

Genetic Engineering in Mycobacteria

D. K. Armianinova^{a, *}, D. S. Karpov^b, M. S. Kotliarova^a, and A. V. Goncharenko^a

^a Federal Research Center “Fundamentals of Biotechnology”, Russian Academy of Sciences, Moscow, 119071 Russia

^b Center of Precision Genome Editing and Genetic Technologies for Biomedicine, Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Moscow, 119991 Russia

*e-mail: dasha.arm@yandex.ru

Received April 20, 2022; revised April 26, 2022; accepted April 28, 2022

Abstract—Genetic tools for targeted modification of the mycobacterial genome contribute to the understanding of the physiology and virulence mechanisms of mycobacteria. Human and animal pathogens, such as the *Mycobacterium tuberculosis* complex, which causes tuberculosis, and *M. leprae*, which causes leprosy, are of particular importance. Genetic research opens up novel opportunities to identify and validate new targets for antibacterial drugs and to develop improved vaccines. Although mycobacteria are difficult to work with due to their slow growth rate and a limited possibility to transfer genetic information, significant progress has been made in developing genetic engineering methods for mycobacteria. The review considers the main approaches to changing the mycobacterial genome in a targeted manner, including homologous and site-specific recombination and use of the CRISPR/Cas system.

Keywords: *Mycobacterium* spp, genome editing, homologous recombination, site-specific recombination, CRISPR/Cas

DOI: 10.1134/S0026893322060036

INTRODUCTION

The genus *Mycobacterium* includes both pathogenic and free-living species. The *M. tuberculosis* complex and *M. leprae* are the most important pathogens, causing tuberculosis and leprosy, respectively, while the *M. avium* complex, *M. abscessus*, *M. ulfans*, and *M. fortuitum* are opportunistic pathogens in humans. Many mycobacteria are free living. The species *M. neoaurum* is considered as a producer of steroid compounds for pharmaceuticals [1].

Mycobacterium tuberculosis causes tuberculosis, which is a widespread dangerous infectious disease and is among the ten main causes of mortality in the global population. *Mycobacterium tuberculosis* can survive and reproduce within macrophages and other immunocompetent cells, causing damage to lymphoid tissue, the lung, and other organs [2, 3]. Tuberculosis requires long-term therapy, and the spreading of broadly resistant and multidrug-resistant strains is a global medical problem. Targeted changes introduced in the microbial genome makes it possible to investigate the relationship between the genotype and physiological features of microorganisms and to understand

their molecular interactions with the environment and the host organism. The issue is of particular importance because functions of a substantial part of *M. tuberculosis* genes remain unknown although their nucleotide sequences were established [4, 5].

Genetic manipulations are rather difficult to perform in mycobacteria. Species of the *M. tuberculosis* complex grow slowly; it takes from 7 days to 3 weeks for them to produce colonies on solid media. Mycobacteria possess a hydrophobic cell wall, which consists of lipids, mycolic acids, polysaccharides, and peptidoglycans. The components allow mycobacteria to grow as a dense film and prevent efficient DNA transfer. Molecular genetic methods were still developed to transform mycobacteria with exogenous DNA, thus increasing the opportunity to modify their genome [6, 7].

Genome editing methods make it possible to change the mycobacterial genome in a targeted manner. Mutant strains with knockouts in particular genes, conditional knockdowns, or overexpression of genes of interest are important to construct in order to perform functional studies and to identify targets for new drugs [8–10]. Introduction of single-point amino acid substitutions allows studies of drug resistance, which may develop as a result of such mutations in the genes that code for drug targets or enzymes responsible for drug modification within the cell, including regulatory gene regions [11].

Abbreviations: BCG, bacille Calmette–Guerin (attenuated live vaccine to prevent tuberculosis); PCR, polymerase chain reaction; CRISPR, clustered regularly interspaced short palindromic repeat; Cas, CRISPR-associated protein; crRNA, CRISPR RNA; tracrRNA, trans-activating CRISPR RNA; PAM, protospacer adjacent motif.

Successful genome editing improved the protective properties of an *M. bovis* bacille Calmette–Guerin (BCG) strain. However, BCG often fails to ensure efficient protection, especially in the adult population [12, 13]. A BCG strain was used to design the VPM1002 vaccine, which shows success in clinical trials [14]. The VPM1002 recombinant strain was obtained by deleting the *ureC* urease gene and integrating *Listeria monocytogenes hly*, which codes for listeriolysin O. Better protective properties are characteristic of the vaccine strain because listeriolysin O facilitates cell exit from the phagosome at lower pH, while the urease gene deletion prevents mycobacteria from increasing the alkalinity of their environment [15]. Also, the recombinant live vaccine was created on the basis of *M. tuberculosis* to have a broader range of protective antigens, thus being advantageous over BCG. To achieve this, vaccine strain MTBVAC was constructed via consecutive unmarked deletions of two genes, *fadD26* and *phoP*, which are important for bacterial virulence [16]. The new vaccine has already been put to phase II clinical trials.

The review describes the main approaches to genome editing in mycobacteria with the use of homologous recombination, mycobacteriophage integrases, and the CRISPR/Cas systems.

HOMOLOGOUS RECOMBINATION-BASED GENOME ENGINEERING IN MYCOBACTERIA

Introduction of targeted changes in the genomes of microorganisms is of immense importance for studying the gene functions and constructing strains with desirable properties. A main approach to identifying the function of a gene is constructing its knockdown or knockout with subsequent complementation in the wild-type genome. Homologous recombination provides a means to replace the allele of interest with its mutant counterpart.

A gene knockout is obtained by allelic exchange, which requires two crossover events to occur on both sides of the region to be mutated (Fig. 1). To this end, cells are transfected with a vector that carries a modified target gene allele (e.g., an allele disrupted by insertion of a resistance gene) flanked with homology regions on both sides (Fig. 1a). Systems of this type were used in early allelic exchange experiments in mycobacteria. Only a limited success was achieved with their use because cell transformation was low efficient, the homologous recombination rate is low in mycobacteria, and illegitimate recombination occurred at a high rate. Use of linear DNA fragments was assumed to increase the homologous recombination rate, but transformation with a linear fragment most often failed to lead to a replacement of the target region in the genome [4]. The allelic exchange rate was found to be higher with ssDNA vectors (alkaline dena-

turation products or phagemide ssDNAs) or UV-irradiated ssDNAs in both fast- and slow-growing species [17].

The presence of marker gene ensuring resistance to a particular antibiotic in the genome of the resulting recombinant strain prevents its further use for biomedical purposes. Also, insertion of an antibiotic resistance gene may exert a polar effect on expression of downstream genes of the operon, thus affecting the resulting phenotype and complicating its characterization. Moreover, integration of the antibiotic resistance cassette in the chromosome excludes the marker from further genetic manipulation [18]. The problems associated with the presence of undesirable antibiotic-resistance genes in constructing serial knockouts are possible to solve by using the site-specific recombination systems obtained from bacteriophages or transposons (Fig. 1b).

Saccharomyces cerevisiae FLP recombinase can be used to remove the antibiotic resistance genes integrated in the bacterial genome as a result of homologous recombination. The enzyme was shown to be functional in *M. smegmatis* cells and mediates site-specific recombination between two FLP recognition target (*FRT*) sites in the chromosome [19]. When the *hyg* hygromycin resistance gene is flanked with two *FRT* sites in the direct orientation, *hyg* is specifically removed from the chromosome in *M. smegmatis* mutants expressing the *FLP* gene [19]. Additional codon optimization in the *FLP_m* gene ensured the more efficient function of the system in *M. smegmatis* and provided the possibility of using the system with slow-growing mycobacteria. For example, approximately 40% of resistant clones lost the *hyg* resistance cassette after short-time *FLP_m* expression in experiments with a *M. bovis* BCG strain. Chromosomal DNA sequencing showed that the *FRT-hyg-FRT* cassette was specifically excised by FLP [20].

The same principle underlies the function of a bacteriophage P1 system. A cassette with a resistance marker is flanked with loxP sites, and short-term expression of Cre recombinase ensures excision of the cassette [21]. Another variant utilizes the system of the $\gamma\delta$ transposon. A plasmid carrying a kanamycin resistance cassette flanked with two *res* sites was used to obtain unmarked deletions in *M. smegmatis* and BCG cells. Expression of the *tnpR* resolvase gene of the $\gamma\delta$ transposon in mutant strains led to efficient excision of the resistance gene flanked with *res* sites [18].

A system based on endogenous Xer recombinase can also be used to remove the resistance genes. In the system, an antibiotic resistance cassette is flanked with *dif* sites, which are recognized and resolved by XerC and XerD recombinases. The system does not require additional episomal elements to be introduced and then eliminated, thus providing a very simple and efficient tool [22]. The system was optimized and used to disrupt several genes in various regions of the mycobacterial genome. Such modifications are important

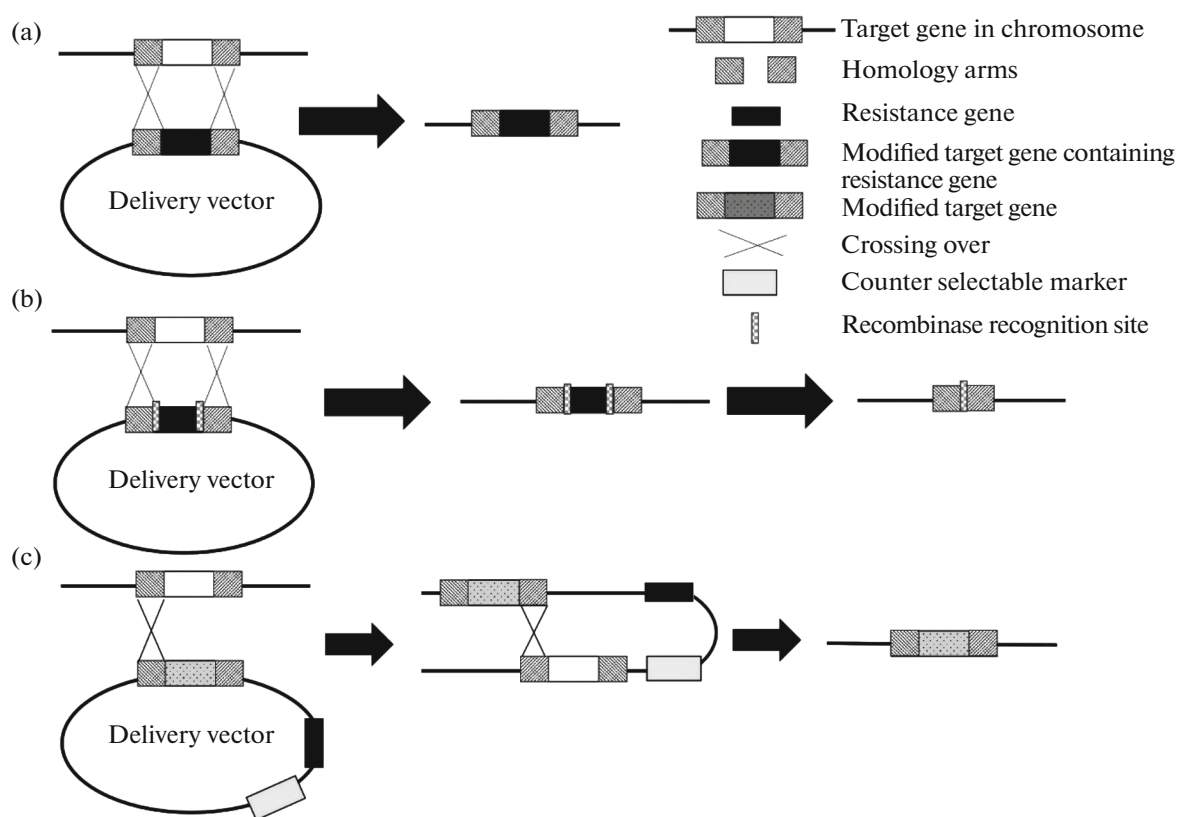


Fig. 1. Construction of knockouts and other target gene modifications via homologous recombination. (a) Construction of a marker-carrying mutant by homologous recombination. Cells are transformed with a vector that carries a modified gene allele containing a resistance gene. Mutants that have undergone crossover events on both sides of the mutation to be introduced are selected. (b) Construction of an unmarked knockout by site-specific recombination. A construct used for modification contains a modified allele with a resistance gene, which is flanked with recombinaise recognition sites. Allelic exchange proceeds in a single step, and the resistance gene is removed with recombinaise after selecting the mutants. (c) Two-step construction of unmarked mutants with the use of a suicide vector. The suicide vector is engineered to harbor a modified allele of the target gene and a cassette containing a marker gene linked with a counterselectable marker. After transformation, cells are plated on a medium that contains the respective antibiotic to select the mutants in which the first crossover event has taken place to integrate the vector body into the chromosome. The resulting mutants undergo a second crossover event and are subject to negative selection, where survival is only possible for the mutants that have lost the counterselectable marker together with the selectable marker.

for constructing attenuated mycobacterial strains and studying the synergistic effect of genes or the functions of duplicated genes [23].

Two-step allelic exchange to construct unmarked mutants became a classical tool in gene engineering of mycobacteria (Fig. 1c). Using this approach, Parish et al. [28, 29] designed the p2Nil/pGoal system to facilitate construct assembly for homologous recombination. Vectors of the p2NIL series are used to clone the homology arms or to modify the target gene; vectors of the pGOAL series contain marker cassettes, such as *lacZ* combined with *sacB* and an antibiotic resistance gene. Insertion of the marker cassette from pGOAL to p2NIL with a modified target gene yields a suicide vector, which is not replicated in mycobacteria. A knockout or modification is achieved in two steps. At the first step, cells are plated on a medium containing the antibiotic and X-gal to isolate blue colonies, in which the plasmid has been incorporated in

the genome via a single crossover event. To isolate the clones that have undergone a second crossover event, cells are plated on an antibiotic-free medium containing sucrose. Because the second crossover event results in a loss of *lacZ* and *sacB*, the colonies are white in color and not sensitive to sucrose. The method was used to construct the *tlyA* Δ *plcABC* Δ unmarked double mutant [24]. The method is widely used to solve various problems in studying the roles of proteins and small RNAs in the physiology of mycobacteria [25–27].

The *sacB* suicide gene (*Bacillus subtilis*) is broadly used to obtain both unmarked and marked mutations in *M. tuberculosis*. However, the frequency of spontaneous *sacB* inactivation may be almost the same as the recombination rate of certain genes, thus complicating mutant selection. Double counter selection (*sacB* with *Escherichia coli galk*) made it possible to achieve nearly 100% selection efficiency in mycobacteria [28]. The principle was used in optimized suicide vectors

(the pKO series), which are part of a uniform cloning platform for genetic manipulations in mycobacteria [10].

Replicating vectors can increase the likelihood of homologous recombination, but are more difficult to remove. Plasmids with a temperature-sensitive origin of replication are capable of replication at a permissive temperature (30–32°C), but are lost rapidly at a higher temperature (39–42°C). A counter selectable marker, such as *sacB*, should also be included in a plasmid designed for homologous recombination. A drawback of the method is that the cell growth rate is low at a permissive temperature [8].

Another system includes two pAL5000-based replicating plasmids, one containing *repA* (primase) and the other, *repB* (a DNA-binding protein). The genes complement each other in *trans* and ensure plasmid replication [29]. When antibiotic pressure is removed, one or both of the plasmids are lost and replication becomes impossible. The approach makes it possible to increase the time for allelic exchange.

A method of specialized transduction ensures almost 100% efficiency of DNA delivery into cells. A vector is assembled from two components, a cosmid vector with a sequence for allelic exchange and a conditionally replicating shuttle plasmid, which is derived from broad host range bacteriophage TM4 [30]. Mycobacteriophages are propagated in *M. smegmatis* at a permissive temperature (30°C), which ensures phage replication. A mycobacterial strain of interest is then transduced and cultured at a restrictive temperature (37°C), which inhibits phage replication. The mutant allele contains a resistance gene flanked with resolvase recognition target sites, and transient expression of resolvase (*tnpR*) eliminates the marker. The method was successfully used in BCG and [31] and *M. tuberculosis* [32] strains. More recently, the method was improved and tested by generating many single or multiple deletion substitutions in a targeted manner [33, 34].

Phage recombination proteins, such as Exo, Beta and Gam, or RecE and RecT, help to increase the homologous recombination rate and allow shorter homology arms to be used. The products of *gp60* and *gp61* of mycobacteriophage Che9c were shown to be homologous to RecE and RecT and to possess exonuclease and DNA-binding activities. The genes were used to construct pJV53 for homologous recombination in mycobacterial cells. Recombination occurs at a homology arm length of 50 bp, but is more efficient when the length exceeds 500 bp. The system makes it simpler to construct mutants in *M. smegmatis* and helps to overcome the illegitimate recombination effect in *M. tuberculosis* [35]. The green fluorescent protein gene (*gfp*) inserted in pJV53 facilitates verification of plasmid loss. With a hygromycin resistance cassette with dif sites, the system can be used to consecutively delete several genes in *M. smegmatis* [36]. A three-plasmid system was also designed to construct

unmarked knockouts. In the system, a helper plasmid carries a temperature-sensitive origin and the *sacB* suicide gene [37]. The pYS2 plasmid is used to engineer the mutant allele, which is modified to include a cassette with the hygromycin resistance and green fluorescent protein genes flanked with loxP sites. A linear substrate is obtained from the plasmid and used to transform cells. The pYS1 plasmid, which is a pJV53 derivative, is used to deliver *gp60* and *gp61* under an inducible promoter. The plasmid additionally harbors the *sacB* counter selectable marker, a temperature-sensitive origin of replication, and the kanamycin resistance gene. After recombination, cells are transformed with a third plasmid, pML2714, to express Cre recombinase and to excise the cassette. Mutant clones lose fluorescence and hygromycin resistance as a result.

Phage Che9c recombinases were used to develop a system for introducing point mutations in mycobacterial genomes with the use of ssDNA. Recombination mediated by *gp61* was efficient enough to introduce point nucleotide substitutions without performing direct selection; mutant strains were identified by PCR. However, the laborious screening procedure to select target clones limits the application of the system [38].

Thus, various homologous recombination methods are successfully employed in gene editing in mycobacteria. The approaches that combine homologous recombination for allelic exchange with site-specific recombination for subsequent elimination of the resistance genes make it possible to design simple and efficient editing systems, including the use of specialized transduction. Recombination engineering with RecE and RecT of phage Che9c ensures a manifold increase in system efficiency.

MYCOBACTERIOPHAGE RECOMBINASES IN GENE ENGINEERING OF MYCOBACTERIA

Site-specific recombination is a genetic recombination type where DNA strands are exchanged in a region between certain sequences. In contrast to homologous recombination, these specific DNA sites lack extended homology regions. To drive recombination, recombinase recognizes and binds the specific sequence to form a synaptic complex, which catalyzes chromosomal DNA cleavage with subsequent rearrangement and ligation of the cleavage site ends [39].

Site-specific recombination often leads to integration of one DNA molecule into the other, rather than to exchange of genetic information between the two molecules.

Recombinase can catalyze the reverse reaction of excising the integrated sequence in bacteriophages. Excision proceeds via site-specific recombination in many temperate bacteriophages. Recombinase alone

is sufficient for some recombination systems, while additional host factors are necessary for others.

Stable integration of mycobacteriophages into the host genome via site-specific recombination was addressed in many studies [40–45]. Mycobacterial phages L5, Ms6, Bxb1, and ϕ Rv1 were studied most comprehensively; the mechanism of their integration is considered below. The mycobacterial integrative element pSAM2 was also investigated, being initially identified in *Streptomyces*.

Temperate phage L5 is the most typical mycobacteriophage and infects both fast- and slow-growing mycobacterial species to produce stable lysogens [45]. The L5 prophage is integrated into a certain chromosome site during the lysogen phase and is excised in the lytic phase. Both integration and excision are catalyzed by phage-encoded Int integrase and require the host-encoded integration host factor (mIHF). The direction of these recombination events is determined by Xis, which is encoded by the phage gene 36. Integration is efficient in the absence of Xis, while excision is Xis dependent [46].

To allow site-specific recombination, a phage attachment site (attP) interacts with a bacterial chromosomal attachment site (attB). The integration reaction was reproduced in vitro and shown to require mIHF and Int [41, 47]. Supercoiling of the attP or attB sequence stimulates integrative recombination, but is not absolutely essential [48].

The minimal sizes of the functional attP and attB sites are approximately 240 and 29 bp, respectively. The attP sequence includes 43 bp that are common for the two sites, strain exchange occurs within this region, and recombination yields the integrated prophage flanked with left (attL) and right (attR) sites (Fig. 2) [49]. The attB site differs in one nucleotide between fast- and slow-growing mycobacteria, but the difference does not substantially affect the capability of mycobacteriophage L5 to efficiently infect mycobacteria of both groups [45].

Bacteriophage L5 integrates in the vicinity of the 3' end of the tRNA^{Gly} gene without altering its sequence. DNase I footprinting revealed an unusually long 413-bp region that serves as an attP site for L5 integrase. A 252-bp fragment sufficient for efficient phage integration was mapped within the attP site in a subsequent deletion analysis [49].

The L5 phage system was used for the first time to construct new recombinant BCG strains [43]. A DNA segment carrying the attP site and the integrase gene (*Int*) of mycobacteriophage L5 was used to replace the mycobacterial origin of replication (*oriM*) in the pMV261 shuttle vector, thus producing the pMV361 vector capable of integration. Because the phage Xis protein is absent, the integrated vector is stably maintained even without antibiotic pressure.

The genome organization is similar in mycobacteriophages Bxb1 and L5 [42]. Phage Bxb1 utilizes ser-

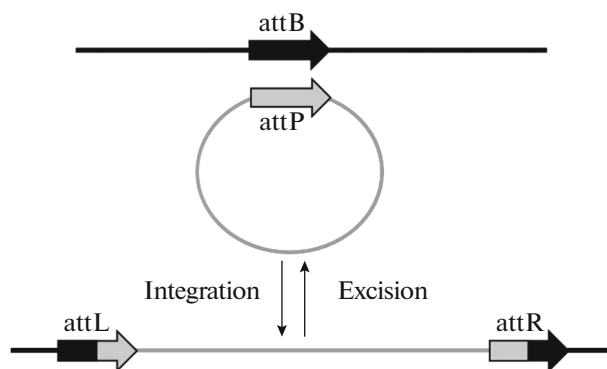


Fig. 2. General scheme of site-specific recombination.

ine integrase to integrate into the functional *groEL1* gene. There are two *groEL* genes in mycobacteria. The genes are highly similar (70%) at the nucleotide sequence level, but Bxb1 integrates only into *groEL1* because sequence specificity is required for integration. In contrast to many other site-specific recombinases, serine integrases lack a strong specificity to the substrate DNA structure and recombine regions within the same DNA molecule (in the head-to-head or head-to-tail orientation) or between different DNA molecules, affecting supercoiled, linear, and even double-stranded molecules [50].

In contrast to tyrosine integrase sites, the attP and attB sites of serine integrases are short, approximately 50 and 40 bp, respectively. The mechanism of DNA strand exchange also differs. Tyrosine integrases introduce single-strand breaks in DNA and exchange only one strand of each site to produce an intermediate structure similar to a Holliday junction [51], while serine integrases induce double-strand breaks in DNA and exchange the strands via a rotation mechanism [52].

Oligonucleotide-mediated recombinering followed by Bxb1 integrase targeting (ORBIT) was developed by combining two molecular tools, homologous recombination and site-specific integration [53, 54]. Target DNA fragments are inserted in two steps in the ORBIT system. A *M. smegmatis* or *M. tuberculosis* strain carrying a plasmid that express RecT recombinase of phage Che9c and integrase of phage Bxb1 is cotransformed with a short synthetic DNA oligonucleotide and a nonreplicating plasmid harboring the Bxb1 attB site, an antibiotic resistance gene to allow transformant selection, and a target sequence. The oligonucleotide is designed so that the phage Bxb1 attP site (48 bp) is flanked with regions of 45–70 bp that are homologous to the target chromosomal region.

At the first step, the oligonucleotide carrying the attP acceptor site for Bxb1 integrase is inserted into a necessary region of the mycobacterial genome via homologous recombination driven by Che9c RecT. At the second step, Bxb1 integrase facilitates site-specific recombination between the plasmid attB and the attP

site of the oligonucleotide insert obtained at the first step. The sequence of the synthetic oligonucleotide determines the insertion site position, while the plasmid serves as a target sequence donor in this system.

Bxb1 integrase functions independently of the host factors, than the success of ORBIT application depends on the efficiency of the first recombination step.

Mycobacteriophage Ms6 forms stable lysogens in *M. smegmatis*. The Ms6 attP site has a high A+T content and harbors many direct and inverted repeats. Ms6 integration into the host genome is mediated by Ms6 integrase, which targets integration to the 3' end of the tRNA^{Ala} gene in both fast- and slow-growing mycobacteria. A 26-bp central region of attP overlaps the 3' end of the tRNA^{Ala} gene, which is conserved in both of the mycobacterial groups [40].

The genomes of *M. tuberculosis* H37Rv and CDC1551 contain two prophage-like elements, ϕ Rv1 and ϕ Rv2. The ϕ Rv2 element encodes tyrosine recombinase, while ϕ Rv1 encodes large serine recombinase [55]. Recombination takes place between a putative attP site and the host chromosome; the attB site is within a redundant repetitive element (REP13E12), which occurs in seven copies in the *M. tuberculosis* genome. It is of interest that both ϕ Rv1 and ϕ Rv2 are absent from nonvirulent *M. bovis* BCG strains. Clinical *M. tuberculosis* isolates do not all have ϕ Rv2, but all seem to harbor at least one copy of ϕ Rv1 or ϕ Rv2. The two related elements presumably play a role in the physiology of *M. tuberculosis* [44]. The functional character of ϕ Rv1 was confirmed by efficient transformation of *M. bovis* BCG cell with a non-replicating plasmid that carried the integrase gene and the attP site (reconstructed from the attL and attR sites of the prophage). Four out of the seven REP13E12 sites present in the BCG genome can be used as attachment sites by vectors of the type, and the vectors can simultaneously occupy more than one site [55].

The pSAM2 element is an 11-kb integrative element and was initially characterized in *Streptomyces ambofaciens*. The pSAM2 recombination system includes integrase of the λ family [56] and the attB/P sites and is similar to the systems described above. The attB site covers the 58-bp region from the anticodon loop to the 3' end of the tRNA^{Pro} gene. The attB site is conserved among actinomycetes, including mycobacteria [57]. Thus, pSAM2 is capable of integration in various actinomycete species.

Several advantages are characteristic of the use of site-specific recombination. First, genes are reproducibly integrated into a known site (or at least a limited number of sites) in the mycobacterial genome. Every effect of the integration site (on gene expression and bacterial biology) is equally applicable to any insert. Second, a single gene copy is integrated, thus reducing the number of artifacts due to the use of mul-

ticopy plasmids. Third, strains that carry integrative plasmids are usually far more stable than strains that carry episomal vectors. Finally, mycobacteriophage integrases facilitate efficient integration of longer DNA fragments as compared with homologous recombination.

CRISPR/Cas SYSTEMS IN GENE ENGINEERING OF MYCOBACTERIA

CRISPR/Cas systems are prokaryotic systems that consist of clustered regularly interspaced short palindromic repeats (CRISPRs) and CRISPR-associated proteins. The systems are responsible for the adaptive immune response against foreign genetic material [58]. Class II systems include a single effector CRISPR protein, which utilizes CRISPR RNA (crRNA) to recognize and hydrolyze target DNA, and are used to edit various genomes in vitro. The type II-A system of *Streptococcus pyogenes* and the type V-A system of *Francisella novicida* are the most common [59, 60]. The *S. pyogenes* genome editing system includes multidomain RNA-dependent DNA endonuclease SpyCas9, whose specificity is determined by a guide RNA. The guide RNA has a 20-nt region known as the spacer, which is complementary to target DNA (a protospacer), and a structural part, which is recognized by SpyCas9 endonuclease. The guide RNA function may be performed by two RNAs, crRNA (encoded by CRISPRs) and trans-activating crRNA (tracrRNA, which is encoded by a separate gene and is necessary for crRNA processing and SpyCas9 binding), or a single guide RNA (sgRNA), which is a fusion product of the above two RNAs [61]. Correct recognition of the target DNA sequence (the protospacer) requires that a short element known as the protospacer adjacent motif (PAM) with the sequence 5'-NGG-3' occurs immediately downstream of the target. When PAM is present and the spacer of the guide RNA nearly perfectly matches the protospacer of genomic DNA, DNase activity of SpyCas9 is activated. SpyCas9 produces a double-strand break with blunt ends, the break activates cell DNA repair mechanisms, and a heritable change in genome sequence arises during repair [60]. The V-A type system derived from *F. novicida* consists of two parts, multidomain RNA-dependent DNA endonuclease FnoCas12a (Cpf1) and a guide RNA, which has a short hairpin recognized by the enzyme and a 23-nt spacer [62]. PAM recognized by FnoCas12a endonuclease has the consensus sequence 5'-BTTV-3' [63] and should occur upstream of the DNA protospacer, unlike in the type II-A *S. pyogenes* system [64]. In addition, FnoCas12a produces ends with overhangs when introducing a double-strand break in DNA, in contrast to SpyCas9 [64].

The *S. pyogenes* CRISPR/Cas system was the first to be used as a heterologous system to edit the mycobacterial genome. However, SpyCas9 endonuclease

was found to be toxic to mycobacteria, especially when its gene was expressed under the control of the potent constitutive promoter of *hsp60* [65, 66]. dSpyCas9 is a nuclease activity devoted form of SpyCas9 endonuclease. It was used in early experiments to inhibit gene expression [67]. It should be noted that the dSpyCas9 gene was expressed under the control of a tetracycline-regulated promoter to avoid toxicity due to dSpyCas9 overproduction. Nuclease SpyCas9 and its derivatives with nickase activity or without nuclease activities were found to be highly toxic to Gram-positive *Corynebacterium glutamicum* with a GC-rich genome [68]. It seems natural to assume that SpyCas9 toxicity in bacteria with GC-rich genomes is associated with nonspecific DNA-binding activity due to the GC-rich respective PAM. The assumption is questionable because SpyCas9 remains toxic when its PAM-binding arginine residues are mutated [66]. Protein–protein interactions may be responsible for SpyCas9 toxicity to a substantial effect. Recent studies with mammalian cells showed that the wild-type SpyCas9 or dSpyCas9 interacts with the Ku78 subunit of the DNA-dependent kinase complex and thus negatively affects activity of DNA repair by nonhomologous end joining (NHEJ) [69], thus stimulating spontaneous mutagenesis. Natural type II-A CRISPR/Cas9 systems were found to inhibit NHEJ activity in bacterial cells [70]. The findings indicate that SpyCas9 is capable of interacting with cell proteins to distort the normal cell processes. A contribution of other factors to SpyCas9 toxicity cannot be excluded as well. For example, toxicity of the CRISPR/SpyCas9 system in *C. glutamicum* cells was recently reduced by using the potent *E. coli* *rrnB* terminator for sgRNA and RecT recombinase of the Rac prophage in combination with a donor DNA fragment as an oligonucleotide [71]. It is possible to assume by analogy that the *S. pyogenes* sgRNA terminator fails to function in mycobacteria and that mycobacterial cells are incapable to efficiently repair the DNA breaks introduced by SpyCas9, although possessing three double-strand break repair pathways [72].

In view of CRISPR/SpyCas9 toxicity, alternative systems were sought to edit the mycobacterial genome. A model system with the *Renilla* luciferase reporter gene was used to study repressor activity of 11 SpyCas9 orthologs that lacked nuclease activity and belonged to type II-A and type II-C CRISPR/Cas systems [66]. A tetracycline-regulated promoter was used to control Cas endonuclease gene expression. The *S. thermophilus* dSt1Cas9 protein was identified as the most efficient and nontoxic repressor in the system. Experiments with inhibition of mycobacterial gene expression confirmed the results obtained with the model system [66]. Thus, St1Cas9 is a good candidate to be used in experiments with genome editing. However, constitutive expression of the St1Cas9 gene under the control of the *hsp60* promoter was found to be toxic to *M. smegmatis* [65]. Controlled expression of the St1-

Cas9 gene is therefore necessary to ensure, for example, by using a tetracycline-regulated promoter when designing a St1Cas9-based system for genome editing in mycobacteria. Meijers et al. [73] performed such experiments. St1Cas9 toxicity was reduced not only by using a tetracycline-regulated promoter, but also by expressing St1Cas9 from a single copy of its gene, which was inserted in the mycobacterial genome with phage L5 integrase. The optimal PAM sequence is 5'-NNAGAAW-3' (where degenerate W is A or T) for this editing tool and is extremely rare in the GC-rich mycobacterial genome. However, PAMs with minor differences from the consensus can be recognized by St1Cas9 occurring at higher concentrations in vitro [74]. In agreement with this finding, St1Cas9 was shown to recognize suboptimal PAMs with the sequences 5'-NNGGAA-3' and 5'-NNAGCAT-3', which substantially increases the number of potential targets in the *M. marium* and *M. tuberculosis* genomes. St1Cas9 was successfully used to construct target gene knockouts (the efficiency was approximately 50%) and gene deletions (the efficiency was 22%). Whole-genome sequencing showed that short-term induction of St1Cas9 expression for 1 h generates virtually no off-target mutations. However, undesirable mutations were detected after longer (for 10 days) expression. Promoter leakage was additionally observed for the regulated promoter used in the study; i.e., edited colonies were produced even in the absence of an inducer. In view of background expression of St1Cas9 and its potential to produce off-target mutations upon long-term expression, it was proposed that the St1Cas9 gene be removed from the mycobacterial genome by replacing it with the red fluorescent protein tdTomato reporter gene via homologous recombination [73].

The V-A type system derived from *F. novicida* was the first to be used for genome editing purposes in mycobacteria [75], as well as in *C. glutamicum* [68]. Double-strand breaks introduced in DNA by Fno-Cas12a were found to be lethal in mycobacteria. The finding confirmed the assumption that mycobacterial repair pathways are incapable of efficiently repairing the DNA breaks introduced by heterologous CRISPR/Cas systems. Recombinase of mycobacteriophage Che9c and a recombination donor DNA fragment were used to overcome toxicity of double-strand breaks introduced in mycobacteria by Fno-Cas12a [75]. The green fluorescent protein reporter gene was used to estimate the efficiency of genome editing in *M. smegmatis*. The efficiency of site-directed mutagenesis in *M. smegmatis* was 80%, and an efficiency of 37–75% was observed for fragment insertion and deletion via recombination with double-stranded DNA fragments of approximately 1000 bp. Single-stranded oligonucleotides of 59 and 79 nt facilitated insertions and deletions of no more than 7 and 20 nt, respectively, with an efficiency of 70–80%; the efficiencies achieved with 418- and 1000-bp oligonucleotides were 17.4 and 8.2%, respectively.

Fourteen Cas proteins were tested as tools for editing the mycobacterial genome [65]. Expression of genes for the Cas proteins of *Treponema denticola* (TdCas9), *Neisseria meningitidis* (NmCas9), and *E. tularensis (novicida)* (FnCpf1) under the control of the potent constitutive *hsp60* promoter was observed to exert no effect on the *M. smegmatis* growth rate. When the FnCpf1 gene was optimized according to the codon usage in *M. smegmatis*, the product was toxic to mycobacteria. It was assumed that high-level expression of optimized genes leads to Cas accumulation and increases the frequency of Cas-induced nonspecific single-strand DNA breaks to a critical level [75]. The mRNA of the *FnCpf1_cg* gene optimized for expression in *C. glutamicum* is apparently translated at a low rate and thus prevents FnCpf1_cg from accumulating in dangerous amounts. Of the three Cas proteins that were nontoxic in mycobacteria, only FnCpf1_cg showed high activity with various guide RNAs and ensured a high efficiency (up to 79%) of mycobacterial genome editing. It should be noted that the NHEJ pathway was utilized to edit the mycobacterial genome in that work [65], while homologous recombination was used in the study considered above [75]. Because known NHEJ genes (*ku* and *ligD*) served as targets, it is possible to assume that other known alternative pathways facilitated double-strand break repair [72].

Base editors were recently added to the CRISPR/Cas systems available for genome engineering in mycobacteria. Ding et al. [76] designed the MtbCBE two-plasmid system that utilizes a cytidine editor to efficiently change the *M. tuberculosis* genome. In the system, one of the plasmids codes for protein inhibitors of the RecA- and NucS-dependent DNA repair pathways, while the other plasmid codes for *S. thermophilus* nCas9_{Sth1} nickase fused with APOBEC1 cytidine deaminase and the uracil-DNA glycosylase inhibitor (UGI). The study showed that deaminated base repair involves not only uracil-DNA glycosylases, but also the homologous recombination (RecA is a key protein) and mismatch (NucS nuclease is a key protein) repair pathways, which substantially reduce the extent of deaminase-dependent genome editing. The plasmid that carries the genes for repair pathway inhibitors therefore serves to improve the efficiency of base editing.

The set of potential targets of the editor is limited to the relatively long Cas9_{Sth1} PAM and cytidine-to-thymine substitutions. New base editors were recently designed to replace cytidines with guanines [77] or adenines [78]. The editors include uracil-DNA glycosylase, which excises a deaminated base to produce an apyrimidinic site. Further repair of the apyrimidinic site mostly generates the C → A or C → G transversions. Adaptation of similar base editors to mycobacteria are now in progress and is expected to substantially extend the range of mutations possible to introduce.

Class I type III-A CRISPR/Cas system was found in the genomes of pathogenic mycobacteria, such as *M. tuberculosis* and *M. bovis*, by a bioinformatics analysis [79, 80]. In *M. tuberculosis* H37Rv, the system includes nine genes for Csm2-6, Cas10, Cas6, Cas1, and Cas2 and two CRISPR loci. Cas6 nuclease initiates maturation of crRNA by cleaving its precursor within repeats. The resulting intermediate products have 8 nt of the repeat at the 5' end, a spacer, and a full-length repeat at the 3' end [81]. A ribonucleoprotein complex is assembled on the intermediate products to include the Csm proteins and Cas10 nuclease, while Cas6 is displaced. Then part of the repeat is cleaved from the 3' end to yield mature crRNA [81]. It should be noted that both DNA (Cas10 exerts nuclease activity in this case) and RNA (Csm3 exerts nuclease activity) can be used as targets by a type III-A system in contrast to type II-A and V-A systems considered above [82].

Early studies of the system structure and, mostly, the repeats and spacers of CRISPR loci assumed that the elements are suitable for genotyping mycobacterial strains in epidemiological research [83, 84]. It was not until recently that the mechanism of function of the CRISPR/Cas system and the expression regulation of its components came to be studied in mycobacteria [85, 86]. The system was shown to be active against foreign DNA elements, and crRNA biogenesis was investigated in detail [85, 87]. The *M. tuberculosis* CRISPR system produced in an *E. coli* heterologous expression system displayed activity against both DNA and RNA targets both in vivo and as a system reconstructed from purified proteins [88]. Studies of the structure and mechanisms of function made it possible to reprogram the endogenous CRISPR/Cas system for editing the cognate *M. tuberculosis* genome [89], as was earlier the case, for example, with endogenous CRISPR/Cas systems of clostridia [90, 91]. For the purpose, *M. tuberculosis* was transformed with a plasmid that coded for guide RNAs against target genome regions. Experiments with structurally various recombination donors demonstrated the possibility of inactivating genes and inserting reporter genes in the *M. tuberculosis* genome [89]. Moreover, a method was developed to inhibit gene expression by targeting the endogenous CRISPR/Cas10 system to mRNA. As demonstrated earlier, mRNA, rather than DNA, is hydrolyzed by the effector Cas10 when the sequence 5'-GAAAC-3' is introduced in the 5' end of the crRNA spacer and is complementary to mRNA [92]. The approach was used to efficiently inhibit expression of individual genes (*katG*, *dcD*, or *esxT*) and several genes simultaneously (*lpqE*, *katG*, and *inhA*). In addition, screening studies identified the genes that affect the *M. tuberculosis* reproduction rate in culture and within macrophages [89]. Thus, it was demonstrated that mycobacterial endogenous CRISPR/Cas systems are principally possible to use to edit the cognate genome,

to regulate expression of individual genes or gene sets, and to carry out screening studies.

CONCLUSIONS

Substantial progress was made in studying mycobacterial genetics over the past years, and molecular tools were developed to allow efficient targeted genetic manipulation in mycobacteria. Targeted changes in the mycobacterial genome are possible to introduce by the methods considered in the review, as is necessary for various biomedical applications.

FUNDING

D.S. Karpov acknowledges support from the Ministry of Science and Higher Education of the Russian Federation (contract no. 075-10-2021-113 in a digital budget system, project ID RF-193021X0001).

COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that they have no conflict of interest. This article does not contain any studies involving animals or human subjects performed by any of the authors.

OPEN ACCESS

This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- Wei W., Wang F., Fan S., Wei D. 2010. Inactivation and augmentation of the primary 3-ketosteroid- $\Delta 1$ -dehydrogenase in *Mycobacterium neoaurum* NwIB-01: bio-transformation of soybean phytosterols to 4-androstene-3,17-dione or 1,4-androstadiene-3,17-dione. *Appl. Environ. Microbiol.* **76**, 4578–4582.
- Sundarsingh J.A.T., Ranjitha J., Rajan A., Shankar V. 2020. Features of the biochemistry of *Mycobacterium smegmatis*, as a possible model for *Mycobacterium tuberculosis*. *J. Infect. Public Health.* **13**, 1255–1264.
- Philips J.A., Ernst J.D. 2012. Tuberculosis pathogenesis and immunity. *Annu. Rev. Pathol. Mech. Dis.* **7**, 353–384.
- Choudhary E., Lunge A., Agarwal N. 2016. Strategies of genome editing in mycobacteria: achievements and challenges. *Tuberculosis.* **98**, 132–138.
- Zhang Y.J., Ioerger T.R., Huttenhower C., Long J.E., Sassetti C.M., Sacchettini J.C., Rubin E.J. 2012. Global assessment of genomic regions required for growth in *Mycobacterium tuberculosis*. *PLoS Pathog.* **8**, e1002946.
- Lamrabet O., Drancourt M. 2012. Genetic engineering of *Mycobacterium tuberculosis*: a review. *Tuberculosis.* **92**, 365–376.
- Borgers K., Vandewalle K., Festjens N., Callewaert N. 2019. A guide to *Mycobacterium* mutagenesis. *FEBS J.* **286**, 3757–3774.
- Kana B.D., Mizrahi V. 2004. Molecular genetics of *Mycobacterium tuberculosis* in relation to the discovery of novel drugs and vaccines. *Tuberculosis.* **84**, 63–75.
- Gopinath K., Warner D.F., Mizrahi V. 2015. Targeted gene knockout and essentiality testing by homologous recombination. In *Mycobacteria Protocols. Methods in Molecular Biology*. Parish T., Roberts D.M., Eds. New York: Humana Press, **1285**, 131–149.
- Arnold F.M., Hohl M., Remm S., Koliwer-Brandl H., Adenau S., Chusri S., Sander P., Hilbi H., Seeger M.A. 2018. A uniform cloning platform for mycobacterial genetics and protein production. *Sci. Rep.* **8**, 9539.
- Safi H., Sayers B., Hazbón M.H., Alland D. 2008. Transfer of *embB* codon 306 mutations into clinical *Mycobacterium tuberculosis* strains alters susceptibility to ethambutol, isoniazid, and rifampin. *Antimicrob. Agents Chemother.* **52**, 2027–2034.
- Brazier B., McShane H. 2020. Towards new TB vaccines. *Semin. Immunopathol.* **42**, 315–331.
- Bastos R.G., Alzan H.F., Rathinasamy V.A., Cooke B.M., Dellagostin O.A., Barletta R.G., Suarez C.E. 2022. Harnessing *Mycobacterium bovis* BCG trained immunity to control human and bovine babesiosis. *Vaccines.* **10**, 1–18.
- Loxton A.G., Knaul J.K., Grode L., Gutschmidt A., Meller C., Eisele B., Johnstone H., Van Der Spuy G., Maertzdorf J., Kaufmann S.H.E., Hesselung A.C., Walzl G., Cotton M.F. 2017. Safety and immunogenicity of the recombinant *Mycobacterium bovis* BCG vaccine VPM1002 in HIV-unexposed newborn infants in South Africa. *Clin. Vaccine Immunol.* **24**, 1–16.
- Grode L., Seiler P., Baumann S., Hess J., Brinkmann V., Eddine A.N., Mann P., Goosmann C., Bandermann S., Smith D., Bancroft G.J., Reyat J.M., Van Soolingen D., Raupach B., Kaufmann S.H.E. 2005. Increased vaccine efficacy against tuberculosis of recombinant *Mycobacterium bovis* bacille Calmette–Guérin mutants that secrete listeriolysin. *J. Clin. Invest.* **115**, 2472–2479.
- Arbues A., Aguilo J.I., Gonzalo-Asensio J., Marinova D., Uranga S., Puentes E., Fernandez C., Parra A., Cardona P.J., Vilaplana C., Ausina V., Williams A., Clark S., Malaga W., Guilhot C., Gicquel B., Martin C. 2013. Construction, characterization and preclinical evaluation of MTBVAC, the first live-attenuated *M. tuberculosis*-based vaccine to enter clinical trials. *Vaccine.* **31**, 4867–4873.
- Hinds J., Mahenthiralingam E., Kempell K.E., Duncan K., Stokes R.W., Parish T., Stoker N.G. 1999. En-

- hanced gene replacement in mycobacteria. *Microbiology*. **145**, 519–527.
18. Malaga W., Perez E., Guilhot C. 2003. Production of unmarked mutations in mycobacteria using site-specific recombination. *FEMS Microbiol. Lett.* **219**, 261–268.
 19. Stephan J., Stemmer V., Niederweis M. 2004. Consecutive gene deletions in *Mycobacterium smegmatis* using the yeast FLP recombinase. *Gene*. **343**, 181–190.
 20. Song H., Niederweis M. 2007. Functional expression of the FLP recombinase in *Mycobacterium bovis* BCG. *Gene*. **399**, 112–119.
 21. Song H., Wolschendorf F., Niederweis M. 2009. Construction of unmarked deletion mutants. In *Mycobacteria Protocols. Methods in Molecular Biology*. Parish T., Brown A., Eds. Totowa, N.J.: Humana Press, **465**, 279–295.
 22. Cascioferro A., Boldrin F., Serafini A., Provvedi R., Palù G., Manganelli R. 2010. Xer site-specific recombination, an efficient tool to introduce unmarked deletions into mycobacteria. *Appl. Environ. Microbiol.* **76**, 5312–5316.
 23. Yang F., Tan Y., Liu J., Liu T., Wang B., Cao Y., Qu Y., Lithgow T., Tan S., Zhang T. 2014. Efficient construction of unmarked recombinant mycobacteria using an improved system. *J. Microbiol. Methods*. **103**, 29–36.
 24. Parish T., Stoker N.G. 2000. Use of flexible cassette method to generate a double unmarked *Mycobacterium tuberculosis* tlyA plcABC mutant by gene replacement. *Microbiology*. **146**, 1969–1975.
 25. Brown A.C. 2021. Gene switching and essentiality testing. In *Mycobacteria Protocols. Methods in Molecular Biology*. Parish T., Kumar A., Eds. New York: Humana Press, **2314**, 285–299.
 26. Onwueme K.C., Ferreras J.A., Buglino J., Lima C.D., Quadri L.E.N. 2004. Mycobacterial polyketide-associated proteins are acyltransferases: proof of principle with *Mycobacterium tuberculosis* PapA5. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 4608–4613.
 27. Grigorov A., Bychenko O., Salina E.G., Skvortsova Y., Mazurova A., Skvortsov T., Kaprelyants A., Azhikina T. 2021. Small rna f6 provides *Mycobacterium smegmatis* entry into dormancy. *Int. J. Mol. Sci.* **22** (21), 11536.
 28. Barkan D., Stallings C.L., Glickman M.S. 2011. An improved counterselectable marker system for mycobacterial recombination using galK and 2-deoxy-galactose. *Gene*. **470**, 31–36.
 29. Pashley C.A., Parish T., McAdam R.A., Duncan K., Stoker N.G. 2003. Gene replacement in mycobacteria by using incompatible plasmids. *Appl. Environ. Microbiol.* **69**, 517–523.
 30. Bardarov S., Bardarov S., Pavelka M.S., Sambandamurthy V., Larsen M., Tufariello J.A., Chan J., Hatfull G., Jacobs W.R. 2002. 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.
 31. Ang M.L.T., Siti Z.Z.R., Shui G., Dianišková P., Madacki J., Lin W., Koh V.H.Q., Gomez J.M.M., Sudarkodi S., Bendt A., Wenk M., Mikušová K., Koruláková J., Pethe K., Alonso S. 2014. An ethA-ethR-deficient *Mycobacterium bovis* BCG mutant displays increased adherence to mammalian cells and greater persistence in vivo, which correlate with altered mycolic acid composition. *Infect. Immun.* **82**, 1850–1859.
 32. Bourai N., Jacobs W.R., Narayanan S. 2012. Deletion and overexpression studies on DacB2, a putative low molecular mass penicillin binding protein from *Mycobacterium tuberculosis* H37Rv. *Microb. Pathog.* **52**, 109–116.
 33. Jain P., Hsu T., Arai M., Biermann K., Thaler D.S., Nguyen A., González P.A., Tufariello J.M., Kriakov J., Chen B., Larsen M.H., Jacobs W.R. 2014. Specialized transduction designed for precise high-throughput unmarked deletions in *Mycobacterium tuberculosis*. *MBio*. **5** (3), e01245-14.
 34. Alonso M.N., Malaga W., Mc Neil M., Jackson M., Romano M.I., Guilhot C., Santangelo M.P. 2020. Efficient method for targeted gene disruption by homologous recombination in *Mycobacterium avium* subspecies paratuberculosis. *Res. Microbiol.* **171**, 203–210.
 35. van Kessel J.C., Hatfull G.F. 2007. Recombineering in *Mycobacterium tuberculosis*. *Nat. Methods*. **4**, 147–152.
 36. Mao X., Yan M., Zhu H., Guo X., Sun Y. 2016. Efficient and simple generation of multiple unmarked gene deletions in *Mycobacterium smegmatis*. *Sci. Rep.* **14** (6), 22922.
 37. Shenkerman Y., Elharar Y., Vishkautzan M., Gur E. 2014. Efficient and simple generation of unmarked gene deletions in *Mycobacterium smegmatis*. *Gene*. **533**, 374–378.
 38. van Kessel J.C., Hatfull G.F. 2008. Efficient point mutagenesis in mycobacteria using single-stranded DNA recombineering: characterization of antimycobacterial drug targets. *Mol. Microbiol.* **67**, 1094–1107.
 39. Niranjala Muttucumar D.G., Parish T. 2004. The molecular biology of recombination in mycobacteria: what do we know and how can we use it? *Curr. Issues Mol. Biol.* **6** (2), 145–157.
 40. Freitas-Vieira A., Anes E., van Kessel Moniz-Pereira J. 1998. The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the tRNA (Ala) gene of *Mycobacterium* spp. *Microbiology*. **144**, 3397–3406.
 41. Pedulla M.L., Lee M.H., Lever D.C., Hatfull G.F. 1996. A novel host factor for integration of mycobacteriophage L5. *Proc. Natl. Acad. Sci. U. S. A.* **93** (26), 15411–15416.
 42. Mediavilla J., Jain S., Kriakov J., Ford M.E., Duda R.L., Jacobs W.R., Hendrix R.W., Hatfull G.F. 2002. Genome organization and characterization of mycobacteriophage Bxb1. *Mol. Microbiol.* **38**, 955–970.
 43. Springer B., Sander P., Sedlacek L., Ellrott K., Böttger E.C. 2001. Instability and site-specific excision of integration-proficient mycobacteriophage L5 plasmids: development of stably maintained integrative vectors. *Int. J. Med. Microbiol.* **290**, 669–675.
 44. Bibb L.A., Hatfull G.F. 2002. Integration and excision of the *Mycobacterium tuberculosis* prophage-like element, ϕ Rv1. *Mol. Microbiol.* **45**, 1515–1526.
 45. Kim A.I., Ghosh P., Aaron M.A., Bibb L.A., Jain S., Hatfull G.F. 2003. Mycobacteriophage Bxb1 integrates into the *Mycobacterium smegmatis* groEL1 gene. *Mol. Microbiol.* **50**, 463–473.

46. Lewis J.A., Hatfull G.F. 2003. Control of directionality in L5 integrase-mediated site-specific recombination. *J. Mol. Biol.* **326**, 805–821.
47. Lee M.H., Hatfull G.F. 1993. Mycobacteriophage L5 integrase-mediated site-specific integration in vitro. *J. Bacteriol.* **175** (21), 6836–6841.
48. Peña C.E.A., Kahlenberg J.M., Hatfull G.F. 1998. The role of supercoiling in mycobacteriophage L5 integrative recombination. *Nucl. Acids Res.* **26**, 4012–4018.
49. Peña C.E.A., Lee M.H., Pedulla M.L., Hatfull G.F. 1997. Characterization of the mycobacteriophage L5 attachment site, attP. *J. Mol. Biol.* **266** (1), 76–92.
50. Stark W.M. 2017. Making serine integrases work for us. *Curr. Opin. Microbiol.* **38**, 130–136.
51. Landy A. 2015. The λ integrase site-specific recombination pathway. *Microbiol. Spectr.* **3** (2), MDNA3-0051-2014.
52. Stark W.M. 2014. The serine recombinases. *Microbiol. Spectr.* **2** (6), MDNA3-0046-2014.
53. Murphy K.C., Nelson S.J., Nambi S., Papavinasandaram K., Baer C.E., Sasseti C.M. 2018. Orbit: a new paradigm for genetic engineering of mycobacterial chromosomes. *MBio.* **9** (6), e01467-18.
54. Murphy K.C. 2021. Oligo-mediated recombineering and its use for making SNPs, knockouts, insertions, and fusions in *Mycobacterium tuberculosis*. In *Mycobacteria Protocols. Methods in Molecular Biology*, Parish T., Kumar A., Eds. New York: Humana Press, **2314**, 301–321.
55. Bibb L.A., Hancox M.I., Hatfull G.F. 2005. Integration and excision by the large serine recombinase ϕ Rv1 integrase. *Mol. Microbiol.* **55**, 1896–1910.
56. Seoane A., Navas J., García Lobo J.M. 1997. Targets for pSAM2 integrase-mediated site-specific integration in the *Mycobacterium smegmatis* chromosome. *Microbiology.* **143**, 3375–3380.
57. Mazodier P., Thompson C., Bocard F. 1990. The chromosomal integration site of the *Streptomyces* element pSAM2 overlaps a putative tRNA gene conserved among actinomycetes. *Mol. Gen. Genet.* **222** (2–3), 431–434.
58. Marraffini L.A. 2015. CRISPR-Cas immunity in prokaryotes. *Nature.* **526**, 1–13.
59. Bayat H., Modarressi M.H., Rahimpour A. 2018. The conspicuity of CRISPR-Cpf1 system as a significant breakthrough in genome editing. *Curr. Microbiol.* **75** (1), 107–115.
60. Roy B., Zhao J., Yang C., Luo W., Xiong T., Li Y., Fang X., Gao G., Singh C.O., Madsen L., Zhou Y., Kristiansen K. 2018. CRISPR/Cascade 9-mediated genome editing—challenges and opportunities. *Front. Genet.* **9**, 240.
61. Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. 2012. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* **337**, 816–821.
62. Zetsche B., Gootenberg J.S., Abudayyeh O.O., Slaymaker I.M., Makarova K.S., Essletzbichler P., Volz S.E., Joung J., van der Oost J., Regev A., Koonin E.V., Zhang F. 2015. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* **163** (3), 759–771.
63. Leenay R.T., Maksimchuk K.R., Slotkowski R.A., Agrawal R.N., Gomaa A.A., Briner A.E., Barrangou R., Beisel C.L. 2016. Identifying and visualizing functional PAM diversity across CRISPR-Cas systems. *Mol. Cell.* **62** (1), 137–147.
64. Swarts D.C., Jinek M. 2018. Cas9 versus Cas12a/Cpf1: Structure–function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA.* **9** (5), e1481.
65. Sun B., Yang J., Ye R.D., Jiang Y., Chen D., Yang S. 2018. A CRISPR-Cpf1-assisted non-homologous end joining genome editing system of *Mycobacterium smegmatis*. *Biotechnol. J.* **13** (9), e1700588.
66. Rock J.M., Hopkins F.F., Chavez A., Diallo M., Chase M.R., Gerrick E.R., Pritchard J.R., Church G.M., Rubin E.J., Sasseti C.M., Schnappinger D., Sarah M.F. 2017. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat. Microbiol.* **2**, 16274.
67. Singh A.K., Carette X., Potluri L.P., Sharp J.D., Xu R.F., Priscic S., Husson R.N. 2016. Investigating essential gene function in *Mycobacterium tuberculosis* using an efficient CRISPR interference system. *Nucl. Acids Res.* **44**, e143.
68. Jiang Y., Qian F., Yang J., Liu Y., Dong F., Xu C., Sun B., Chen B., Xu X., Li Y., Wang R., Yang S. 2017. CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* **8**, 15179
69. Xu S., Kim J., Tang Q., Chen Q., Liu J., Xu Y., Fu X. 2020. Cas9 is a genome mutator by directly disrupting DNA-PK dependent DNA repair pathway. *Protein Cell.* **11** (5), 352–365.
70. Bernheim A., Calvo-Villamanan A., Basier C., Cui L., Rocha E.P.C., Touchon M., Bikard D. 2017. Inhibition of NHEJ repair by type II-A CRISPR-Cas systems in bacteria. *Nat. Commun.* **8** (1), 2094.
71. Liu J., Wang Y., Lu Y., Zheng P., Sun J., Ma Y. 2017. Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Fact.* **16**, 205.
72. Singh A. 2017. Guardians of the mycobacterial genome: a review on DNA repair systems in *Mycobacterium tuberculosis*. *Microbiology.* **163**, 1740–1758.
73. Meijers A.S., Troost R., Ummels R., Maaskant J., Speer A., Nejentsev S., Bitter W., Kuijl C.P. 2020. Efficient genome editing in pathogenic mycobacteria using *Streptococcus thermophilus* CRISPR1-Cas9. *Tuberculosis (Edinb.)*. **124**, 101983.
74. Karvelis T., Gasiunas G., Young J., Bigelyte G., Silanskas A., Cigan M., Siksnys V. 2015. Rapid characterization of CRISPR-Cas9 protospacer adjacent motif sequence elements. *Genome Biol.* **16**, 253.
75. Yan M.-Y., Yan H.-Q., Ren G.-X., Zhao J.-P., Guo X.-P., Sun Y.-C. 2017. CRISPR-Cas12a-assisted recombineering in bacteria. *Appl. Environ. Microbiol.* **83**, e00947–00917.
76. Ding X.Y., Li S.S., Geng Y.M., Yan M.Y., Li G.B., Zhang G.L., Sun Y.C. 2021. Programmable base editing in *Mycobacterium tuberculosis* using an engineered CRISPR RNA-guided cytidine deaminase. *Front. Genome Ed.* **3**, 734436.

77. Kurt I.C., Zhou R., Iyer S., Garcia S.P., Miller B.R., Langner L.M., Grunewald J., Joung J.K. 2021. CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* **39** (1), 41–46.
78. Zhao D., Li J., Li S., Xin X., Hu M., Price M.A., Rosser S.J., Bi C., Zhang X. 2021. Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* **39** (1), 35–40.
79. He L., Fan X., Xie J. 2012. Comparative genomic structures of *Mycobacterium* CRISPR-Cas. *J. Cell Biochem.* **113** (7), 2464–2473.
80. Supply P., Marceau M., Mangenot S., Roche D., Rouanet C., Khanna V., Majlessi L., Criscuolo A., Tap J., Pawlik A., Fiette L., Orgeur M., Fabre M., Parmentier C., Frigui W., Simeone R., Boritsch E.C., Debie A.S., Willery E., Walker D., Quail M.A., Ma L., Bouchier C., Salvignol G., Sayes F., Cascioferro A., Seemann T., Barbe V., Locht C., Gutierrez M.C., Leclerc C., Bentley S.D., Stinear T.P., Brisse S., Medigue C., Parkhill J., Cruveiller S., Brosch R. 2013. Genomic analysis of smooth tubercle bacilli provides insights into ancestry and pathoadaptation of *Mycobacterium tuberculosis*. *Nat. Genet.* **45** (2), 172–179.
81. Liu T., Pan S., Li Y., Peng N., She Q. 2018. Type III CRISPR-Cas system: introduction and its application for genetic manipulations. *Curr. Iss. Mol. Biol.* **26**, 1–14.
82. Pyenson N.C., Marraffini L.A. 2017. Type III CRISPR-Cas systems: when DNA cleavage just isn't enough. *Curr. Opin. Microbiol.* **37**, 150–154.
83. Gomgnimbou M.K., Abadia E., Zhang J., Refregier G., Panaiotov S., Bachiyska E., Sola C. 2012. “Spoligotyping,” a dual-priming oligonucleotide-based direct-hybridization assay for tuberculosis control with a multianalyte microbead-based hybridization system. *J. Clin. Microbiol.* **50**, 3172–3179.
84. Botelho A., Canto A., Leao C., Cunha M.V. 2015. Clustered regularly interspaced short palindromic repeats (CRISPRs) analysis of members of the *Mycobacterium tuberculosis* complex. *Methods Mol. Biol.* **1247**, 373–389.
85. Zhang Y., Yang J., Bai G. 2018. Regulation of the CRISPR-associated genes by Rv2837c (CnpB) via an Orn-like activity in TB complex mycobacteria. *J. Bacteriol.* **200** (8), e00743–17.
86. Huang Q., Luo H., Liu M., Zeng J., Abdalla A.E., Duan X., Li Q., Xie J. 2016. The effect of *Mycobacterium tuberculosis* CRISPR-associated Cas2 (Rv2816c) on stress response genes expression, morphology and macrophage survival of *Mycobacterium smegmatis*. *Infect. Genet. Evol.* **40**, 295–301.
87. Wei W., Zhang S., Fleming J., Chen Y., Li Z., Fan S., Liu Y., Wang W., Wang T., Liu Y., Ren B., Wang M., Jiao J., Chen Y., Zhou Y., Zhou Y., Gu S., Zhang X., Wan L., Chen T., Zhou L., Chen Y., Zhang X.E., Li C., Zhang H., Bi L. 2019. *Mycobacterium tuberculosis* type III-A CRISPR/Cas system crRNA and its maturation have atypical features. *FASEB J.* **33** (1), 1496–1509.
88. Gruschow S., Athukoralage J.S., Graham S., Hoogboom T., White M.F. 2019. Cyclic oligoadenylate signalling mediates *Mycobacterium tuberculosis* CRISPR defence. *Nucl. Acids Res.* **47** (17), 9259–9270.
89. Rahman K., Jamal M., Chen X., Zhou W., Yang B., Zou Y., Xu W., Lei Y., Wu C., Cao X., Tyagi R., Ahsan Naem M., Lin D., Habib Z., Peng N., Fu Z.F., Cao G. 2021. Reprogramming *Mycobacterium tuberculosis* CRISPR system for gene editing and genome-wide RNA interference screening. *Genomics, Proteomics Bioinf.* S1672-0229 (21)00249-7. <https://doi.org/10.1101/2020.03.09.983494>
90. Pyne M.E., Bruder M.R., Moo-Young M., Chung D.A., Chou C.P. 2016. Harnessing heterologous and endogenous CRISPR-Cas machineries for efficient markerless genome editing in *Clostridium*. *Sci. Rep.* **6**, 25666.
91. Zhang J., Zong W., Hong W., Zhang Z.T., Wang Y. 2018. Exploiting endogenous CRISPR-Cas system for multiplex genome editing in *Clostridium tyrobutyricum* and engineer the strain for high-level butanol production. *Metab. Eng.* **47**, 49–59.
92. You L., Ma J., Wang J., Artamonova D., Wang M., Liu L., Xiang H., Severinov K., Zhang X., Wang Y. 2019. Structure studies of the CRISPR-Csm complex reveal mechanism of co-transcriptional interference. *Cell.* **176** (1–2), 239–253.

Translated by T. Tkacheva