: 30538179; PubMed Central PMCID: PMC6299477.

110. Marinelli LJ, Piuri M, Swigonova Z, Balachandran A, Oldfield LM, van Kessel JC, et al. BRED: a simple and powerful tool for constructing mutant and recombinant bacteriophage genomes. PLoS ONE. 2008; 3(12):e3957. https://doi.org/10.1371/journal.pone.0003957 PMID: 19088849.

111. Marinelli LJ, Hatfull GF, Piuri M. Recombineering: A powerful tool for modification of bacteriophage genomes. Bacteriophage. 2012; 2(1):5–14. Epub 2012/06/06. https://doi.org/10.4161/bact.18778 2011BACTERIOPHAGE0041R [pii]. PMID: 22666652; PubMed Central PMCID: PMC3357384.

112. Marinelli LJ, Piuri M, Hatfull GF. Genetic Manipulation of Lytic Bacteriophages with BRED: Bacteriophage Recombineering of Electroporated DNA. Methods Mol Biol. 2019; 1898:69–80. Epub 2018/12/21. https://doi.org/10.1007/978-1-4939-8940-9_6 PMID: 30570724.

113. Wetzel KS, Guerrero-Bustamante CA, Dedrick RM, Ko CC, Freeman KG, Aull HG, et al. CRISPY-BRED and CRISPY-BRIP: efficient bacteriophage engineering. Sci Rep. 2021; 11(1):6796. Epub 2021/03/26. https://doi.org/10.1038/s41598-021-86112-6 PMID: 33762639; PubMed Central PMCID: PMC7990910.

114. Housby JN, Mann NH. Phage therapy. Drug Discov Today. 2009; 14(11–12):536–40. https://doi.org/10.1016/j.drudis.2009.03.006 PMID: 19508915.

115. Dedrick RM, Smith BE, Cristinziano M, Freeman KG, Jacobs-Sera D, Belessis Y, et al. Phage Therapy of Mycobacterium Infections: Compassionate-use of Phages in Twenty Patients with Drug-Resistant Mycobacterial Disease. Clin Infect Dis. 2022; Forthcoming.

116. Nick JA, Dedrick RM, Gray AL, Vladar EK, Smith BE, Freeman KG, et al. Host and Pathogen Response to Bacteriophage Engineered Against Mycobacterium abscessus Lung Infection. Cell. 2022; Forthcoming. https://doi.org/10.1016/j.cell.2022.04.024 PMID: 35568033

117. Little JS, Dedrick RM, Freeman KG, Cristinziano M, Smith BE, Benson CA, et al. Bacteriophage treatment of disseminated cutaneous Mycobacterium chelonae infection. Nat Commun. 2022; 13(1):2313. Epub 2022/05/04. https://doi.org/10.1038/s41467-022-29689-4 PMID: 35504908; PubMed Central PMCID: PMC9064978.

118. Dedrick RM, Freeman KG, Nguyen JA, Bahadirli-Talbott A, Smith BE, Wu AE, et al. Potent antibody-mediated neutralization limits bacteriophage treatment of a pulmonary Mycobacterium abscessus infection. Nat Med. 2021; 27(8):1357–61. Epub 2021/07/10. https://doi.org/10.1038/s41591-021-01403-9 PMID: 34239133.

119. Dedrick RM, Freeman KG, Nguyen JA, Bahadirli-Talbott A, Cardin ME, Cristinziano M, et al. Nebulized bacteriophage in a patient with refractory Mycobacterium abscessus lung disease. Open Forum Infect Dis. 2022; Forthcoming.

120. Trigo G, Martins TG, Fraga AG, Longatto-Filho A, Castro AG, Azeredo J, et al. Phage therapy is effective against infection by Mycobacterium ulcerans in a murine footpad model. PLoS Negl Trop Dis. 2013; 7(4):e2183. Epub 2013/05/03. https://doi.org/10.1371/journal.pntd.0002183 PMID: 23638204; PubMed Central PMCID: PMC3636042.

121. Hatfull GF. Mycobacteriophages and tuberculosis. In: Eisenach K, Cole ST, Jacobs WR Jr, McMurray D, editors. Tuberculosis. Washington, DC: ASM Press; 2004. p. 203–18.

122. Hatfull GF. Mycobacteriophages: windows into tuberculosis. PLoS Pathog. 2014; 10(3):e1003953. https://doi.org/10.1371/journal.ppat.1003953 PMID: 24651299; PubMed Central PMCID: PMC3961340.

123. Diacon AH, Guerrero-Bustamante CA, Rosenkranz B, Rubio Pomar FJ, Vanker N, Hatfull GF. Mycobacteriophages to Treat Tuberculosis: Dream or Delusion? Respiration. 2022; 101(1):1–15. Epub 2021/11/24. https://doi.org/10.1159/000519870 PMID: 34814151.

124. Shield CG, Swift BMC, McHugh TD, Dedrick RM, Hatfull GF, Satta G. Application of Bacteriophages for Mycobacterial Infections, from Diagnosis to Treatment. Microorganisms. 2021; 9(11). Epub 2021/11/28. https://doi.org/10.3390/microorganisms9112366 PMID: 34835491; PubMed Central PMCID: PMC8617706.

125. Broxmeyer L, Sosnowska D, Miltner E, Chacon O, Wagner D, McGarvey J, et al. Killing of Mycobacterium avium and Mycobacterium tuberculosis by a mycobacteriophage delivered by a nonvirulent mycobacterium: a model for phage therapy of intracellular bacterial pathogens. J Infect Dis. 2002; 186(8):1155–60. https://doi.org/10.1086/343812 PMID: 12355367.

126. Danelishvili L, Young LS, Bermudez LE. In vivo efficacy of phage therapy for Mycobacterium avium infection as delivered by a nonvirulent mycobacterium. Microb Drug Resist. 2006; 12(1):1–6. https://doi.org/10.1089/mdr.2006.12.1 PMID: 16584300.

127. Foley-Thomas EM, Whipple DL, Bermudez LE, Barletta RG. Phage infection, transfection and transformation of Mycobacterium avium complex and Mycobacterium paratuberculosis. Microbiology. 1995; 141(Pt 5):1173–81. https://doi.org/10.1099/13500872-141-5-1173 PMID: 7773411

128. Gorski A, Miedzybrodzki R, Borysowski J, Dabrowska K, Wierzbicki P, Ohams M, et al. Phage as a modulator of immune responses: practical implications for phage therapy. Adv Virus Res. 2012; 83:41–71. Epub 2012/07/04. https://doi.org/10.1016/B978-0-12-394438-2.00002-5 PMID: 22748808.

129. Nguyen S, Baker K, Padman BS, Patwa R, Dunstan RA, Weston TA, et al. Bacteriophage Transcytosis Provides a Mechanism To Cross Epithelial Cell Layers. MBio. 2017; 8(6). Epub 2017/11/23. https://doi.org/10.1128/mBio.01874-17 PMID: 29162715; PubMed Central PMCID: PMC5698557.

130. Freeman KG, Wetzel KS, Zhang Y, Zack KM, Jacobs-Sera D, Walters SM, et al. A Mycobacteriophage-Based Vaccine Platform: SARS-CoV-2 Antigen Expression and Display. Microorganisms. 2021; 9 (12). Epub 2021/12/25. https://doi.org/10.3390/microorganisms9122414 PMID: 34946016; PubMed Central PMCID: PMC8704799.