INNOVATION

Recombineering mycobacteria and their phages

Julia C. van Kessel, Laura J. Marinelli and Graham F. Hatfull

Abstract | Bacteriophages are central components in the development of molecular tools for microbial genetics. Mycobacteriophages have proven to be a rich resource for tuberculosis genetics, and the recent development of a mycobacterial recombineering system based on mycobacteriophage Che9c-encoded proteins offers new approaches to mycobacterial mutagenesis. Expression of the phage exonuclease and recombinase substantially enhances recombination frequencies in both fast- and slow-growing mycobacteria, thereby facilitating construction of both gene knockout and point mutants; it also provides a simple and efficient method for constructing mycobacteriophage mutants. Exploitation of host-specific phages thus provides a general strategy for recombineering and mutagenesis in genetically naive systems.

Genetic manipulation of *Mycobacterium tuberculosis* is complicated by its pathogenesis, slow growth rate, inefficient DNA uptake and high level of illegitimate recombination. Construction of gene knockout mutants can be achieved using plasmid-based systems and by phage-mediated specialized transduction, but the simple introduction of a linear double-stranded DNA (dsDNA) substrate by electroporation leads to a high number of ectopic integration events regardless of homology between the targeting substrate and the bacterial chromosome[1,2]. Other types of mutagenesis, such as the construction of defined isogenic strains with single point mutations, are further complicated by the lack of generalized transducing phages that infect *M. tuberculosis*[3].

A range of techniques for constructing gene knockout or replacement mutants have been described that are designed to overcome the high level of illegitimate recombination (compared with homologous recombination) that is observed with linear dsDNA allelic exchange substrates (AESs). These include the use of long, linear dsDNA substrates[4], as well as various plasmid- and cosmid-based strategies[4–8], and these are often coupled with the use of counter-selectable genetic markers, such as *sacB*[9,10]. Specialized transducing phages that are based on conditionally replicating shuttle phasmids represent an alternative and highly effective approach to mutagenesis[11], including the construction of isogenic *M. tuberculosis* strains that differ by a single known point mutation[12]. Although most of these methods are effective, they are constrained by the requirement for complex genetic constructions and/or multiple steps of manipulation and screening. Given the slow growth rate of *M. tuberculosis* (doubling time of 24 hours), alternative methods that simplify its genetic manipulation are highly desirable.

Recombineering — genetic engineering using recombination proteins[13] — is a powerful system for mutagenesis in *Escherichia coli* (as well as species of *Salmonella* and *Shigella*[14,15]), in which recombination systems encoded by the λ Red system or the *recE* and *recT* genes of the Rac prophage greatly enhance the frequencies of homologous recombination. These high recombination efficiencies can be exploited in various ways, including the construction of chromosomal gene knockouts, point mutations, deletions, small insertions, *in vivo* cloning, mutagenesis of bacterial artificial chromosomes and genomic libraries[13,16–19]. Only short segments (50 bp) of DNA homology are required, and mutagenic substrates constructed by PCR can be readily introduced by electroporation into *E. coli* strains that express the phage recombination proteins.

Recombineering in *E. coli* using the λ Red system involves three phage-encoded proteins, Exo, Beta and Gam, whereas the Rac prophage encodes only RecE and RecT, which are functionally equivalent to λ Exo and Beta, respectively[20,21] (FIG. 1a). Exo and RecE are 5′→3′ exonucleases that degrade a single strand of a linear dsDNA substrate[22,23], thereby exposing a 3′ single-stranded DNA (ssDNA) tail to which Beta or RecT can bind[24,25]. Recombination is then mediated by association of these complexes with resident chromosomal or plasmid targets through the pairing of complementary sequences, strand exchange or strand invasion[26–29]. λ Beta and RecT proteins are members of a large family of ssDNA annealing proteins (SSAPs)[30]; two other families of SSAPs that function similarly are defined by *Salmonella* phage P22 (P22) Arf and the eukaryotic protein, Rad52. Although the RecE and RecT and λ Red systems confer similar levels of recombination in *E. coli*, the Beta and RecT proteins have a strong preference for their cognate Exo and RecE proteins, and it is likely that they function as protein complexes[31,32]. *E. coli* dsDNA recombineering requires both an exonuclease and its associated SSAP, whereas recombination using ssDNA substrates requires only the SSAP. The ability to recombineer with short, oligonucleotide-derived ssDNA substrates is especially useful[19], and frequencies can be sufficiently high to enable identification of mutants without the need for direct genetic selection, especially when using strains that are defective in mismatch repair[33]. The third protein in the λ system, Gam, increases recombineering frequencies by binding to the RecB subunit of RecBCD and inhibiting the degradation of dsDNA substrates[18,19,34,35]. Although other Gam functional analogues have been characterized, such as Abc2 of P22 (REF. 36), these proteins are rare compared with the large superfamilies of SSAPs and their associated exonucleases[30].

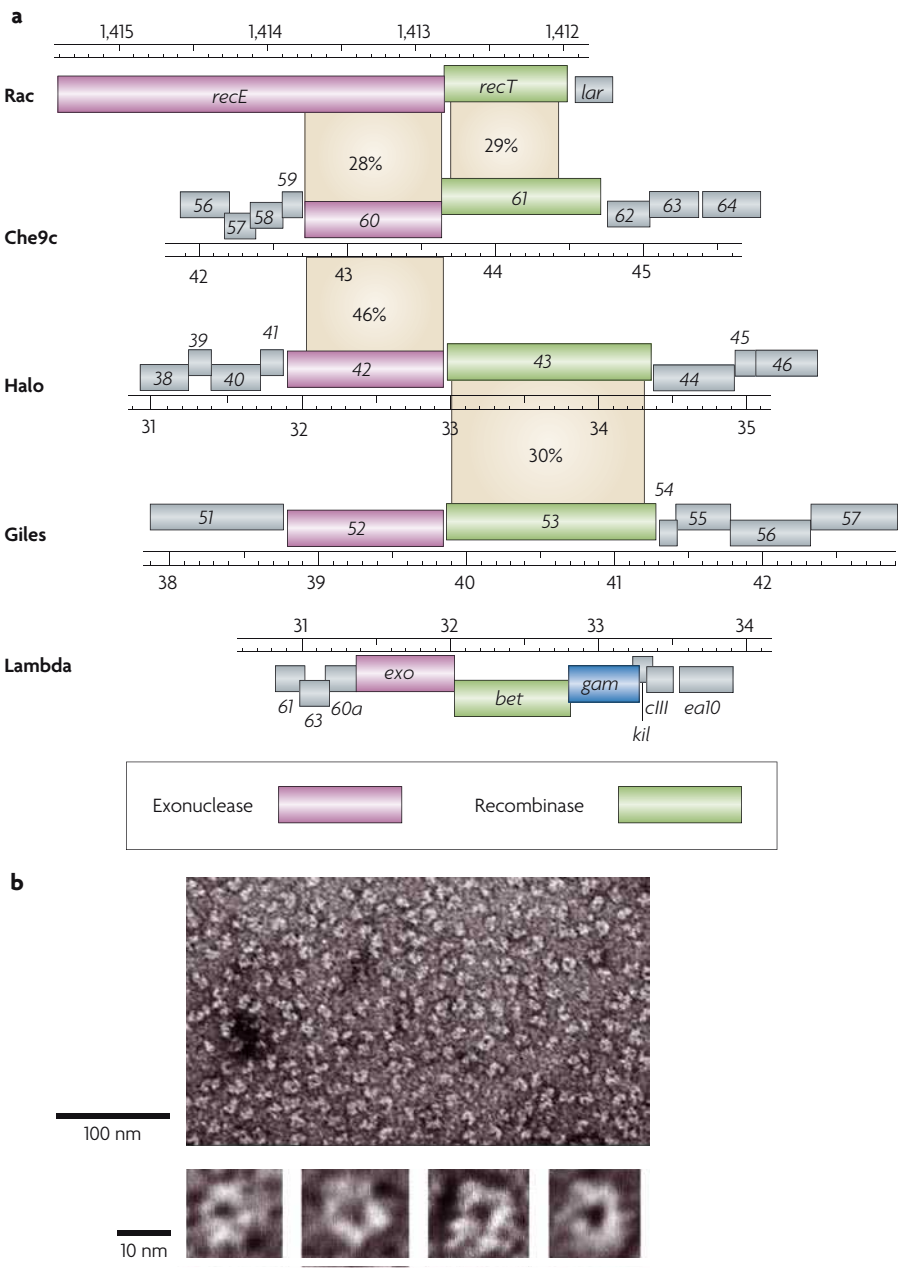

Figure 1 | **Mycobacteriophage-encoded recombination proteins. a** | Mycobacteriophage Che9c gene products 60 (gp60) and gp61 are *Escherichia coli* Rac prophage RecE and RecT homologues, respectively. Mycobacteriophage Halo encodes a RecE homologue, whereas mycobacteriophage Giles gp52 belongs to the YqaJ family of exonucleases. Halo gp43 is similar to Giles gp53, but these are only distantly related to other phage RecT-like proteins. The λ Red recombination proteins are not closely related to the mycobacteriophage proteins or the RecE or RecT proteins. None of the mycobacteriophages or the *E. coli* Rac prophage encodes a λ Gam homologue (blue). **b** | Electron micrographs of Che9c gp61 protein multimers in the presence of single-stranded DNA (ssDNA). Reaction mixtures that contain gp61 protein (1.2 μM) incubated with ssDNA (100 nucleotides; 2 μM) were adsorbed to glow discharged 400 mesh formvar carbon-coated copper grids, stained with 2% uranyl acetate, and examined by transmission electron microscopy. Four examples of torroid structures are shown; images were collected at a magnification of x140,000 (bottom).

Mycobacteriophages have helped us to circumvent some of the major challenges of genetics in *M. tuberculosis*, including shuttle-phasmid development[37] for transposon mutagenesis[38], reporter-gene delivery[39,40] and specialized transduction[11], as well as the production of integration-proficient vectors for stable introduction of foreign genes[41–44]. Exploitation of these phages has been simplified by the complete sequencing of more than 50 mycobacteriophage genomes[43–46], and the high level of genetic diversity discovered has provided us with at least 1,500 phamilies of unique genes. Because the *E. coli*-based systems seem to function less well in distantly related bacteria, we turned to this collection of mycobacteriophage genomes to identify mycobacteriophage-encoded proteins that could be developed for mycobacterial recombineering. Although RecE and RecT homologues are rare among mycobacteriophages, Mycobacteriophage Che9c (Che9c) encodes both proteins, and we have exploited these proteins to develop recombineering systems for both fast- and slow-growing mycobacteria[47,48]. This approach also provides a powerful strategy for constructing mutant derivatives of lytically replicating mycobacteriophage genomes, including in-frame deletions, point mutations and small insertions.

Identification of recombineering functions

Mycobacteriophages are genetically diverse, possess architecturally mosaic genomes[43–46,49] and are replete with predicted open reading frames of unknown function that have no detectable similarity to known proteins. While constructing shuttle-phasmid derivatives of Mycobacterium phage TM4 (TM4), Jacobs and colleagues[37] noted a high prevalence of recombinants following electroporation of phage libraries into *Mycobacterium smegmatis*, suggesting the presence of a phage-encoded recombination system. However, bioinformatic analysis of the TM4 genomic sequence[50] provides no clues as to which genes might encode recombination functions.

A search of the predicted open reading frames of all completely sequenced mycobacteriophage genomes reveals a small number of homologues of the *E. coli* RecE and RecT proteins but none that is related to the λ Red proteins. Che9c encodes for homologues of both RecE and RecT, the products of genes *60* and *61*, respectively[46,47] (FIG. 1a); both are distant relatives and share less than 30% amino-acid identity. Che9c gene product 60 (gp60; 314 amino acids) is also smaller than *E. coli* RecE (866 amino acids) and corresponds to the carboxy (C)-terminal part that encompasses the RecB-family nuclease domain (FIG. 1a); the amino (N)-terminal region of RecE that is absent in Che9c gp60 is not necessary for RecE exonuclease activity[51]. Che9c gp61 shares 29% amino-acid identity with *E. coli* RecT, and is therefore clearly a member of the

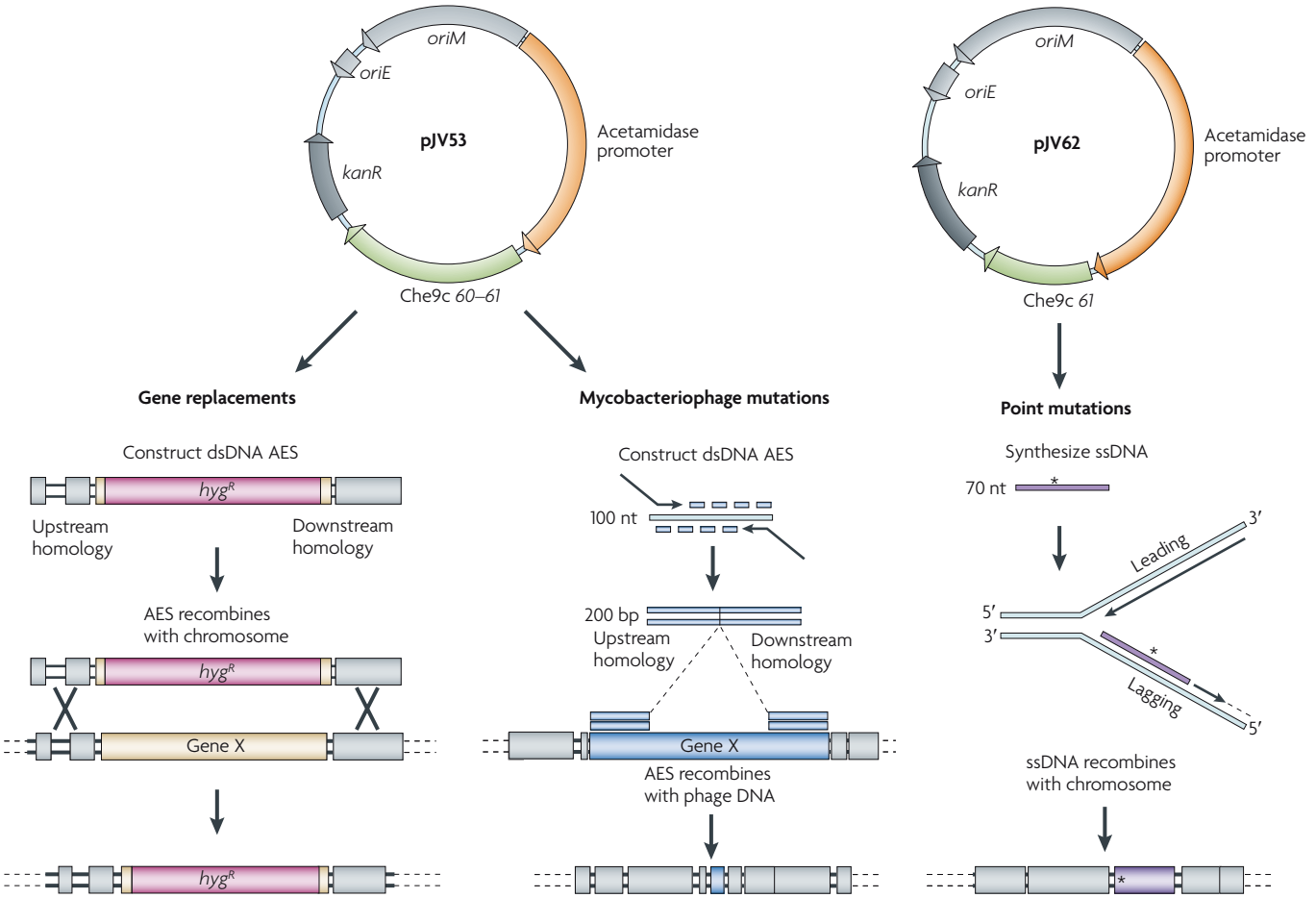

Figure 2 | **Strategies for mycobacterial recombineering.** Plasmids pJV53 and pJV62 either express both Che9c gene products 60 (gp60) and gp61 or only gp61 to facilitate double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) recombination, respectively. Gene knockouts are made by targeted gene replacement with a linear dsDNA allelic exchange substrate (AES) that contains 500 bp of homology to the target locus that flanks an antibiotic resistance gene (left; for example, hygromycin resistance (hyg^R)). Mycobacteriophage mutants (either point mutants or deletion mutants) are made by synthesizing a 200 bp dsDNA AES by PCR using a 100 nucleotide (nt) template and extender primers (75 nt) that contain 25 nt homology to the template and 50 nt extension at the 5′ end (middle). Point mutants are constructed by synthesizing a 50–70 nt ssDNA substrate and centring the desired point mutation, which will anneal to the lagging strand to allow DNA replication to undergo recombination. An asterisk represents the point mutation.

superfamily of SSAPs[30] but with a longer C-terminal extension (FIG. 1a). Although Che9c gp60 and gp61 are distant relatives of RecE and RecT and the λ Exo and Beta proteins, they possess similar biochemical functions[47]. Che9c gp60 has exonuclease activity *in vitro* that is dependent on the presence of dsDNA ends, and Che9c gp61 binds short (20 nucleotides) ssDNA as well as dsDNA substrates in the absence of Mg^{2+}. Che9c gp61 was also observed to form toroidal multimers in the presence of ssDNA by electron microscopy (J.C.v.K. and G.F.H., unpublished observations; FIG. 1b), a property exhibited by other SSAPs[52,53].

Two other mycobacteriophages, Mycobacteriophage Halo (Halo) and Mycobacteriophage Giles (Giles), encode similar recombination systems. Halo gp42 is 46% identical to Che9c gp60 and 30% identical to RecE, whereas Giles gp52 contains a domain from the YqaJ family of phage-encoded exonucleases (FIG. 1a). However, even though Halo gp43 and Giles gp53 are 30% identical, they are even more distantly related to other phage-encoded RecT homologues than Che9c gp61 (FIG. 1a). Interestingly, a prophage in the sequenced genome of *Mycobacterium avium* strain 104 also encodes RecE and RecT-like proteins (MAV_0830 and MAV_0829, respectively) that are similar to Che9c gp60 and gp61 (41% and 29% identical) and are also related to the Rac prophage proteins RecE and RecT (23% and 40% identical). Thus, although in each case the mycobacteriophage putative SSAP proteins are associated with a gene that encodes an exonuclease, these systems exhibit the same modularity observed in most other phage genomes[54]. None of these mycobacteriophage systems encodes homologues of the λ Gam protein, and it is unknown if there are any mycobacteriophage functional analogues that are RecBCD inhibitors similar to the Abc2 protein of P22.

Construction of recombineering strains

Mycobacterial strains have been constructed for recombineering with both dsDNA and ssDNA substrates. These contain an extrachromosomal plasmid in which the phage recombination genes are under the control of an inducible acetamidase promoter[55], and the most widely used configurations use the Che9c genes *60* and *61* for dsDNA

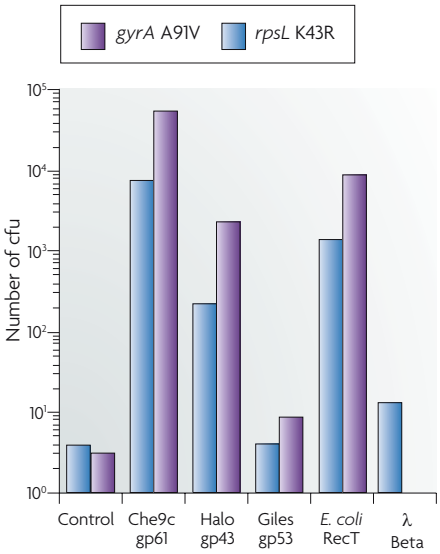

Figure 3 | **Mutagenesis of mycobacterial genomes.** Recombinases (single-stranded DNA (ssDNA) annealing proteins (SSAPs)) were assayed for ssDNA recombineering activity in *Mycobacterium smegmatis* by their ability to introduce streptomycin-resistance (*rpsL* K43R) and ofloxacin-resistance (*gyrA* A91V) mutations. SSAP genes are expressed from the acetamidase promoter and translationally fused to the pLAM12 vector[47] rather than using their native translation-initiation signals[48]. The number of drug-resistant colony forming units (cfu) obtained for each drug target is shown.

recombineering[47] and only Che9c *61* for ssDNA manipulations[48]. Optimal levels of recombination are obtained using plasmid derivatives that express Che9c genes *60* and *61* from their endogenous translational signals (for example, pJV53), although it should be noted that the higher levels of expression of some alternative constructions do not necessarily provide higher levels of recombineering[47]. This has also been observed for the λ Red system, the recombination activity of which correlates poorly with expression levels, particularly for the exonuclease[32].

Plasmid pJV53 serves as a basic recombineering plasmid for regulated expression of gp60 and gp61 (FIG. 2), although other derivatives of this plasmid have been made, including those with different selectable markers or those with a *sacB* cassette to simplify removal of the plasmid by counter-selection following mutagenesis (plasmid pJV48). The acetamidase promoter is 'leaky', and there is a detectable level of Che9c gp61 expression in *M. tuberculosis* in some strains that are not induced, although recombination frequencies are low without induction[47]. High expression levels of these proteins are toxic to *M. smegmatis*, and plasmids in which the constitutive *Mycobacterium bovis* BCG *hsp60* promoter is linked to Che9c *60* and *61* do not transform mycobacteria. This toxicity seems to be derived mainly from the Che9c gp60 exonuclease, as plasmids that express only gp61 constitutively are tolerated, although they grow slowly. For ssDNA recombineering, plasmids such as pJV62 (FIG. 2) have been constructed that express gp61 using its own translation signals and the acetamidase promoter. In preparation for recombineering experiments, *M. smegmatis* or *M. tuberculosis* strains that contain a plasmid are induced in mid-logarithmic growth with acetamide and harvested for electroporation[47,48]. Alternatively, *M. tuberculosis* strains can be grown in standard medum, washed into induction medium and incubated for 24 hours to induce expression before harvesting. This protocol yields similar recombineering frequencies, but reduces the total number of cells owing to losses during the washing steps (J.C.v.K., L.J.M. and G.F.H., unpublished observations).

Constructing gene replacements

Targeted gene-replacement mutants can be readily constructed by electroporation of linear dsDNA substrates into either *M. smegmatis* or *M. tuberculosis* recombineering strains (FIG. 2). The AESs typically contain regions of homology upstream and downstream of the target gene that flanks a cassette for antibiotic resistance, such as the hygromycin resistance cassette (*hyg*^{R}). Although SSAP recombinases, such as gp61, can bind short regions of ssDNA, the frequency of recombineering is considerably lower with 50 bp than with 500 bp of homology, and with shorter lengths the DNA-uptake frequency becomes limiting; we therefore recommend using a minimum of 500 bp of homology[47]. A simple protocol entails transformation of electrocompetent mycobacterial cells that contain pJV53 and are induced for gp60 and gp61 expression with 100 ng of a linearized AES[47,56]. This typically yields 50–200 drug-resistant colonies of which >90% contain correctly targeted gene replacements for both *M. smegmatis* and *M. tuberculosis*; control transformations using 50 ng plasmid DNA with the same competent cells typically yield approximately 1x10^5 colonies. Similar frequencies are observed when extrachromosomal plasmids are targeted, emphasizing the need to ensure that DNA with potential similarity to the pJV53 plasmid backbone (such as the *E. coli* origin of replication) is removed from the AES substrate. The AES can be generated using either PCR or restriction digests to remove the plasmid backbone; if using restriction digests, using enzymes with incompatible sites can prevent re-circularization of the plasmid and undesired insertional duplication. AESs can also be designed to subsequently unmark mutants; for example, by including γδ resolvase sites that flank the antibiotic resistance cassette and a *sacB* cassette for counter selection[11].

The dsDNA recombineering frequencies seem to be limited primarily by poor DNA-uptake efficiencies of mycobacteria rather than poor protein expression or DNA degradation by host nuclease systems. With either *M. tuberculosis*[57] or a high-efficiency-transformation *M. smegmatis* strain[58] competencies as high as 1–5 x 10^6 transformants per μg of plasmid DNA can be achieved[47,59], but only 1 viable cell in 1,000 takes up DNA productively. Inactivation of the *M. smegmatis recBCD* system only modestly increases dsDNA recombineering frequencies, and expression of λ Gam does not influence the recombineering frequencies in any strain background that we have tested (J.C.v.K. and G.F.H., unpublished observations). In practice, the recombination and DNA-uptake frequencies with the Che9c system yield sufficient recombinants for the construction of targeted gene-replacement mutants in both *M. smegmatis* and *M. tuberculosis*.

Construction of point mutations

Recombineering using ssDNA substrates offers a simple method for constructing point mutations in mycobacterial genomes[48]. The overall efficiency of ssDNA recombineering is substantially higher than with dsDNA substrates for both chromosomal and plasmid targets. Moreover, the exonuclease function (Che9c gp60) is not required for ssDNA substrates, and strains that contain plasmid, such as pJV62 (FIG. 2), are suitable[48].

In ssDNA recombineering, a choice needs to be made as to which DNA strand to target. In mycobacterial recombineering this is a crucial issue, as ssDNA substrates that target the same sequence, but on different strands, can differ in recombineering efficiencies by up to 10,000-fold[48]; this exceeds by far the 2–50-fold strand biases observed in the λ Red system[19,31]. Recombination frequencies are always higher using a ssDNA substrate that anneals to the template for discontinuous DNA synthesis (lagging strand), presumably owing to the greater availability of ssDNA chromosomal DNA

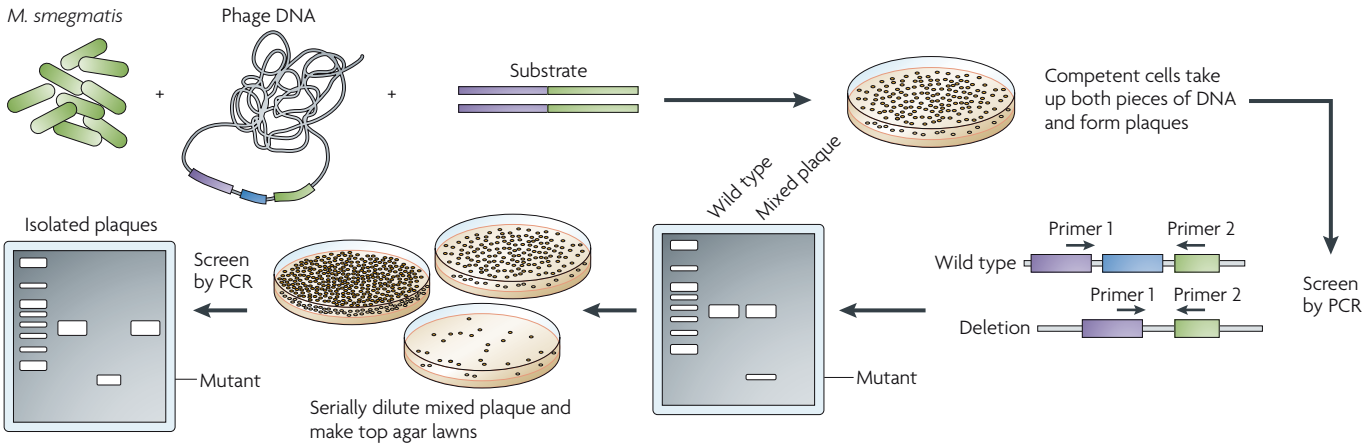

Figure 4 | **Mycobacteriophage recombineering.** Schematic of the strategy used to adapt mycobacterial recombineering for use on mycobacteriophages. Induced electrocompetent *Mycobacterium smegmatis* pJV53 cells are co-electroporated with phage DNA (50–100 ng) and a 200 bp double-stranded DNA recombineering substrate (50–500 ng), mixed with *M. smegmatis* plating cells and plated for individual plaques. Approximately 10–40% of plaques are mixed (they contain both the mutant and wild-type alleles) and are distinguishable by PCR. A mixed plaque can be diluted and plated to analyse individual plaques or prepare a lysate. If a non-essential gene is deleted a mutant phage derivative can be readily identified by PCR, but if the gene is essential, the mutant allele is no longer present in the lysate.

for pairing with gp61–ssDNA substrate complexes. It is not clear why the biases are so much larger in mycobacteria than in *E. coli*, but it may reflect fundamental differences in the DNA replication systems or in how gp61 interacts with the replication machinery. Substrates as short as 48 nucleotides provide optimal recombination frequencies, and this recombination is also independent of host RecA functions[48].

A simple experiment to test for ssDNA recombineering uses an oligonucleotide substrate to introduce a chromosomal point mutation that confers drug resistance. In a typical experiment that uses an oligonucleotide to target the lagging strand, 100 ng of ssDNA yields a similar number of drug-resistant colonies as that obtained using 100 ng of extrachromosomal plasmid DNA (~10^5 colonies) in *M. smegmatis*; this frequency is lower in *M. tuberculosis* (fivefold to tenfold). Thus, in contrast to dsDNA recombineering, a high proportion of cells that take up ssDNA undergo recombination. The question then arises as to whether the overall frequency is high enough to construct point mutations that cannot be directly selected, bearing in mind that 99.9% of cells do not take up DNA.

This problem can be conveniently circumvented using a counter-selection strategy, in which a plasmid is co-transformed with the ssDNA and plasmid transformants are selected; this effectively counter selects against non-competent cells in the population. The transformants can then be examined by PCR, and we have found the mismatch amplification mutation assay (MAMA) PCR[60,61] that selectively amplifies the mutant allele to be particularly effective. In our experience with *M. smegmatis*, 5–10% of the plasmid transformants also contain the newly introduced point mutation; the frequencies in *M. tuberculosis* are also slightly lower using this technique. A similar outcome can be achieved using a double-oligonucleotide configuration, in which two ssDNA substrates are used in the same electroporation: one introduces the desired point mutation in the chromosome, whereas the other repairs a mutant *hyg*^{R} gene that is present on the recombineering plasmid. The proportion of Hyg^R colonies that contain the point mutation is lower (~3%), but can still be readily identified by MAMA PCR; the advantage is that only a single plasmid subsequently needs to be removed from strains for further analyses. A plasmid (pJV128) has been constructed that contains Che9c gp61 fused to the acetamidase promoter, a Kan^R selectable marker, a mutated *hyg*^{R} gene and a *sacB* counter-selectable gene for this purpose[48].

This counter-selection strategy can also be used to make unmarked deletions with dsDNA substrates, and we have successfully constructed a deletion of the *M. smegmatis leuD* gene using a 200 bp dsDNA substrate with 100 bp *leuD* homology on each side of the deletion (J.C.v.K. and G.F.H., unpublished observations). The optimal parameters for making deletions — other than the requirement for dsDNA substrates instead of ssDNA substrates — have not yet been analysed, but could include the effects of deletion length and substrate length on recombineering efficiencies, and thus far this method has only been tested in *M. smegmatis*. However, we think this will be an important addition to the mycobacterial recombineering technology, particularly for the construction of vaccine strains for which antibiotic resistance genes should be avoided.

An intriguing use of the ssDNA recombineering technology is the study of mutations that confer antibiotic resistance. This has been shown by introducing several well characterized point mutations that confer resistance to isoniazid, rifampicin, streptomycin and ofloxacin in *M. smegmatis*, as well as rifampicin and streptomycin in *M. tuberculosis*[48]. Recombineering with ssDNA thus offers a simple approach that allows us to determine whether any point mutation identified in a clinical strain of *M. tuberculosis* contributes to its drug-susceptibility profile.

Identification of phage recombinases

Although homologues of known recombinases are rare in mycobacteriophages, the Halo gp43 and Giles gp53 proteins are candidates for recombination activity (FIG. 1a). Introduction of point mutations that confer drug resistance by ssDNA recombineering provides a simple assay for testing these mutations as well as SSAPs from non-mycobacteriophage sources, including λ Beta and RecT (FIG. 3). Halo gp43 is less active than Che9c gp61, and Giles gp53 is even less active than both Halo gp43 and Che9c gp61, which further illustrates the usefulness of a large reservoir of phage genomes to optimize potential genetic tools[48]. Anecdotal reports that the λ Red system functions

poorly in mycobacteria are confirmed by the low activity of λ Beta. RecT works well, however, in *M. smegmatis* (FIG. 3), although its activity is still reduced compared with Che9c gp61. Related studies report that Che9c gp61 performs poorly in *E. coli* ssDNA recombineering assays[62]. The molecular basis for differing activities of SSAPs is not clear, although an intriguing possibility is that they act by interacting directly with the host DNA-replication machinery.

Mycobacteriophage mutations

Constructing mutant derivatives of bacteriophages is often more difficult than manipulating the host chromosome, mostly because drug selection is not useful in lytically propagated viruses. Currently, the most powerful method for manipulating mycobacteriophages is through the construction of shuttle phasmids[63]. However, although such plasmids have many applications, they are of only limited use for determining phage gene functions. In one example, TM4 shuttle phasmids were used to make mutations in the TM4 *tape measure* gene, which allowed mutagenesis to be performed in *E. coli* (by λ Red recombineering) and the mutant phage to be recovered in *M. smegmatis*[64]. However, most mycobacteriophage genomes are too large for shuttle-phasmid construction, and this approach is not broadly applicable for functional genomic studies on mycobacteriophages. A homologous recombination approach was used to construct a fire-fly luciferase recombinant of mycobacteriophage L5 (REF. 65), but the frequency of host-mediated recombination was so low that construction was inefficient and time-consuming. Mycobacterial recombineering can be used to manipulate prophages, but is not generally applied because few mycobacteriophages form stable lysogens.

The mycobacterial recombineering systems described above offer potentially powerful approaches for the construction of mutants of any lytically growing mycobacteriophage. Recombineering of phage genomes has been described for *E. coli* phages, such as the *E. coli* phage λ (lambda), using a strategy in which bacterial cells are infected with phage, competent cells are prepared, dsDNA or ssDNA substrate is introduced by electroporation and plaques are analysed for the presence of the mutation[66]. Because the frequency of DNA uptake in mycobacteria is substantially lower than the efficiency of phage infection, we have adopted an alternative approach that is based on the co-selection methods described above. Although these initial experiments are still being developed, they could be a highly effective approach for mycobacteriophage functional genomics (L.J.M. and G.F.H., unpublished observations).

The scheme we have developed for mycobacteriophage mutagenesis is shown in FIG. 4. An *M. smegmatis* strain that carries a recombineering plasmid, such as pJV53 or pJV62 (FIG. 2), is induced with acetamide and electrocompetent cells are prepared. Electroporation is performed with two DNAs: phage genomic DNA of the virus to be manipulated and a PCR-generated or synthetic DNA substrate. Plaques are recovered in an infectious centre assay (before lysis) such that cells that have taken up phage DNA give rise to plaques on a lawn of *M. smegmatis* plating cells; approximately 100 plaques are obtained from 50–100 ng of phage DNA in such an experiment. All plaques contain wild-type phage DNA, but 10–40% also contain the mutant allele and can be easily identified by PCR screening of 12–18 individual plaques. Provided that the deleted region is non-essential for phage growth, purified mutant phage derivatives can be isolated by plating serial dilutions of the mixed plaques and PCR analysis of the purified plaques (FIG. 4). Typically, analysis of 10–18 isolated plaques from this plaque purification is sufficient to identify a homogenously mutant isolate.

We have used phage recombineering to manipulate several different mycobacteriophage genomes and to generate gene knockouts, in-frame internal deletions, point mutations and small insertions. When constructing deletion mutants, we have found that a higher proportion of primary plaques contain the desired mutant allele when using dsDNA substrates than when using ssDNA substrates. Larger dsDNA substrates (200 bp) also seem to work better than shorter ones (100 bp) as with mycobacterial chromosomal recombineering[47]. We have also adapted this approach to determine whether mycobacteriophage genes are required for lytic growth. In one experiment, we used a 200 bp dsDNA substrate to generate a deletion in the lysin A gene of Giles that seemed to be an essential function. PCR analysis showed that approximately 30% of the initial plaques recovered contained mutants, but when lysates were prepared after the re-plating of these mixed plaques, the mutant allele could not be identified even using a sensitive MAMA PCR approach (discussed above). We conclude that the gene is essential for the growth of Giles, and that the mutation could be constructed and the mutant propagated in the presence of wild-type phage that acts as a helper phage in the primary plaque, even though single particles do not give rise to plaques in the secondary plating. The full usefulness of these approaches for mycobacteriophage mutagenesis has perhaps yet to be realized, but the ability to modify mycobacteriophage genomes using mycobacterial recombineering will, for the first time, allow us to use broad functional genomic studies on mycobacteriophages.

Development of recombineering systems

The ability to stimulate homologous recombination in both slow- and fast-growing mycobacteria by expressing mycobacteriophage-encoded recombination systems could be exploited in various ways. First, it seems likely that the same approaches could be used to manipulate a range of mycobacterial species other than those described here (*M. smegmatis* and *M. tuberculosis*), including other bacterial species that are closely related to mycobacteria. Second, the high efficiencies of chromosomal recombineering could facilitate the construction of ordered gene knockout or replacement mutants, particularly because the identification of mutants by PCR is compatible with genome-wide robotic strategies. Last, the coupling of efficient point mutagenesis with mycobacterial nonsense suppressor strains[67] is a simple method that could allow us to produce conditional lethal mutants in mycobacterial genomes and their phages.

Julia C. van Kessel, Laura J. Marinelli and Graham F. Hatfull are at the Pittsburgh Bacteriophage Institute and Department of Biological Sciences, 376 Crawford Hall, 4249 Fifth Avenue, University of Pittsburgh, Pittsburgh, Pennsylvania 15260, USA.
Correspondence to G.F.H.
e-mail: gfh@pitt.edu

doi:10.1038/nrmicro2014

1. Aldovini, A., Husson, R. N. & Young, R. A. The *uraA* locus and homologous recombination in *Mycobacterium bovis* BCG. *J. Bacteriol.* **175**, 7282–7289 (1993).
2. Kalpana, G. V., Bloom, B. R. & Jacobs, W. R. Jr. Insertional mutagenesis and illegitimate recombination in mycobacteria. *Proc. Natl Acad. Sci. USA* **88**, 5433–5437 (1991).
3. Hatfull, G. F. in *Tuberculosis* (eds Eisenach, K., Cole, S. T., Jacobs, W. R. Jr & McMurray, D.) 203–218 (ASM Press, Washington DC, 2004).
4. Balasubramanian, V. *et al.* Allelic exchange in *Mycobacterium tuberculosis* with long linear recombination substrates. *J. Bacteriol.* **178**, 273–279 (1996).
5. Azad, A. K., Sirakova, T. D., Fernandes, N. D. & Kolattukudy, P. E. Gene knockout reveals a novel gene cluster for the synthesis of a class of cell wall lipids unique to pathogenic mycobacteria. *J. Biol. Chem.* **272**, 16741–16745 (1997).

6. Hinds, J. *et al.* Enhanced gene replacement in mycobacteria. *Microbiology* **145**, 519–527 (1999).
7. Parish, T. & Stoker, N. G. Use of a flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis tlyA plcABC* mutant by gene replacement. *Microbiology* **146**, 1969–1975 (2000).
8. Pashley, C. A., Parish, T., McAdam, R. A., Duncan, K. & Stoker, N. G. Gene replacement in mycobacteria by using incompatible plasmids. *Appl. Environ. Microbiol.* **69**, 517–523 (2003).
9. Pelicic, V., Reyrat, J. M. & Gicquel, B. Positive selection of allelic exchange mutants in *Mycobacterium bovis* BCG. *FEMS Microbiol. Lett.* **144**, 161–166 (1996).
10. Sander, P., Meier, A. & Bottger, E. C. *rpsL*$^+$: a dominant selectable marker for gene replacement in mycobacteria. *Mol. Microbiol.* **16**, 991–1000 (1995).
11. Bardarov, S. *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* **148**, 3007–3017 (2002).
12. Vilcheze, C. *et al.* Transfer of a point mutation in *Mycobacterium tuberculosis inhA* resolves the target of isoniazid. *Nature Med.* **12**, 1027–1029 (2006).
13. Court, D. L., Sawitzke, J. A. & Thomason, L. C. Genetic engineering using homologous recombination. *Annu. Rev. Genet.* **36**, 361–388 (2002).
14. Datta, S., Costantino, N. & Court, D. L. A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**, 109–115 (2006).
15. Ranallo, R. T., Barnoy, S., Thakkar, S., Urick, T. & Venkatesan, M. M. Developing live *Shigella* vaccines using λ Red recombineering. *FEMS Immunol. Med. Microbiol.* **47**, 462–469 (2006).
16. Copeland, N. G., Jenkins, N. A. & Court, D. L. Recombineering: a powerful new tool for mouse functional genomics. *Nature Rev. Genet.* **2**, 769–779 (2001).
17. Sarov, M. *et al.* A recombineering pipeline for functional genomics applied to *Caenorhabditis elegans*. *Nature Methods* **3**, 839–844 (2006).
18. Yu, D. *et al.* An efficient recombination system for chromosome engineering in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **97**, 5978–5983 (2000).
19. Ellis, H. M., Yu, D., DiTizio, T. & Court, D. L. High efficiency mutagenesis, repair, and engineering of chromosomal DNA using single-stranded oligonucleotides. *Proc. Natl Acad. Sci. USA* **98**, 6742–6746 (2001).
20. Zhang, Y., Buchholz, F., Muyrers, J. P. & Stewart, A. F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature Genet.* **20**, 123–128 (1998).
21. Gottesman, M. M., Gottesman, M. E., Gottesman, S. & Gellert, M. Characterization of bacteriophage λ reverse as an *Escherichia coli* phage carrying a unique set of host-derived recombination functions. *J. Mol. Biol.* **88**, 471–487 (1974).
22. Little, J. W. An exonuclease induced by bacteriophage λ. II. Nature of the enzymatic reaction. *J. Biol. Chem.* **242**, 679–686 (1967).
23. Joseph, J. W. & Kolodner, R. Exonuclease VIII of *Escherichia coli*. I. Purification and physical properties. *J. Biol. Chem.* **258**, 10411–10417 (1983).
24. Hall, S. D., Kane, M. F. & Kolodner, R. D. Identification and characterization of the *Escherichia coli* RecT protein, a protein encoded by the *recE* region that promotes renaturation of homologous single-stranded DNA. *J. Bacteriol.* **175**, 277–287 (1993).
25. Kmiec, E. & Holloman, W. K. β protein of bacteriophage λ promotes renaturation of DNA. *J. Biol. Chem.* **256**, 12636–12639 (1981).
26. Li, Z., Karakousis, G., Chiu, S. K., Reddy, G. & Radding, C. M. The beta protein of phage λ promotes strand exchange. *J. Mol. Biol.* **276**, 733–744 (1998).
27. Noirot, P. & Kolodner, R. D. DNA strand invasion promoted by *Escherichia coli* RecT protein. *J. Biol. Chem.* **273**, 12274–12280 (1998).
28. Hall, S. D. & Kolodner, R. D. Homologous pairing and strand exchange promoted by the *Escherichia coli* RecT protein. *Proc. Natl Acad. Sci. USA* **91**, 3205–3209 (1994).
29. Rybalchenko, N., Golub, E. I., Bi, B. & Radding, C. M. Strand invasion promoted by recombination protein β of coliphage λ. *Proc. Natl Acad. Sci. USA* **101**, 17056–17060 (2004).
30. Iyer, L. M., Koonin, E. V. & Aravind, L. Classification and evolutionary history of the single-strand annealing proteins, RecT, Redβ, ERF and RAD52. *BMC Genomics* **3**, 8 (2002).
31. Zhang, Y., Muyrers, J. P., Rientjes, J. & Stewart, A. F. Phage annealing proteins promote oligonucleotide-directed mutagenesis in *Escherichia coli* and mouse ES cells. *BMC Mol. Biol.* **4**, 1 (2003).
32. Muyrers, J. P., Zhang, Y., Buchholz, F. & Stewart, A. F. RecE/RecT and Redα/Redβ initiate double-stranded break repair by specifically interacting with their respective partners. *Genes Dev.* **14**, 1971–1982 (2000).
33. Costantino, N. & Court, D. L. Enhanced levels of λ Red-mediated recombinants in mismatch repair mutants. *Proc. Natl Acad. Sci. USA* **100**, 15748–15753 (2003).
34. Murphy, K. C. The λ Gam protein inhibits RecBCD binding to dsDNA ends. *J. Mol. Biol.* **371**, 19–24 (2007).
35. Marsic, N., Roje, S., Stojiljkovic, I., Salaj-Smic, E. & Trgovcevic, Z. *In vivo* studies on the interaction of RecBCD enzyme and λ Gam protein. *J. Bacteriol.* **175**, 4738–4743 (1993).
36. Murphy, K. C. Bacteriophage P22 Abc2 protein binds to RecC increases the 5′ strand nicking activity of RecBCD and together with λ bet, promotes Chi-independent recombination. *J. Mol. Biol.* **296**, 385–401 (2000).
37. Jacobs, W. R. Jr, Tuckman, M. & Bloom, B. R. Introduction of foreign DNA into mycobacteria using a shuttle phasmid. *Nature* **327**, 532–535 (1987).
38. Bardarov, S. *et al.* Conditionally replicating mycobacteriophages: a system for transposon delivery to *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **94**, 10961–10966 (1997).
39. Banaiee, N. *et al.* Rapid identification and susceptibility testing of *Mycobacterium tuberculosis* from MGIT cultures with luciferase reporter mycobacteriophages. *J. Med. Microbiol.* **52**, 557–561 (2003).
40. Jacobs, W. R. Jr *et al.* Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. *Science* **260**, 819–822 (1993).
41. Kim, A. I. *et al.* Mycobacteriophage Bxb1 integrates into the *Mycobacterium smegmatis groEL1* gene. *Mol. Microbiol.* **50**, 463–473 (2003).
42. Lee, M. H., Pascopella, L., Jacobs, W. R. Jr & Hatfull, G. F. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette–Guerin. *Proc. Natl Acad. Sci. USA* **88**, 3111–3115 (1991).
43. Morris, P., Marinelli, L. J., Jacobs-Sera, D., Hendrix, R. W. & Hatfull, G. F. Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. *J. Bacteriol.* **190**, 2172–2182 (2008).
44. Pham, T. T., Jacobs-Sera, D., Pedulla, M. L., Hendrix, R. W. & Hatfull, G. F. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology* **153**, 2711–2723 (2007).
45. Hatfull, G. F. *et al.* Exploring the mycobacteriophage metaproteome: phage genomics as an educational platform. *PLoS Genet.* **2**, e92 (2006).
46. Pedulla, M. L. *et al.* Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**, 171–182 (2003).
47. van Kessel, J. C. & Hatfull, G. F. Recombineering in *Mycobacterium tuberculosis*. *Nature Methods* **4**, 147–152 (2007).
48. van Kessel, J. C. & Hatfull, G. F. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Mol. Microbiol.* **67**, 1094–1107 (2008).
49. Morris, P., Marinelli, L. J., Jacobs-Sera, D., Hendrix, R. W. & Hatfull, G. F. Genomic characterization of mycobacteriophage Giles: evidence for phage acquisition of host DNA by illegitimate recombination. *J. Bacteriol.* **190**, 2172–2182 (2008).
50. Ford, M. E., Stenstrom, C., Hendrix, R. W. & Hatfull, G. F. Mycobacteriophage TM4: genome structure and gene expression. *Tuber. Lung Dis.* **79**, 63–73 (1998).
51. Chang, H. W. & Julin, D. A. Structure and function of the *Escherichia coli* RecE protein, a member of the RecB nuclease domain family. *J. Biol. Chem.* **276**, 46004–46010 (2001).
52. Passy, S. I., Yu, X., Li, Z., Radding, C. M. & Egelman, E. H. Rings and filaments of β protein from bacteriophage λ suggest a superfamily of recombination proteins. *Proc. Natl Acad. Sci. USA* **96**, 4279–4284 (1999).
53. Thresher, R. J., Makhov, A. M., Hall, S. D., Kolodner, R. & Griffith, J. D. Electron microscopic visualization of RecT protein and its complexes with DNA. *J. Mol. Biol.* **254**, 364–371 (1995).
54. Hendrix, R. W., Smith, M. C., Burns, R. N., Ford, M. E. & Hatfull, G. F. Evolutionary relationships among diverse bacteriophages and prophages: all the world's a phage. *Proc. Natl Acad. Sci. USA* **96**, 2192–2197 (1999).
55. Parish, T., Mahenthiralingam, E., Draper, P., Davis, E. O. & Colston, M. J. Regulation of the inducible acetamidase gene of *Mycobacterium smegmatis*. *Microbiology* **143**, 2267–2276 (1997).
56. van Kessel, J. C. & Hatfull, G. F. Mycobacterial recombineering. *Methods Mol. Biol.* **435**, 203–215 (2008).
57. Wards, B. J. & Collins, D. M. Electroporation at elevated temperatures substantially improves transformation efficiency of slow-growing mycobacteria. *FEMS Microbiol. Lett.* **145**, 101–105 (1996).
58. Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T. & Jacobs, W. R. Jr. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol. Microbiol.* **4**, 1911–1919 (1990).
59. Bibb, L. A. & Hatfull, G. F. Integration and excision of the *Mycobacterium tuberculosis* prophage-like element, φRv1. *Mol. Microbiol.* **45**, 1515–1526 (2002).
60. Cha, R. S., Zarbl, H., Keohavong, P. & Thilly, W. G. Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. *PCR Methods Appl.* **2**, 14–20 (1992).
61. Swaminathan, S. *et al.* Rapid engineering of bacterial artificial chromosomes using oligonucleotides. *Genesis* **29**, 14–21 (2001).
62. Datta, S., Costantino, N., Zhou, X. & Court, D. L. Identification and analysis of recombineering functions from Gram-negative and Gram-positive bacteria and their phages. *Proc. Natl Acad. Sci. USA* **105**, 1626–1631 (2008).
63. Jacobs, W. R. Jr *et al.* Genetic systems for mycobacteria. *Meth. Enzymol.* **204**, 537–555 (1991).
64. Piuri, M. & Hatfull, G. F. A peptidoglycan hydrolase motif within the mycobacteriophage TM4 tape measure protein promotes efficient infection of stationary phase cells. *Mol. Microbiol.* **62**, 1569–1585 (2006).
65. Sarkis, G. J., Jacobs, W. R. Jr & Hatfull, G. F. L5 luciferase reporter mycobacteriophages: a sensitive tool for the detection and assay of live mycobacteria. *Mol. Microbiol.* **15**, 1055–1067 (1995).
66. Oppenheim, A. B., Rattray, A. J., Bubunenko, M., Thomason, L. C. & Court, D. L. *In vivo* recombineering of bacteriophage λ by PCR fragments and single-strand oligonucleotides. *Virology* **319**, 185–189 (2004).
67. Ghosh, P., Wasil, L. R. & Hatfull, G. F. Control of phage Bxb1 excision by a novel recombination directionality factor. *PLoS Biol.* **4**, e186 (2006).

Acknowledgements

This work was supported by US National Institutes of Health grant AI067649. We thank W. Jacobs Jr and D. Court for helpful discussions, and M. Bochman for help with electron microscopy.

DATABASES

Entrez Genome: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genome
Che9c | Halo | Giles | lambda | P22 | TM4

Entrez Genome Project: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomeprj
Escherichia coli | *Mycobacterium avium* | *Mycobacterium bovis* BCG | *Mycobacterium tuberculosis* | *Mycobacterium smegmatis*

ALL LINKS ARE ACTIVE IN THE ONLINE PDF