Molecular Therapy
Nucleic Acids
Original Article

A CRISPR-guided mutagenic DNA polymerase strategy for the detection of antibiotic-resistant mutations in *M. tuberculosis*

Siyuan Feng,[1,2,3,11] Lujie Liang,[1,2,3,11] Cong Shen,[4,5,11] Daixi Lin,[1,2,3] Jiachen Li,[1,2,3] Lingxuan Lyu,[1,2,3] Wanfei Liang,[1,2,3] Lan-lan Zhong,[1,2,3] Gregory M. Cook,[6,7] Yohei Doi,[8,9] Cha Chen,[4,5] and Guo-bao Tian[1,2,3,10]

[1]Program in Pathobiology, The Fifth Affiliated Hospital, Zhongshan School of Medicine, Sun Yat-Sen University, 74 Zhongshan 2nd Road, Guangdong 510080, China; [2]Advanced Medical Technology Center, The First Affiliated Hospital, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan 2nd Road, Guangzhou 510080, China; [3]Key Laboratory of Tropical Diseases Control (Sun Yat-sen University), Ministry of Education, 74 Zhongshan 2nd Road, Guangzhou 510080, China; [4]The Second Clinic Medical College, Guangzhou University of Chinese Medicine, Guangzhou 510006, China; [5]Clinical Laboratory, The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangdong Provincial Hospital of Traditional Chinese Medicine, Guangzhou 510006, China; [6]Department of Microbiology and Immunology, University of Otago, Dunedin, New Zealand; [7]Maurice Wilkins Centre for Molecular Biodiscovery, University of Auckland, Auckland, New Zealand; [8]Departments of Microbiology and Infectious Diseases, Fujita Health University School of Medicine, Aichi 470-1192, Japan; [9]University of Pittsburgh School of Medicine, Pittsburgh, PA 15261, USA; [10]School of Medicine, Xizang Minzu University, Xianyang, Shaanxi 712082, China

A sharp increase in multidrug-resistant tuberculosis (MDR-TB) threatens human health. Spontaneous mutation in essential gene confers an ability of *Mycobacterium tuberculosis* resistance to anti-TB drugs. However, conventional laboratory strategies for identification and prediction of the mutations in this slowly growing species remain challenging. Here, by combining XCas9 nickase and the error-prone DNA polymerase A from *M. tuberculosis*, we constructed a CRISPR-guided DNA polymerase system, CAMPER, for effective site-directed mutagenesis of drug-target genes in mycobacteria. CAMPER was able to generate mutagenesis of all nucleotides at user-defined loci, and its bidirectional mutagenesis at nick sites allowed editing windows with lengths up to 80 nucleotides. Mutagenesis of drug-targeted genes in *Mycobacterium smegmatis* and *M. tuberculosis* with this system significantly increased the fraction of the antibiotic-resistant bacterial population to a level approximately 60- to 120-fold higher than that in unedited cells. Moreover, this strategy could facilitate the discovery of the mutation conferring antibiotic resistance and enable a rapid verification of the growth phenotype-mutation genotype association. Our data demonstrate that CAMPER facilitates targeted mutagenesis of genomic loci and thus may be useful for broad functions such as resistance prediction and development of novel TB therapies.

INTRODUCTION

Mycobacterium tuberculosis, the causative agent of tuberculosis (TB), is the leading cause of death from infectious diseases. Most people with TB (drug-susceptible or drug-resistant) are cured after a medication regimen over 6–24 months.[1,2] However, untimely treatment interruption or inappropriate use of antimicrobial drugs can result in drug resistance and treatment failure.[3,4] Drug-resistant TB, particularly multidrug-resistant TB (MDR-TB, resistant to isoniazid and rifampicin), often requires prolonged treatment with expensive and toxic drugs, thus posing an even higher risk of treatment failure. In 2018, approximately 214,000 deaths from MDR-TB were reported.[5] According to the 2019 global TB report from the World Health Organization, 3.4% of the new TB cases and 18% of the previously treated cases were MDR.[6] Despite the many strategies for controlling the emergence of MDR-TB, recent years have witnessed a growing number of MDR-TB cases, and there is an ongoing urgent need to develop novel treatment strategies for the elimination of MDR-TB.

In *M. tuberculosis*, drug resistance typically occurs because of spontaneous mutations in the gene's product targeted by the anti-TB drug.[2,7] Identification of all drug-resistance-related mutations in a certain gene may contribute to the design of appropriate treatment strategies and diagnostic tools.[7] For example, GeneXpert, a diagnostic tool for the identification of rifampicin-resistant mutated isolates, works on the basis of detailed information about the mutations leading to drug resistance.[8] Rapid identification of drug-resistant mutations facilitates the development of anti-TB drugs.[7,9] However, high-efficiency strategies to identify these mutations remain a challenge.

Continuous culture (passage) and randomized mutagenesis are the two most common laboratory methods used for the verification of

Received 9 November 2021; accepted 8 July 2022;
https://doi.org/10.1016/j.omtn.2022.07.004.

[11]These authors contributed equally

Correspondence: Guo-Bao Tian, Ph.D, Professor, Department of Microbiology, Zhongshan School of Medicine, Sun Yat-sen University, 74 Zhongshan 2nd Road, Guangzhou 510080, China.
E-mail: tiangb@mail.sysu.edu.cn

drug-resistance-related mutations. Continuous culture has been applied for the identification and prediction of drug-resistance mutations and is an effective way to isolate target mutants under antibiotic selection pressure.[10,11] However, the mutation genotypes are very limited with this method because the spontaneous mutation rate in *M. tuberculosis* is extremely low (2×10^{-10}).[12] Another strategy, randomized mutagenesis, is mainly based on error-prone PCR and requires complicated procedures, such as deletion of the chromosomal target gene and replacement with the mutated pool through homologous recombination (HR).[13] Several genetic engineering strategies have been modified to improve the efficiency of HR in *M. tuberculosis*, such as using a suicide vector for one-step or two-step allelic exchange.[14] However, these methods are limited by the highly activated illegitimate recombination pathway in *M. tuberculosis*,[15] the inefficient HR pathway,[16] and the slow growth of the bacterium, such that 3–6 months is usually required to obtain a single recombinant mutant.[14] Hence, developing advanced strategies to rapidly identify resistance-associated mutations is becoming increasingly important to support the development of appropriate TB treatment strategies and diagnostic tools.

CRISPR-guided DNA polymerase has the potential to facilitate targeted diversification of genes of interest in *Escherichia coli* and *Saccharomyces cerevisiae*.[17,18] This approach incorporates expression of a single guide RNA (sgRNA) targeting the gene of interest in bacteria harboring Cas9 nickase (nCas9)-DNA polymerase (PolA) fusion protein. The Cas9 nickase first recognizes and cleaves the target site of a specific gene. The gap generated by the nickase binds DNA polymerase and initiates nick translation. The formation of a nick allows the DNA polymerase to repair mismatches on both the leading and lagging strands.[19] If the nick translation is finished by error-prone DNA polymerase, more substitution mutations are introduced into the new DNA fragment, resulting in random substitution mutations in the target region.[17] Therefore, the combination of an sgRNA-guided Cas9 nickase and error-prone DNA polymerase can allow site-directed mutagenesis of target genes *in vivo*. Unlike *E. coli*, *M. tuberculosis* lacks plasmids.[20] It has been convincingly shown by many groups that *M. tuberculosis* clinical isolates become resistant through point mutations in chromosomal genes and that these mutations arise spontaneously at low frequency.[21] Given these features of *M. tuberculosis*, it seems that this CRISPR-guided mutagenic DNA polymerase strategy may appropriately perform targeted mutagenesis for the identification of drug-resistant mutants in *M. tuberculosis*.

In the current study, we used a CRISPR-guided DNA polymerase system in both fast-growing *Mycobacterium smegmatis* and slow-growing *M. tuberculosis*. Through protein structure alignment, we used DNA polymerase A from *M. tuberculosis* to generate an error-prone DNA polymerase. By combining the highly compatible XCas9 nickase, a Cas9 variant recognizing NGN as the protospacer adjacent motif (PAM),[22] with the error-prone DNA polymerase A from *M. tuberculosis*, the CAMPER system was able to introduce random substitution mutations within an 80-bp-long editing window. Moreover, CAMPER not only efficiently induced site-directed mutagenesis in *M. smegmatis* and *M. tuberculosis* but also enabled the detection of the novel resistant mutants that likely involve inhibitor binding regions of a drug-targeted gene. Finally, when combined with high-throughput sequencing, the system successfully facilitated verification of the growth phenotype-mutation genotype association.

RESULTS

Construction of a CRISPR-guided DNA polymerase system in mycobacteria

With inspiration from the method previously reported by Halperin et al.,[17] we aimed at constructing a CRISPR-guided error-prone DNA polymerase system in mycobacteria for random substitution mutations in a target gene. First, we used *M. tuberculosis* DNA polymerase A (PolA) to generate error-prone DNA polymerase. The PolA possesses two functional domains: a flap endonuclease domain and a DNA polymerase domain.[23] Point mutations in the large fragment of DNA polymerase A, the functional domain responsible for polymerization and proofreading, have been established to decrease the fidelity of the enzyme.[24,25] For example, D424A at the metal ion site B of the 3′-5′ exonuclease domain inactivates the proofreading function of *E. coli* PolA and increases the global mutation rate.[24,26] The combination of I709F with D424A, a mutation inactivating the 3′-5′ exonuclease activity of *E. coli* PolA, results in a 400-fold higher mutation rate than that of the wild type. In addition, the A759R substitution mutation increases the error rate of the polymerase in *E. coli*.[27–31] To determine the protein structure of *M. tuberculosis* PolA, we performed structure modeling with *M. smegmatis* PolA (PDB: 6VDE) as a structural template. Through protein structure alignment, we observed a similar structure of the DNA polymerase domain between *E. coli* PolA and *M. tuberculosis* PolA (Figure S1A). Of note, the positions of D424, I709, and A759 in *E. coli* PolA corresponded to D405, I684, and A734 in *M. tuberculosis* PolA, and these amino acids were conserved in *E. coli* PolA and *M. tuberculosis* PolA (Figure S1B), suggesting that D405, I684, and A734 in *M. tuberculosis* PolA were critical for DNA synthesis. We therefore constructed PolA3M, a *M. tuberculosis* PolA variant harboring D405A, I684N, and A734R substitution mutations, to be combined with Cas9 nickase. Since PolA has the activities of 5′-3′ exonuclease and 5′-3′ polymerase, the error-prone DNA polymerase A, after nicking by nCas9, may mutate the DNA fragment at the 5′ end of the nick through nick translation (Figure 1A).

We next designed a platform for gene editing containing an integrated plasmid for the expression of codon-optimized nCas9-PolA3M and a plasmid to express the sgRNA for guiding the system and TetR for regulation (Figure S2 and Table S1). The expression of both nCas9-PolA3M and sgRNA was under the control of an anhydrotetracycline (ATc)-inducible promoter. The editing efficiency of the tool was evaluated by measuring repair of a plasmid-encoded bleomycin resistance marker disabled by an early stop codon. For the system constructed by Cas9 cutting double-stranded DNA, the mutation frequency (7.81×10^{-7}) was almost the same as that of the off-target control. However, for the system constructed with Cas9 nickase, the ratio of BleoR colony-forming units (CFUs)/viable

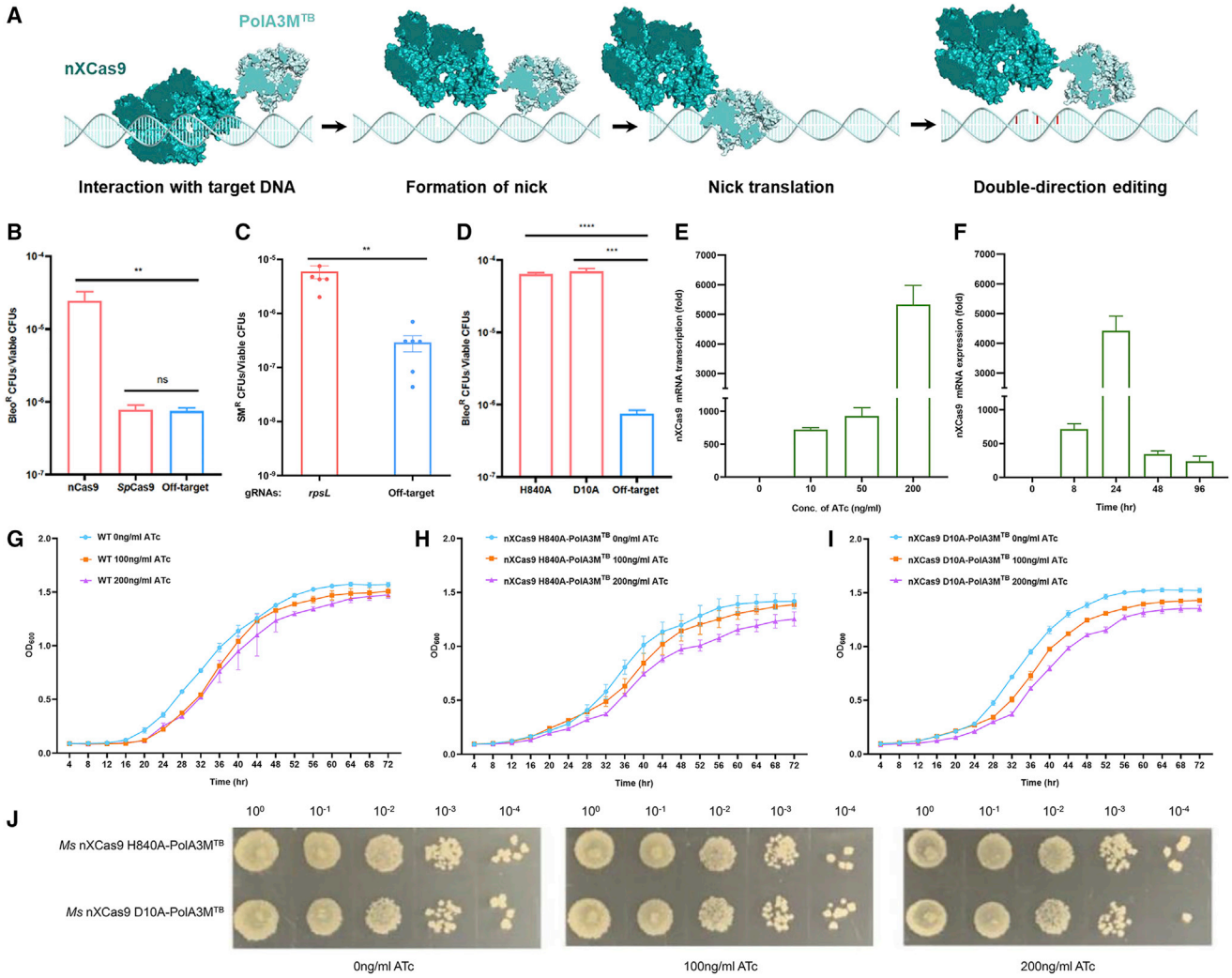

Figure 1. Editing efficiency of XCas9 H840A/D10A combining with PolA3M^{TB}

The mutagenesis tool CAMPER consisting of XCas9 nickase (nXCas9) and error-prone DNA polymerase A from *Mtb* (PolA3M^{TB}) was applied in *Mycobacterium*. (A) The CAMPER system constructed with XCas9 nickase and error-prone DNA polymerase A from *M. tuberculosis* nicked the target locus under the specific guidance of a gRNA and performed bidirectional nick translation for mutagenesis. (B) *M. tuberculosis* error-prone DNA polymerase A (PolA3M^{TB}) fused with nCas9 (*Streptococcus pyogenes* Cas9 containing a single point mutation, D10A) and *Sp*Cas9 (*S. pyogenes* Cas9). Compared with the off-target control, only the expression of PolA3M^{TB} fused with nCas9 led to an increase in bleomycin-resistant CFUs. (C) Evaluation of editing efficiency of CAMPER toward the chromosomal *rpsL* gene in *M. smegmatis*. A 100-μL electroporation culture was subcultured in 7H9 medium for mutant frequency calculations. (D) CAMPER for expression of XCas9 nickase H840A or XCas9 nickase D10A, under specific guidance of a gRNA, resulted in an increased rate of bleomycin-resistant cells compared with that in the off-target control. (E) Expression of CAMPER in *M. smegmatis* was induced with ATc at 0, 10, 50, or 200 ng/mL, and the transcriptional level was determined at 24 h with qPCR. The transcriptional level of MATURETB was highest under induction with 200 ng/mL ATc. (F) With the induction of 200 ng/mL ATc, the transcriptional level of CAMPER in *M. smegmatis* was measured after 8, 24, 48, and 96 h, which was highest at 24 h but decreased at 48 h. (G–I) Growth curves were measured to evaluate the cytotoxicity of CAMPER H840A/D10A with or without ATc induction. The growth curves were measured every 4 h, and the cell density at each time point was presented as the value of OD_{600}. (J) The consequences of nXCas9-PolA3MTB expression were monitored by spotting dilutions of each culture on LB agar. All of the above experiments were performed at least three times with similar results. Error bars indicate SEM for three biological replicates. A two-tailed unpaired t test was performed to determine the statistical significance of the data. ns, no significant difference; $^{**}p < 0.01$; $^{***}p < 0.001$; $^{****}p < 0.0001$.

CFUs for target editing (2.44×10^{-5}) was 32.6-fold higher than that of the off-target control (7.48×10^{-7}) on average (Figure 1B). We also applied this system to target the chromosomal locus of *rpsL*, an essential gene that encodes the S12 ribosomal protein and is the primary site of streptomycin resistance mutations.[32] The K88R mutation of RpsL is associated with high resistance to streptomycin,[33] and an sgRNA targeting this region was designed. The mutant frequency calculations showed that the mutation rate of target editing on K88 (5.98×10^{-6}) was 20.66-fold higher than that of the off-target control (2.89×10^{-7}) (Figure 1C). The results strongly suggested that this

strategy effectively repaired the stop codon and led to the substitution mutation of the target gene.

Cas9 from *Streptococcus pyogenes* recognizes NGG as a PAM for cleavage activity, thus limiting the target site recognition of the system. To eliminate this constraint, we changed the Cas9 nickase to xCas9 nickase, a variant of Cas9 that carries point mutations at the PAM-identification domain and recognizes NGN as a PAM site.[22] We further verified the editing efficiency for the system constructed from two types of XCas9 nickase, XCas9 (H840A) and XCas9 (D10A), which cleave different DNA strands at the same recognition site.[34] We found that the editing efficiency was 3.70×10^{-5} for nXCas9(H840A)-PolA3M and 4.32×10^{-5} for nXCas9(D10A)-PolA3M, respectively 49.45-fold and 57.82-fold higher than that of the off-target control (Figure 1D). This result demonstrated that the system, which we called Cas9 and Mutagenic Polymerase for Evolving Resistance (CAMPER), constructed by XCas9 nickase, effectively mutated the target gene.

To assess the impact of ATc concentration on nXCas9-PolA3M^{TB} expression, *M. smegmatis* cells expressing nXCas9-PolA3M^{TB} and off-target sgRNA were treated with a range of ATc concentrations, and the expression of nXCas9 in *M. smegmatis* was examined by quantitative real-time PCR (qRT-PCR). Under induction with 200 ng/mL ATc, the nXCas9-PolA3M^{TB} achieved maximal induction, to an expression level 43.45-fold higher than that of the control without ATc induction (Figure 1E). We also found that the maximal expression of nXCas9-PolA3M^{TB} persisted in the cells for approximately 24 h but decreased dramatically after 48 h (Figure 1F). Importantly, maintaining maximal induction of nXCas9-PolA3M^{TB} resulted in slight growth defects (Figures 1G–1I), but no significant change on bacterial morphology was observed (Figure 1J), thus suggesting slight proteotoxicity for expression of nXCas9-PolA3M^{TB} fusion protein in mycobacteria.

Subsequently, the targeted *bleo*^{R} gene was analyzed via high-throughput targeted amplicon sequencing (Figure 2A). We found an increase in mutagenesis within a 50-nt window at the 5′ side of the nick and interestingly, a relatively low elevation in mutation frequency within a 10- to 15-bp window ~20 bp from the nick site in the opposite direction. No such increase was observed in cells expressing off-target sgRNA and nXCas9-PolA3M^{TB} (Figure 2B). This result suggested that the fusion protein nXCas9-PolA3M^{TB} advanced along the DNA strand in both 5′-3′ and 3′-5′ directions from the nick. To confirm this feature, we used the system to repair two early stop codons located at different sides of the nick in *bleo*^{R}. For the plasmid with a stop codon in the 5′ window of the nick, the mutation frequency for target editing (3.6×10^{-4}) increased by 100-fold over that of the off-target control (3.67×10^{-6}), whereas the mutation frequency for the plasmid harboring a stop codon at the opposite side of the nick was 9.8×10^{-5} (Figure 2C). The genotypes for repaired clones exhibiting bleomycin resistance were confirmed by Sanger sequencing (Table S3). The results indicated that the CRISPR-guided system edited target fragments at both sides of the nick. Importantly, this system also displayed a double-direction editing feature when heterogeneous expression of enCas9-PolA3M^{TB} occurred in *E. coli*, and its editing efficiency appeared to be higher than that of EvolvR[23] (Figure 2D). In addition, although the target site is a GC-rich region, we observed that two other nucleotides were mutated (Figure 2E). It appeared that CAMPER was more likely to mutate G/C into A/T, rendering it more applicable in G+C-rich bacteria such as *Mycobacterium*.

Off-target effects of nXCas9-PolA3M^{TB} are of great concern. A previous study has shown that off-target effects of Cas9 can occur at sites with five to six mismatches out of 20 bp sequence relative to the protospacer sequence of sgRNA.[35] To evaluate the off-target ratio, we amplified potential target sites with four mismatches relative to the sgRNA targeting site of *bleo*^{R} from ten *bleo*^{R} isolates expressing nXCas9 and *bleo*^{R}-specific sgRNA. Using Sanger sequencing, we observed no difference between the wild-type sequence and these sequences (Table S4), in agreement with previous findings.[35] It seems likely that CAMPER did not reveal significant off-target effects.

Testing CAMPER in *M. smegmatis* and *M. tuberculosis*

Next, we tested CAMPER for its mutagenesis editing ability in *M. smegmatis* through several types of mutant frequency calculation assay. We first evaluated the mutagenesis editing on a plasmid gene. A plasmid for expression of nXCas9-PolA3M^{TB} and another plasmid for expression of *bleo*^{R} disabled by an early stop codon as well as an sgRNA targeting the early stop codon were transformed into *M. smegmatis*. Different volumes of electroporation cultures were subcultured into 1 mL of medium and cultivated until saturation. For the groups with a 100-μL subculture volume, the value of Bleo^{R} CFUs/viable CFUs was 2.83×10^{-5}, which was 118-fold higher than that of the off-target control (2.41×10^{-7}). No significant difference was observed in the mutation frequency among groups with different subculture volumes (Figure 3A). These results indicated that CAMPER could efficiently edit plasmid genes. To test its editing efficiency on chromosomal genes, we decided to target the *rpoB* gene, which encodes the RNA polymerase β subunit whose mutations confer resistance to rifampicin.[36] Via an sgRNA targeting *rpoB*, the target-editing strain displayed a 91-fold higher mutation rate (1.31×10^{-4}) than that of the off-target control strain (1.44×10^{-6}) with a 100-μL subculture volume, and no significant differences were observed in the mutation rates among groups with different subculture volumes (Figure 3B).

Inspired by the success of CAMPER application in *M. smegmatis*, we next explored its programmability for genome editing in *M. tuberculosis*. To simplify the procedure in the slowly growing species, we generated an optimized system without TetR. In agreement with the observations in *M. smegmatis*, constitutive expression of nXCas9-PolA3M with an sgRNA targeting K88 of *rpsL* resulted in a 68.87-fold higher streptomycin-resistant population fraction (1.38×10^{-5}) than that of the off-target control (2.0×10^{-7}) (Figures 3C and 3D). Also, with the expression of sgRNA targeting I66 of *atpE*, which encodes the c subunit of the mycobacterial F-ATP synthase whose

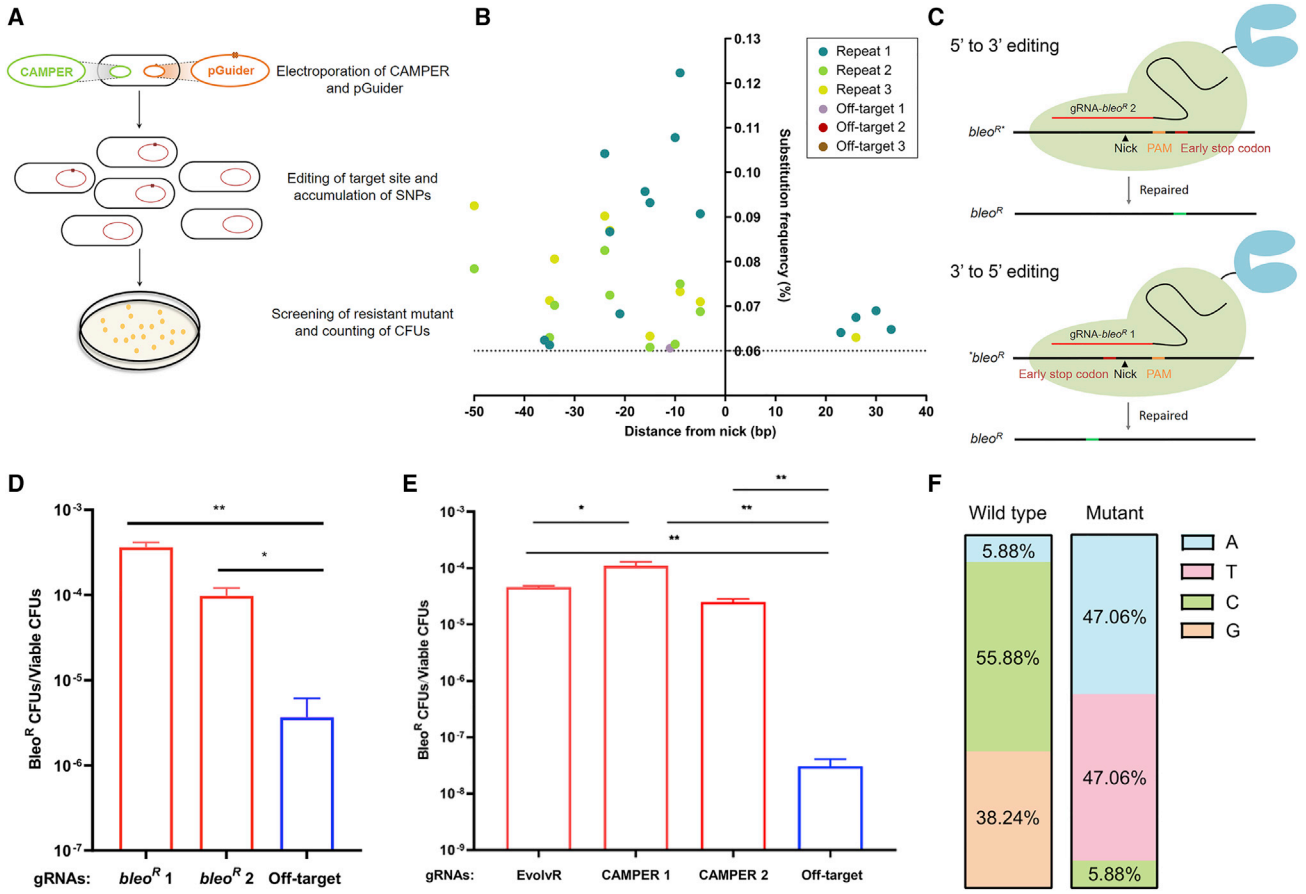

Figure 2. Construction of a CRISPR-guided DNA polymerase system in mycobacteria

(A) Schematic diagram of the mutant frequency calculation assay workflow for calculation of the ratio of the resistant mutant versus the viable population. (B) High-throughput sequencing showing that CAMPER with an on-target gRNA resulted in random substitution mutation within ~40-bp windows at both the 3′ side and 5′ side of the nick. The zero point represents the site nicked by CAMPER. (C and D) Two plasmids harboring *bleo*^{R} with an early stop codon were generated to verify the bidirectional editing feature of CAMPER. One stop codon, W23*, was introduced at the 3′ side of the nick, requiring 3′-to-5′ nick translation for repair (C). Another E77* was introduced at the 5′ end of the nick, requiring 5′ to 3′ nick translation for repair (D). Expression of gRNA-*bleo*^{R} 1 and 2, compared with off-target gRNA, both increased the bleomycin-resistant CFUs. (E) The editing efficiency of CAMPER in *E. coli* was assessed. We constructed a plasmid harboring bleoR with an early stop codon, B68*. One gRNA targeting at the 5′ side of the stop codon was designed (CAMPER 1), requiring 5′-to-3′ nick translation for repair. Another gRNA targeting at the 3′ side of the stop codon was also designed (CAMPER 2), requiring 3′-to-5′ nick translation for repair. Expression of gRNA CAMPER 1 and 2, compared with off-target gRNA, both increased the bleomycin-resistant CFUs. In addition, an EvolvR system was applied to repair the same stop codon from 5′ side to 3′ side, and the editing efficiency was similar to that of the CAMPER in *E. coli*. (F) Distribution of the nucleotides for *bleo*^{R} with an early stop codon (left) and mutated *bleo*^{R} repaired by CAMPER (right). All substitutions were derived from three independent biological replicates with the target strand as reference. All of the above experiments were performed at least three times with similar results. Error bars indicate SEM for three biological replicates. A two-tailed unpaired t test was performed to determine the statistical significance of the data. ns, no significant difference; *p < 0.1; **p < 0.01.

mutations confer resistance to bedaquiline, the percentage of bedaquiline-resistant population (4.83×10^{-5}, Figure 3E) was 22.31-fold higher than that of the off-target control (2.17×10^{-6}). More importantly, consistent with the observation in *M. smegmatis* and *M. tuberculosis* H37Ra, the system also exhibited a capacity for efficient genome editing in *M. tuberculosis* H37Rv (Figures 3F and 3G).

All of the above results demonstrated that the CRISPR-guided DNA polymerase system effectively edited a targeted chromosomal gene in both *M. smegmatis* and *M. tuberculosis*.

Rapid discovery of novel drug-resistant mutations via CAMPER

Elucidating the interaction between drug-resistant mutations and antimicrobial drugs may facilitate the design of novel antibiotics before their clinical applications. We wondered whether our strategy might facilitate the screening of drug-resistant mutations. To this end, we chose AtpE, the target protein of the latest anti-TB drug bedaquiline, as a proof of principle. Bedaquiline inhibits bacterial ATP synthase by targeting the c subunit of the mycobacterial F-ATP synthase AtpE[37–39] (Figure 4A), which is used to treat MDR-TB along with other anti-TB drugs and represents the first novel class of anti-TB agents in more than 40 years.[40]

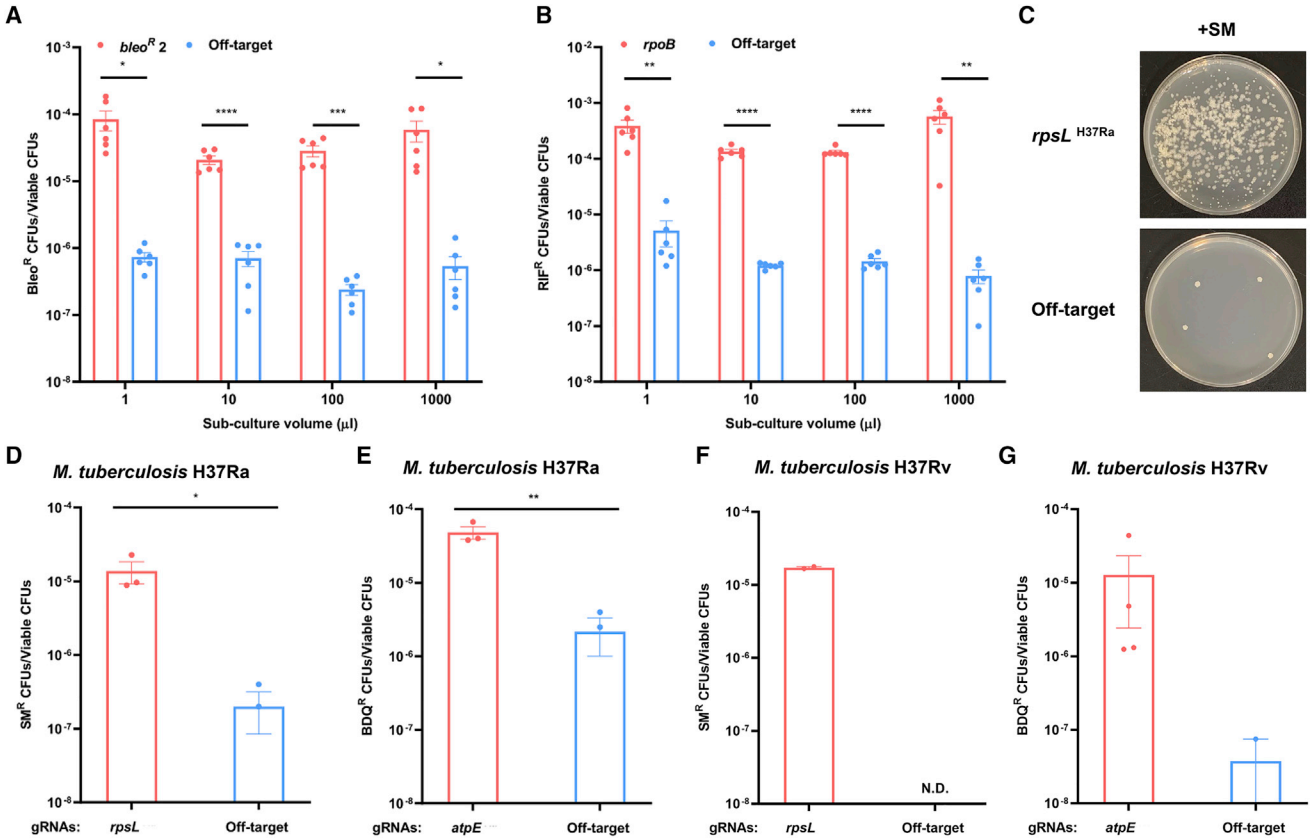

Figure 3. CAMPER enabled mutagenesis of target genes in mycobacteria

(A) A guide targeting *bleo*^{R} with an early stop codon from plasmid was expressed in *M. smegmatis*. We subcultured 1, 10, 100, or 1,000 μL of electroporation culture in 7H9 medium for mutant frequency calculation assays. The percentage of bleomycin-resistant mutants per viable CFU was calculated. (B) A guide targeting chromosomal *rpoB* gene was expressed in *M. smegmatis*. We subcultured 1, 10, 100, or 1,000 μL of electroporation culture in 7H9 medium for mutant frequency calculation assays. The percentage of rifampicin-resistant mutants per viable CFU was calculated. For the assessment of CAMPER gene-editing efficiency in *M. tuberculosis*, four gRNAs were targeted to chromosomal *rpsL* (C and D) and *atpE* (E) from H37Ra, and *rpsL* (F) and *atpE* (G) from H37Rv. All of the above experiments were performed at least three times with similar results, and each dot in the bar graph represents one repeat. N.D, not detected. Error bars indicate SEM for all the repeats. A two-tailed unpaired t test was performed to determine the statistical significance of the data. ns, no significant difference; *p < 0.1; **p < 0.01; ***p < 0.001; ****p < 0.0001.

To detect the drug-resistant mutations, we designed six sgRNAs to cover the 261-bp *atpE* coding sequence (Figure 4B). After culture saturation, bacteria cells were challenged by growth on agar plates containing various concentrations of bedaquiline. We observed that the strains harboring the off-target control, sgRNA-13R, sgRNA-39R, or sgRNA-153R, did not grow under bedaquiline selection, even at the lowest concentration (Figure 4B). For the strains with sgRNA-215F targeting the E65-N71 region of AtpE (c subunit), which has been demonstrated to interact with bedaquiline,[41] the resistance level was highest (Figure 4C). A recent study has demonstrated that the occurrence of a mutation at I66 of AtpE is associated with a 4- to 64-fold bedaquiline minimum inhibitory concentration (MIC) increase in clinical *M. tuberculosis* isolates,[42,43] corresponding to I70 in *M. smegmatis*. To further confirm the resistance phenotype response to genotype, we verified mutation type I70M from ten clones isolated from bedaquiline-resistant strains harboring sgRNA-215F. As expected, the mutation of I70M conferred bedaquiline resistance (Figure S3).

We further used the capacity of CAMPER to diversify the rifampicin resistance determining region (RRDR) of *rpoB* gene to verify the mutation that confers rifampicin resistance in *M. smegmatis* (Figure 4D). In total, 31 sgRNAs were designed to cover the coding region, and the saturated cultures were plated on agar plates containing rifampicin (Figure 4E). The percentage of rifampicin-resistant mutants per viable CFU was calculated. The strains carrying sgRNA-3, sgRNA-12, and sgRNA-16 displayed higher mutation frequencies than that of the remainder, while the expression of sgRNA-11 resulted in the highest mutation frequency, which was 97.43-fold higher than that of the off-target. The corresponding regions targeted by the above sgRNAs were RRDR-N, RRDR-II, RRDR-III, and RRDR-I, respectively. These results are consistent with the report from previous research that the resistance toward rifampicin was related to the spontaneous mutations within five RRDRs in RpoB containing the N-terminal cluster and clusters I, II, and III, with the vast majority of mutations being concentrated in cluster I (RRDR-I)[36] (Figure 4F).

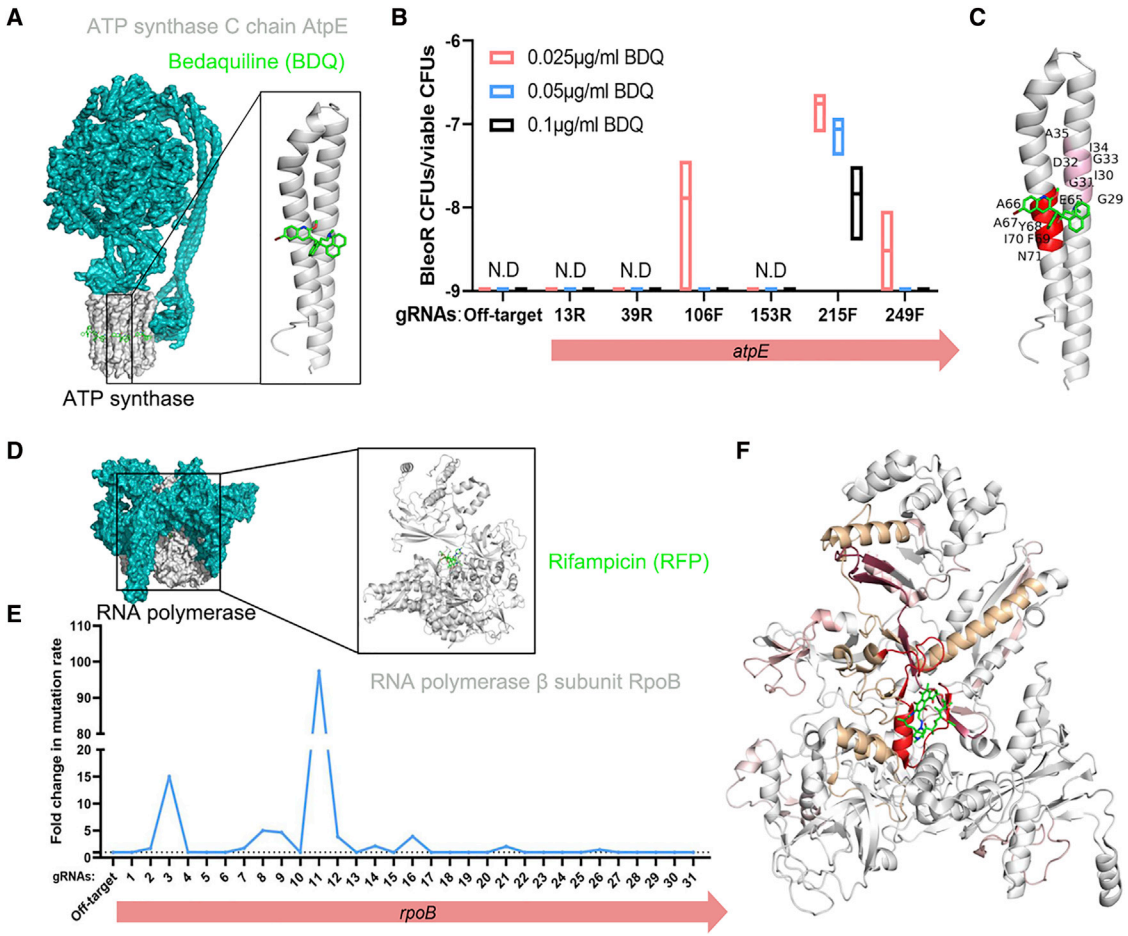

Figure 4. Characterization of drug-resistance-related mutations with CAMPER

A total of six gRNAs were designed to identify the bedaquiline resistance-related mutations in *atpE* in *M. smegmatis*. The interaction between AtpE and bedaquiline is shown in (A). The region targeted by each gRNA and corresponding mutation frequency is shown in (B), where the mutation frequency of each region was determined with CFU assays. With the expression of gRNAs-215F, CAMPER increased the ratio of bedaquiline-resistant cells among the whole viable population. The locations of the mutations are shown in (C), which are labeled with red and the color shaded according to the mutation frequency of each region. Three biological replicates were tested. N.D, not detected. A total of 31 gRNAs were designed to identify the rifampicin resistance-related mutations in *rpoB*. The interaction between RpoB and rifampicin is shown in (D). The region targeted by each gRNA and corresponding mutation frequency are shown in (E), where the mutation frequency of each region was determined with CFU assays. The regions targeted by gRNA-3, gRNA-8, gRNA-9, gRNA-11, gRNA-12, and gRNA-16 displayed significantly higher mutation frequencies than that of the off-target control. The locations of the mutations are shown in (F), which are labeled with red and the color shaded according to the mutation frequency of each region. All of the above experiments were performed at least three times with similar results.

Collectively, this strategy should be useful for rapidly discovering drug-resistant mutations.

Using CAMPER to rapidly analyze the growth phenotype-mutation genotype association

Drug-resistance mutations are often associated with a fitness cost as compared with the antibiotic susceptibility of bacteria in the absence of antibiotic selection.[44,45] Prediction of future drug-resistant TB epidemics depends on the fitness cost of drug-resistance mutations.[46,47] Identification of the fitness of mutation genotype that occurs for essential genes is of great concern in mycobacteria. We assumed that, combined with high-throughput sequencing, our strategy might contribute to rapid analysis of the correlation between growth phenotype and mutation genotype. We chose the RRDR region of *M. smegmatis rpoB* as a proof of concept, because the amino acid sequence of the RRDR is highly conserved between *M. tuberculosis* and *M. smegmatis*. It is well known that rifampicin resistance in *M. tuberculosis* conferring mutation in the RRDR imposes a fitness cost when grown in the absence of rifampicin.[48]

Using CAMPER, we developed a framework to rapidly evolve *rpoB* and determine the fitness of its variants (Figure 5A). The sgRNAs with high editing efficiency were chosen for library construction (Figure S4). Subsequently, we screened out the rifampicin-resistant

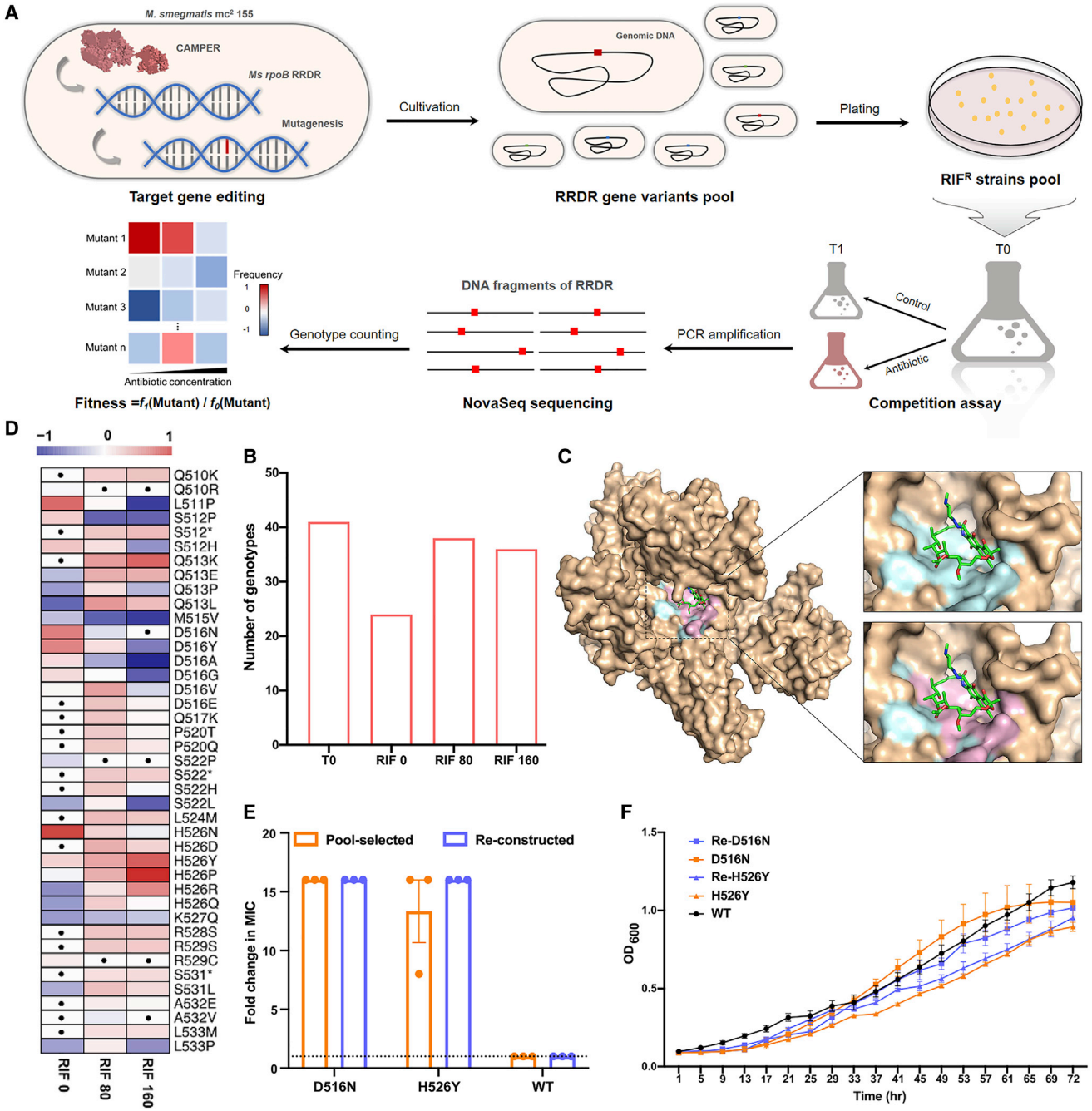

Figure 5. Using CAMPER to rapidly evolve *rpoB* RRDR of *M. smegmatis*

(A) Schematic illustration of the competitive assay workflow for determining the fitness landscape of RRDR. pCAMPER and pGuider were both transformed into *M. smegmatis*. With the expression of gRNAs for target editing, CAMPER performed random substitution mutation among the 81-bp loci of RRDR. During cultivation, the system mutated the target region continuously, owing to induction with ATc. Mutagenesis accumulated among the entire viable population. After the saturation stage, all cultures were plated on agar containing rifampicin to screen out rifampicin-resistant (RIF^R) mutants, and were considered as the T0 library. For competitive assays, equal amounts of RIF^R mutants were subcultured in 7H9 medium containing rifampicin, and cultures without rifampicin were used as a control. Samples were collected at T1, and the 81-bp DNA fragment of RRDR was amplified via PCR. Genotypes and related frequencies were further verified via Illumina HiSeq sequencing. The fitness of each genotype was evaluated as the increase in frequency under antibiotic selection relative to that of wild-type RRDR. (B) The number of genotypes changed under treatment with different concentrations of rifampicin. (C) Surface representation of the *M. smegmatis* RNA polymerase β subunit and steric hindrance of rifampicin binding to the RRDR. The RRDR is in blue, and the region with mutations detected at T0 is in pink. (D) Fitness landscape of RRDR under treatment with rifampicin at 0, 80, or 160 μg/mL at T1. Each tile

(legend continued on next page)

mutants by plating the pool on Luria-Bertani (LB) agar plates containing rifampicin. After incubation, all colonies were collected from the resistant plate for *in vitro* competitive experiments, and this time point was considered time zero (T0). Competitive experiments were performed under three concentrations of rifampicin (0, 80, and 160 μg/mL). The indicated samples were collected in log phase (Figure S5A). To accurately estimate the fitness of the variants, we included 41 non-synonymous mutation genotypes with frequencies $\geq$0.06% at T0 in further analysis, discarding multiple mutations, indels, or deletions (Figure 4B and Table S5). The genotype frequencies were highly correlated between biological replicates under treatment with rifampicin at 0, 80, and 160 μg/mL, and the mean *R* of all replicate pairs was 0.9999 (Figure S5B). In addition, we also verified ten mutation genotypes directly isolated from the T0 rifampicin-resistant strains (RIF^R) pool through Sanger sequencing (Table S6). Among these mutations, nine were identical to the mutation genotypes identified from the T0 library, one of which carried three point mutations, and the serine at position 522 was substituted with valine, which is an RIF^R-related mutation that has not been reported in clinical *M. tuberculosis* isolates to date.

The number of mutation genotypes decreased to 24 under the condition without antibiotics, and this number increased to 38 under treatment with 80 μg/mL rifampicin and to 36 under selection with 160 μg/mL rifampicin (Figure 5B). A total of 74.1% (20/27) of the amino acid mutations in RRDR were represented in the T0 library (Figure 5C). Next, we verified the fitness change of the 41 mutants under competitive assays (Figure 5D). Notably, all single nucleotide mutation types for the positions D516 and H526 were represented in our library, suggesting that CAMPER conferred an ability to substitute all four nucleotides. The fitness of mutations in D516 decreased in a dose-dependent manner. In contrast, mutations in D526 displayed an increase in fitness with increasing rifampicin concentration. In addition, three stop codon mutations, S512*, S522*, and S531*, were represented in our library, and these mutants exhibited reduced competitive ability patterns in the absence of rifampicin but reversed patterns under rifampicin selection. To our knowledge, these are the first reported findings identifying such mutation genotypes with extremely low occurrence in clinical settings involving *M. tuberculosis* isolates by means of laboratory methods in mycobacteria. Moreover, to confirm the resistance-related genotype and growth phenotype of the *rpoB* mutant, we selected two mutants, D516N and H526Y, with significant fitness changes during competitive assays, from the T0 library. These two mutants (Re-D516N and Re-H526Y) were also generated in wild-type *M. smegmatis* through HR (Figure S6). Both the pool-selected and reconstructed mutants exhibited similarly increased resistance to rifampicin (Figure 5E) and similar growth patterns (Figure 5F).

On the basis of these data, we conclude that CAMPER combined with high-throughput sequencing enables rapid determination of the relationship between growth phenotype and mutation genotype in mycobacteria.

DISCUSSION

In this study, we developed an approach for mutagenesis of targeted genomic loci in mycobacteria by using a CRISPR-guided DNA polymerase system. This system may significantly enable TB research, particularly on the evolution of antimicrobial resistance and anti-TB drug development. First, this approach enables targeted diversification of all nucleotides at user-defined loci in *M. tuberculosis* within 2 months. Moreover, whereas conventional approaches rely on laborious genomic engineering procedures, diversification of the genomic region edited by our system requires only co-expression of nXCas9-PolA3M^{TB} and a 20-nt sgRNA, thus making this system efficient in slowly growing mycobacteria. Second, CAMPER can be useful for understanding protein-drug interactions. As demonstrated by our data, this system can be used to rapidly identify the drug-resistant mutations of gene responses to its protein function. Finally, as demonstrated in this study, predictions of drug-resistance mutations in mycobacteria according to our framework may have different potentials for clinical applications. For example, this framework could be used to perform a rapid preliminary evaluation of whether novel mutations are likely to confer the ability to resist therapy, thus contributing to genotype-phenotype predictions in diagnostics. In addition, this framework could be used to map novel variants, which may emerge after deployment in clinical practice, for assessment of novel agents.

Two caveats in our study should be discussed. First, we observed that high expression of nxCas9-PolA3M^{TB} slightly retarded bacterial growth, suggesting that some degree of proteotoxicity may result from nxCas9-PolA3M^{TB} expression. In fact, first-generation Cas9spy proteotoxicity has been reported in mycobacteria. Although second-generation Cas9sth1 exhibits low toxicity and robust utility in gene knockdown and knockout in mycobacteria,[49,50] $Cas9_{sth1}$ is unable to target certain specific sequences because of the limitation of its unique PAM sequences. Therefore, exploring novel Cas9 with low toxicity and high flexibility in mycobacteria would enable further improvements in the performance of this tool. Second, our data demonstrated that the editing window of nxCas9-PolA3M^{TB} was ~80 bp. Also, the mutation frequency relies on the flexibility of the DNA polymerase domain of PolA. Rationally designed mutation of the DNA polymerase domain to increase the mutation frequency and window length will be highly important in future research.

Interestingly, we observed that CAMPER has an extended mutagenesis window length. This phenomenon in mycobacteria was not

represents a variant with one single-base substitution (x axis) at one specific position (y axis), whose relative growth is indicated by the color of the tile scale according to the corresponding color scale bar on top. In addition, mutants with a frequency <0.06% are marked with a black dot. (E) Drug susceptibility comparison of pool-selected and reconstructed strains for rifampicin-resistant mutants D516N and H526Y (MIC fold change RIF^R/parent strain). Each dot in the bar graph represents one repeat. (F) Growth curve pattern comparison among pool-selected and reconstructed strains for the rifampicin-resistant mutants D516N and H526Y. The cell density is presented as the value of OD_{600} and was measured every 4 h. All of the above experiments were performed at least three times with similar results. Error bars indicate SEM for all the repeats.

previously observed in *E. coli*, in which mutagenesis induced by *E. coli* PolA only occurs at 3′ of the nick sites. CAMPER also displayed a double-direction editing feature with heterogeneous expression of enCas9-PolA3M^{TB} in *E. coli*, indicating that a difference in structure between PolA^{TB} and *E. coli* PolA, rather than in genetic backgrounds (*E. coli* versus mycobacteria) may result in this editing feature. Although recent work[51] has reported a similar protein structure between *E. coli* PolA and *M. smegmatis* PolA, the 3′-5′ exonuclease subdomain of *M. smegmatis* PolA lacks several carboxylate-containing amino acids that interact with essential metal ions at the 3′-5′ exonuclease active site of the *E. coli* PolA Klenow large fragment.[51] This distinctive feature results in a vestigial 3′-5′ exonuclease and a lack of proofreading activity in *M. smegmatis* PolA. In addition, the vestigial 3′-5′ exonuclease subdomain of *M. smegmatis* PolA contains a unique single-manganese ion binding site, which cannot engage with other metal ions and is not conserved in *E. coli* PolA.[62] These findings may to some extent contribute to the action of *M. tuberculosis* PolA observed herein, because the interaction with metal ions is essential for the activity of 3′-5′ exonuclease and its binding to single-stranded DNA. However, the mechanism underlying the bidirectional nick translation activity of PolA3M^{TB} remains to be elucidated.

MATERIALS AND METHODS

Strains and media

M. tuberculosis H37Ra and H37Rv were cultured at 37°C in Difco Middlebrook 7H9 broth or 7H10 plates supplemented with 0.2% glycerol (7H9) or 0.5% glycerol (7H10), 0.05% Tween 80, and OADC (oleic acid-albumin-dextrose-catalase). *M. smegmatis* MC^2 155 was grown at 37°C in LB agar plates, or 7H9 broth supplemented with 0.05% Tween 80 and 10% oleic acid ADC. When indicated, antibiotics and inducers were used at the following concentrations: kanamycin (25 μg/mL), hygromycin (100 μg/mL), bleomycin (25 μg/mL), streptomycin (25 μg/mL), rifampicin (80 μg/mL), bedaquiline (0.0125 μg/mL), and ATc (200 ng/mL). The *E. coli* strain DH5α was used for plasmid cloning and was cultivated at 37°C in LB broth or agar plates. When indicated, antibiotics were used at the following concentrations: kanamycin (50 μg/mL), hygromycin (150 μg/mL), and streptomycin (50 μg/mL).

Plasmid construction

All plasmids (Table S1) were constructed with a one-step cloning strategy. The integrating vector pCAMPER consisted of a kanamycin resistance cassette, integrase from phage L5, a high-copy-number origin of replication, and genes encoding nXCas9-PolA3M^{TB} under control of the uv15tetO promoter. To allow for simple cloning of sgRNA, we constructed a replicative vector, pGuider, derived from pRH2521.[35] In brief, an sfGFP cassette under the control of the J23115 promoter was cloned into the *Bbs*I-digested pRH2521. The pGuider2 was identical to pGuider except that it contained a *bleo^R* gene carrying an E77* mutation. The pGuider3 was identical to pGuider except that it contained a *bleo^R* gene carrying a W23* mutation. The full plasmid sequences are provided in Table S2. Details on plasmid construction are available upon request.

mRNA quantification

To identify the optimal ATc concentration for nXCas9-PolA3M^{TB} expression induction, we grew *M. smegmatis* harboring pCAMPER and pGuider to log phase and performed induction with ATc at 0, 10, 50, or 200 μg/mL, respectively. After cultivation at 37°C for 48 h, we collected 10 mL of culture for each sample, and three replicates were used for each condition. To further verify the steadiness of the fusion protein expression in *M. smegmatis*, *M. smegmatis* harboring pCAMPER and pGuider for an off-target control was grown to log phase. After induction with ATc at 200 ng/mL, 10 mL of culture was collected at 0, 8, 24, 48, and 96 h, respectively. The collected cultures were harvested by centrifugation. After removal of the supernatant, the cell pellet was resuspended in RNA-easy Isolation Reagent (Vazyme) and disrupted by bead beating with a TissueLyser-II (Qiagen). Total RNA was precipitated by the addition of isopropanol and collected by centrifugation, the supernatant was discarded, and the mRNA pellet was washed with 75% ethanol. After drying of the pellet, mRNA was dissolved in RNase-free H_2O. The contaminating genomic DNA was digested with gDNA wiper Mix (Vazyme). The cDNA was prepared from purified mRNA with HiScriptII qRT SuperMix II (Vazyme) through reverse transcription. The cDNA levels of target genes were then quantified by qRT-PCR on a CFX96 cycler (Bio-Rad) with AceQ Universal SYBR qPCR Master Mix (Vazyme) as a dye for fluorescence signal detection. All qPCR primers were determined to be >95% efficient, and the cDNA masses tested were experimentally validated to be within the linear dynamic range of the assay. Signals were normalized to those of the housekeeping *sigA* transcript and quantified with the ΔΔCt method. Error bars were 95% confidence intervals of the three technical replicates.

Gene editing in *M. smegmatis* and *M. tuberculosis*

For *M. smegmatis*, a total of 100 ng of pGuider was transformed into *M. smegmatis* harboring pCAMPER electroporation-competent cells. After 2 h of recovery at 37°C, unless otherwise stated, 100 μL of culture was inoculated into 1 mL of 7H9 medium containing kanamycin, hygromycin, and ATc. The cultures were cultivated at 37°C until reaching the saturation stage, and 100 μL of culture was plated on an LB agar plate containing corresponding antibiotic for target mutant screening. For counting of viable CFUs, 100 μL of 1:10^6 diluted culture was plated on an LB agar plate containing kanamycin. After incubation at 37°C for 72 h, resistant CFUs and viable CFUs were counted, and the mutation frequencies were verified according to the values of resistant CFUs/viable CFUs. Off-target gRNA was transformed as a negative control.

For *M. tuberculosis*, a total of 100 ng of pGuider for expression of sgRNA targeting the 3′-terminal regions of *rpoB* RRDR, K88 of *rpsL*, or I66 of *atpE* was transformed into *M. tuberculosis* harboring pCAMPER electroporation-competent cells. For both *M. tuberculosis* H37Ra and H37Rv, 100-mL logarithmic phase cultures of were collected by centrifugation (5,000 × *g*, 10 min) and washed three times with gradually decreasing volumes (40, 20, and 10 mL) of 10% glycerol at room temperature, and finally resuspended in 1 mL of 10% glycerol. For high yield of transformants, 100 ng of plasmid

was mixed thoroughly with 200 μL of fresh electrocompetent bacilli, which were then transferred to a chilled 0.2-cm electrode gap GenePulser eletroporation cuvette (Bio-Rad) and subjected to electrotransformation using a GenePulser Cxell electroporator (Bio-Rad) at the following settings: 2.5 kV, 1,000 Ω, and 25 μF. After 24 h of recovery at 37°C, 100 μL of culture was then inoculated into 1 mL of 7H9 medium containing kanamycin, hygromycin, and ATc. The cultures were cultivated at 37°C until saturation stage, and 100 μL of culture was plated on a 7H10 agar plate containing rifampicin, streptomycin, or bedaquiline for target mutant screening. For viable CFU counting, 100 μL of $1:10^6$ diluted culture was plated on a 7H10 agar plate containing kanamycin. After incubation at 37°C for 28 days, resistant CFUs and viable CFUs were counted, and the mutation frequencies were verified by the values of resistant CFUs/viable CFUs.

Sample preparation for high-throughput sequencing

For NovaSeq next-generation sequencing, genomic DNA was extracted from samples of interest with a HiPure Bacterial DNA kit (Magen). The related oligonucleotides were used for the amplification of the target region in a 20-cycle PCR reaction with high-fidelity Phanta DNA polymerase (Vazyme), with 200 ng of genomic DNA as the template. The PCR products were then purified with FastPure DNA Mini Columns (Vazyme). For the second PCR reactions with 100 ng of the first-round PCR product as the template, the gene of interest was amplified with oligonucleotides containing Illumina sequencing adapters, an index for sample separation and primers for sequencing with ten thermocycles. The well-prepared samples were submitted to the Novogene Sequencing Center & Clinical Lab (Tianjin, China) for quality control and sequencing. The quality control included fragment size analysis and concentration measurement. The pooled sample was run on an Illumina NovaSeq instrument by means of a 150-bp paired-end read NovaSeq Reagent Kit. The resulting raw data were then converted into fastq format for splicing, filtration, and analysis.

High-throughput sequencing data analysis

The raw reads were filtered with fastp v.0.2 (https://doi.org/10.1093/bioinformatics/bty560) through the following procedures: (1) primers and adapter sequences were removed (default); (2) the first and last five nucleotides were trimmed (-f 5 -t 5); (3) reads with N bases (-n 0) or low-quality bases (-u 100) were removed; (4) paired-end reads whose R1 did not perfectly match part of the reverse complement R2 were removed (–overlap_diff_limit 0 –overlap_diff_percent_limit 0); and (5) each pair of reads were merged into a single read if they overlapped (-m), or were removed otherwise. Merged reads were mapped to the wild-type *rpoB-RRDR* gene (GenBank: ABK70312) or $bleo^R$ gene (GenBank: CAA02067) to generate the alignment files via bwa v0.7 (https://doi.org/10.1093/bioinformatics/btp324) and samtools v0.1.18 (https://doi.org/10.1093/bioinformatics/btp352). Variant calling was performed with VarScan v.2.3.9 with parameters: –min-coverage 1 –min-reads2 1 –p-value 0.99 –variants 1 –min-var-freq 0 (https://doi.org/10.1093/bioinformatics/btp373). Because the highest frequency variant of the empty vector control was 0.06%, the variants in the experimental group were discarded from subsequent analysis if the frequency was below 0.06% (–min-var-freq 0.0006).

Construction of an RRDR gene variant pool

To construct a mutation library covering all loci of RRDR, we designed six sgRNAs for mutagenesis. Three targeted the 5′ terminus (101F, 113F, and 91R) and the remainder targeted the 3′ terminus (40R, 67F, and 74F). The pGuider plasmids for expression of either the six sgRNAs or off-target gRNA were transformed into *M. smegmatis* harboring pCAMPER electroporation-competent cells. For each condition, 100 μL of culture was inoculated into 1 mL of 7H9 medium containing kanamycin, hygromycin, and ATc. The cultures were cultivated at 37°C until reaching saturation stage. To verify the editing efficiency of the systems guided by different sgRNAs, we calculated the mutation frequencies with CFU assays: 100 μL of culture was plated on an LB agar plate containing rifampicin for target mutant screening. For viable CFU counting, 100 μL of $1:10^6$ diluted culture was plated on an LB agar plate containing kanamycin. After incubation at 37°C for 72 h, resistant CFUs and viable CFUs were counted, and the mutation frequencies were verified according to the values of resistant CFUs/viable CFUs. Three replicates were used for each condition. All remaining cultures were prepared for competition assays.

Competition assays

The well-prepared RRDR mutant library was first plated on LB agar containing kanamycin and rifampicin to screen out rifampicin resistance-related isolates, then 100 μL of culture was plated on each plate. After incubation at 37°C for 72 h, all colonies were collected from the LB agar plate by washing with saline and were considered as the rifampicin-resistant strain T0 library. A 5-mL sample was then collected for genotype analysis. Three replicates were used. Aliquots (100 μL) of the pooled culture were subcultured in 100 mL of 7H9 medium containing rifampicin at 0, 80, or 160 μg/mL. Three biological repeats were examined for each condition. A total of 50 mL of sample culture was collected at the indicated time points, where OD_{600} was 0.6. At each time point, the cell concentration was measured with a Nano-photometer (IMPLEN) and represented by the value of OD_{600}.

Estimation of the growth phenotype-mutation genotype association

To determine the frequency of each barcode within a sample, we performed paired-end 150-bp sequencing on each sample on the Illumina NovaSeq platform. RRDR sequences of 81 nt were extracted from the sequencing reads, and the sequences with nonidentical sequences from matching forward and reverse sequencing reads were excluded from further analysis. To ensure accurate estimation of the fitness, we analyzed 41 genotypes with frequency $\geq$0.06% from the technical repeats at T0. The fitness at each time point of each competitive growth assay was calculated for individual mutations as f_1(Mutant)/f_0(Mutant), where f_1(Mutant) is the frequency of sequencing reads of a certain genotype at T1 and f_0(Mutant) is the frequency of sequencing reads of the same genotype at T0.

Strain construction

Deletion of *rpoB* was generated by transformation with a second copy of the gene and integration at the L5 phage integration sites, and subsequent mycobacterial recombineering as previously described[52] to delete the endogenous copy.

To confirm the phenotype of the *rpoB* mutants observed in competitive experiments, we constructed two variants of *rpoB* by one-step cloning. In brief, two adjacent mutated target fragments with 20–15 bp homologous sequences were amplified by PCR. The PCR products were verified by electrophoresis and purified with a gel extraction kit (Omega). The pGMCZ-rpoB-SmR plasmid was digested with FastDigest NsiI (Thermo Fisher) and FastDigest XhoI (Thermo Fisher) at 37°C for 1 h. Ligation was performed at a vector/insert ratio of 1:3 with a Seamless Cloning Kit (Beyotime, China). The product was transformed into DH5-α-competent cells and selected on streptomycin (100 μg/mL). Recombinant plasmids were purified, and the corresponding *rpoB* variant was sequenced to confirm the mutation. The plasmid containing the *rpoB* mutation was transformed into *M. smegmatis* Δ*rpoB attB*::*rpoB bleo*^{R} competent cells. L5 swaps were performed as previously described.[53]

Three-dimensional analysis of RpoB variant proteins

Three-dimensional structural data for *M. smegmatis* RpoB (PDB: 6CCV), ATP synthase (PDB: 7JG8), and *E. coli* PolA (PDB: 1KLN) with high resolution were downloaded from the PDB database. The architecture of full-length *M. tuberculosis* PolA was modeled with Swiss-Model. The PolA of *M. smegmatis* (PDB: 6VDE) served as a structural template. The protein structure was visualized with PyMOL software.

MIC assays

To determine the rifampicin susceptibility of the RRDR variants, we performed rifampicin MIC assays. In brief, single colonies of the target RRDR mutants were inoculated in 7H9 medium separately and cultivated overnight at 37°C. The cultures were then adjusted to an OD_{600} of 0.5 with saline and diluted at a ratio of 1:10. Next, 20 μL of diluted culture was added into 180 μL of 7H9 medium containing rifampicin at 0, 4, 8, 16, 32, 64, 128, or 256 μg/mL in a 96-well plate. A Nano-photometer (Implen) was used to measure the optical density of each well after 72 h of growth at 37°C. Three biological replicates of each RRDR mutant under each rifampicin concentration were verified, and a wild-type *M. smegmatis* MC^{2} 155 strain was used for comparison.

Measurement of growth curve

Fresh single colonies of the target strains were inoculated in 7H9 medium separately and cultivated overnight at 37°C. The cultures were then adjusted to an OD_{600} of 0.5 with saline, and diluted at a ratio of 1:10. Next, 20 μL of diluted culture was added to 180 μL of 7H9 medium in a 96-well plate. A Nano-photometer (Implen) was used to measure the optical density of each well every 3 h. Three replicates were used for each strain. The growth curve was plotted in Prism 8.

Statistical analysis

Statistical analysis was performed in Prism (version 8.0; GraphPad Software). Statistical significance of paired comparison data was assessed with Student's paired t test. The statistical significance of data with multiple confounding factors was assessed with analysis of variance. A p value of 0.05 or lower was considered statistically significant.

DATA AND CODE AVAILABILITY

The data that support the findings of this study are available from the corresponding authors upon request. High-throughput sequencing data have been deposited as an NCBI BioProject under accession number GenBank: PRJNA716418. Also, the code that supports the findings of this study is available at https://github.com/shencong456/ropB_bleoR.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtn.2022.07.004.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (grant numbers 82061128001, 81830103 to G.-B.T., grant number 82002173 to S.F.), Guangdong Natural Science Foundation (grant number 2017A030306012 to G.-B.T.), National Key Research and Development Program (grant number 2017ZX10302301 to G.-B T.), Project of High-level Health Teams of Zhuhai in 2018 (The Innovation Team for Antimicrobial Resistance and Clinical Infection to G.-B.T.), and the Project 111 (grant number B12003 to G.-B.T.).

AUTHOR CONTRIBUTIONS

S.F., L. Liang, and C.S. contributed equally to this study. G.-B.T. conceived the idea, and designed and supervised the study. L. Liang, S.F., D.L., and L. Liu conducted the experiments and produced the tables and figures. J.L. and W.L. searched the literature. C.S. contributed to data analysis and interpretation. S.F., L. Liang, and G.-B.T. wrote the manuscript, which was reviewed and edited by G.M.C. and Y.D. All authors reviewed, revised, and approved the final report.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Dorothee Heemskerk, M.C., Marais, B., and Farrar., J. (2015). Tuberculosis in adults and children. In Treatment, *Chapter 5* (Springer).
2. Sandgren, A., Strong, M., Muthukrishnan, P., Weiner, B.K., Church, G.M., and Murray, M.B. (2009). Tuberculosis drug resistance mutation database. PLoS Med. *6*, e2. https://doi.org/10.1371/journal.pmed.1000002.
3. Aziz, M.A., Wright, A., Laszlo, A., De Muynck, A., Portaels, F., Van Deun, A., Wells, C., Nunn, P., Blanc, L., Raviglione, M., et al.; WHO/International Union Against Tuberculosis And Lung Disease Global Project on Anti-tuberculosis Drug Resistance Surveillance (2006). Epidemiology of antituberculosis drug resistance (the global Project on anti-tuberculosis drug resistance surveillance): an updated analysis. Lancet *368*, 2142–2154. https://doi.org/10.1016/S0140-6736(06)69863-2.
4. WHO global strategy for containment of antimicrobial resistance. (Geneva: World Health Organization), 2001.

5. Apjok, G., Boross, G., Nyerges, Á., Fekete, G., Lázár, V., Papp, B., Pál, C., and Csörgő, B. (2019). Limited evolutionary conservation of the phenotypic effects of antibiotic resistance mutations. Mol. Biol. Evol. *36*, 1601–1611. https://doi.org/10.1093/molbev/msz109.

6. Global tuberculosis report 2020. (Geneva: World Health Organization), 2020.

7. Walker, T.M., Kohl, T.A., Omar, S.V., Hedge, J., Del Ojo Elias, C., Bradley, P., Iqbal, Z., Feuerriegel, S., Niehaus, K.E., Wilson, D.J., et al.; Modernizing Medical Microbiology MMM Informatics Group (2015). Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. Lancet Infect. Dis. *15*, 1193–1202. https://doi.org/10.1016/S1473-3099(15)00062-6.

8. Steingart, K.R., Schiller, I., Horne, D.J., Pai, M., Boehme, C.C., and Dendukuri, N. (2014). Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst. Rev. CD009593. https://doi.org/10.1002/14651858.CD009593.pub3.

9. Munir, A., Vedithi, S.C., Chaplin, A.K., and Blundell, T.L. (2020). Genomics, computational biology and drug discovery for mycobacterial infections: fighting the emergence of resistance. Front. Genet. *11*, 965. https://doi.org/10.3389/fgene.2020.00965.

10. Evangelopoulos, D., Prosser, G.A., Rodgers, A., Dagg, B.M., Khatri, B., Ho, M.M., Gutierrez, M.G., Cortes, T., and de Carvalho, L.P.S. (2019). Comparative fitness analysis of D-cycloserine resistant mutants reveals both fitness-neutral and high-fitness cost genotypes. Nat. Commun. *10*, 4177. https://doi.org/10.1038/s41467-019-12074-z.

11. Bacon, J., Hatch, K.A., and Allnutt, J. (2010). Application of continuous culture for measuring the effect of environmental stress on mutation frequency in Mycobacterium tuberculosis. Methods Mol. Biol. *642*, 123–140. https://doi.org/10.1007/978-1-60327-279-7_10.

12. Castañeda-García, A., Martín-Blecua, I., Cebrián-Sastre, E., Chiner-Oms, A., Torres-Puente, M., Comas, I., and Blázquez, J. (2020). Specificity and mutagenesis bias of the mycobacterial alternative mismatch repair analyzed by mutation accumulation studies. Sci. Adv. *6*, eaay4453. https://doi.org/10.1126/sciadv.aay4453.

13. Yadon, A.N., Maharaj, K., Adamson, J.H., Lai, Y.P., Sacchettini, J.C., Ioerger, T.R., Rubin, E.J., and Pym, A.S. (2017). A comprehensive characterization of PncA polymorphisms that confer resistance to pyrazinamide. Nat. Commun. *8*, 588. ARTN 588. https://doi.org/10.1038/s41467-017-00721-2.

14. Borgers, K., Vandewalle, K., Festjens, N., and Callewaert, N. (2019). A guide to Mycobacterium mutagenesis. FEBS J. *286*, 3757–3774. https://doi.org/10.1111/febs.15041.

15. Kalpana, G.V., Bloom, B.R., and Jacobs, W.R., Jr. (1991). Insertional mutagenesis and illegitimate recombination in mycobacteria. Proc. Natl. Acad. Sci. USA *88*, 5433–5437. https://doi.org/10.1073/pnas.88.12.5433.

16. Muniyappa, K., Vaze, M.B., Ganesh, N., Sreedhar Reddy, M., Guhan, N., and Venkatesh, R. (2000). Comparative genomics of Mycobacterium tuberculosis and Escherichia coli for recombination (rec) genes. Microbiology (Read.) *146*, 2093–2095. https://doi.org/10.1099/00221287-146-9-2093.

17. Halperin, S.O., Tou, C.J., Wong, E.B., Modavi, C., Schaffer, D.V., and Dueber, J.E. (2018). CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. Nature *560*, 248–252. https://doi.org/10.1038/s41586-018-0384-8.

18. Tou, C.J., Schaffer, D.V., and Dueber, J.E. (2020). Targeted diversification in the S. cerevisiae genome with CRISPR-guided DNA polymerase I. ACS Synth. Biol. *9*, 1911–1916. https://doi.org/10.1021/acssynbio.0c00149.

19. Allen, J.M., Simcha, D.M., Ericson, N.G., Alexander, D.L., Marquette, J.T., Van Biber, B.P., Troll, C.J., Karchin, R., Bielas, J.H., Loeb, L.A., and Camps, M. (2011). Roles of DNA polymerase I in leading and lagging-strand replication defined by a high-resolution mutation footprint of ColE1 plasmid replication. Nucleic Acids Res. *39*, 7020–7033. https://doi.org/10.1093/nar/gkr157.

20. Billington, O.J., McHugh, T.D., and Gillespie, S.H. (1999). Physiological cost of rifampin resistance induced in vitro in Mycobacterium tuberculosis. Antimicrob. Agents Chemother. *43*, 1866–1869. https://doi.org/10.1128/AAC.43.8.1866.

21. Peters, J.S., Ismail, N., Dippenaar, A., Ma, S., Sherman, D.R., Warren, R.M., and Kana, B.D. (2020). Genetic diversity in Mycobacterium tuberculosis clinical isolates and resulting outcomes of tuberculosis infection and disease. Annu. Rev. Genet. *54*, 511–537. https://doi.org/10.1146/annurev-genet-022820-085940.

22. Nishimasu, H., Shi, X., Ishiguro, S., Gao, L., Hirano, S., Okazaki, S., Noda, T., Abudayyeh, O.O., Gootenberg, J.S., Mori, H., et al. (2018). Engineered CRISPR-Cas9 nuclease with expanded targeting space. Science *361*, 1259–1262. https://doi.org/10.1126/science.aas9129.

23. Joyce, C.M., Kelley, W.S., and Grindley, N.D. (1982). Nucleotide sequence of the Escherichia coli polA gene and primary structure of DNA polymerase I. J. Biol. Chem. *257*, 1958–1964.

24. Bebenek, K., Joyce, C.M., Fitzgerald, M.P., and Kunkel, T.A. (1990). The fidelity of DNA synthesis catalyzed by derivatives of Escherichia coli DNA polymerase I. J. Biol. Chem. *265*, 13878–13887.

25. Carroll, S.S., Cowart, M., and Benkovic, S.J. (1991). A mutant of DNA polymerase I (Klenow fragment) with reduced fidelity. Biochemistry *30*, 804–813. https://doi.org/10.1021/bi00217a034.

26. Reha-Krantz, L.J., Stocki, S., Nonay, R.L., Dimayuga, E., Goodrich, L.D., Konigsberg, W.H., and Spicer, E.K. (1991). DNA polymerization in the absence of exonucleolytic proofreading: in vivo and in vitro studies. Proc. Natl. Acad. Sci. USA *88*, 2417–2421. https://doi.org/10.1073/pnas.88.6.2417.

27. Loh, E., Choe, J., and Loeb, L.A. (2007). Highly tolerated amino acid substitutions increase the fidelity of Escherichia coli DNA polymerase I. J. Biol. Chem. *282*, 12201–12209. https://doi.org/10.1074/jbc.M611294200.

28. Shinkai, A., and Loeb, L.A. (2001). In vivo mutagenesis by Escherichia coli DNA polymerase I. Ile(709) in motif A functions in base selection. J. Biol. Chem. *276*, 46759–46764. https://doi.org/10.1074/jbc.M104780200.

29. Troll, C., Alexander, D., Allen, J., Marquette, J., and Camps, M. (2011). Mutagenesis and functional selection protocols for directed evolution of proteins in E. coli. JoVE, e2505. https://doi.org/10.3791/2505.

30. Loh, E., Salk, J.J., and Loeb, L.A. (2010). Optimization of DNA polymerase mutation rates during bacterial evolution. Proc. Natl. Acad. Sci. USA *107*, 1154–1159. https://doi.org/10.1073/pnas.0912451107.

31. Camps, M., Naukkarinen, J., Johnson, B.P., and Loeb, L.A. (2003). Targeted gene evolution in Escherichia coli using a highly error-prone DNA polymerase I. Proc. Natl. Acad. Sci. USA *100*, 9727–9732. https://doi.org/10.1073/pnas.1333928100.

32. Coll, F., Phelan, J., Hill-Cawthorne, G.A., Nair, M.B., Mallard, K., Ali, S., Abdallah, A.M., Alghamdi, S., Alsomali, M., Ahmed, A.O., et al. (2018). Genome-wide analysis of multi- and extensively drug-resistant Mycobacterium tuberculosis. Nat. Genet. *50*, 307–316. https://doi.org/10.1038/s41588-017-0029-0.

33. Pelchovich, G., Zhuravlev, A., and Gophna, U. (2013). Effect of ribosome-targeting antibiotics on streptomycin-resistant Mycobacterium mutants in the rpsL gene. Int. J. Antimicrob. Agents *42*, 129–132. https://doi.org/10.1016/j.ijantimicag.2013.04.001.

34. Nishimasu, H., Ran, F.A., Hsu, P.D., Konermann, S., Shehata, S.I., Dohmae, N., Ishitani, R., Zhang, F., and Nureki, O. (2014). Crystal structure of Cas9 in complex with guide RNA and target DNA. Cell *156*, 935–949. https://doi.org/10.1016/j.cell.2014.02.001.

35. Singh, A.K., Carette, X., Potluri, L.P., Sharp, J.D., Xu, R., Prisic, S., and Husson, R.N. (2016). Investigating essential gene function in Mycobacterium tuberculosis using an efficient CRISPR interference system. Nucleic Acids Res. *44*, e143. https://doi.org/10.1093/nar/gkw625.

36. Molodtsov, V., Scharf, N.T., Stefan, M.A., Garcia, G.A., and Murakami, K.S. (2017). Structural basis for rifamycin resistance of bacterial RNA polymerase by the three most clinically important RpoB mutations found in Mycobacterium tuberculosis. Mol. Microbiol. *103*, 1034–1045. https://doi.org/10.1111/mmi.13606.

37. Andries, K., Verhasselt, P., Guillemont, J., Göhlmann, H.W.H., Neefs, J.M., Winkler, H., Van Gestel, J., Timmerman, P., Zhu, M., Lee, E., et al. (2005). A diarylquinoline drug active on the ATP synthase of Mycobacterium tuberculosis. Science *307*, 223–227. https://doi.org/10.1126/science.1106753.

38. Segala, E., Sougakoff, W., Nevejans-Chauffour, A., Jarlier, V., and Petrella, S. (2012). New mutations in the mycobacterial ATP synthase: new insights into the binding of the diarylquinoline TMC207 to the ATP synthase C-ring structure. Antimicrob. Agents Chemother. *56*, 2326–2334. https://doi.org/10.1128/AAC.06154-11.

39. Haagsma, A.C., Podasca, I., Koul, A., Andries, K., Guillemont, J., Lill, H., and Bald, D. (2011). Probing the interaction of the diarylquinoline TMC207 with its target mycobacterial ATP synthase. PLoS One *6*, e23575. https://doi.org/10.1371/journal.pone.0023575.

40. Karmakar, M., Rodrigues, C.H.M., Holt, K.E., Dunstan, S.J., Denholm, J., and Ascher, D.B. (2019). Empirical ways to identify novel Bedaquiline resistance mutations in AtpE. PLoS One *14*, e0217169. https://doi.org/10.1371/journal.pone.0217169.

41. Preiss, L., Langer, J.D., Yildiz, Ö., Eckhardt-Strelau, L., Guillemont, J.E.G., Koul, A., and Meier, T. (2015). Structure of the mycobacterial ATP synthase Fo rotor ring in complex with the anti-TB drug bedaquiline. Sci. Adv. *1*, e1500106. https://doi.org/10.1126/sciadv.1500106.

42. Nguyen, T.V.A., Anthony, R.M., Bañuls, A.L., Nguyen, T.V.A., Vu, D.H., and Alffenaar, J.W.C. (2018). Bedaquiline resistance: its emergence, mechanism, and prevention. Clin. Infect. Dis. *66*, 1625–1630. https://doi.org/10.1093/cid/cix992.

43. Kadura, S., King, N., Nakhoul, M., Zhu, H., Theron, G., Köser, C.U., and Farhat, M. (2020). Systematic review of mutations associated with resistance to the new and repurposed Mycobacterium tuberculosis drugs bedaquiline, clofazimine, linezolid, delamanid and pretomanid. J. Antimicrob. Chemother. *75*, 2031–2043. https://doi.org/10.1093/jac/dkaa136.

44. Andersson, D.I., and Hughes, D. (2010). Antibiotic resistance and its cost: is it possible to reverse resistance? Nat. Rev. Microbiol. *8*, 260–271. https://doi.org/10.1038/nrmicro2319.

45. Rifat, D., Campodónico, V.L., Tao, J., Miller, J.A., Alp, A., Yao, Y., and Karakousis, P.C. (2017). In vitro and in vivo fitness costs associated with Mycobacterium tuberculosis RpoB mutation H526D. Future Microbiol. *12*, 753–765. https://doi.org/10.2217/fmb-2017-0022.

46. Knight, G.M., Colijn, C., Shrestha, S., Fofana, M., Cobelens, F., White, R.G., Dowdy, D.W., and Cohen, T. (2015). The distribution of fitness costs of resistance-conferring mutations is a key determinant for the future burden of drug-resistant tuberculosis: a model-based analysis. Clin. Infect. Dis. *61*, S147–S154. https://doi.org/10.1093/cid/civ579.

47. Cohen, T., Sommers, B., and Murray, M. (2003). The effect of drug resistance on the fitness of Mycobacterium tuberculosis. Lancet Infect. Dis. *3*, 13–21. https://doi.org/10.1016/s1473-3099(03)00483-3.

48. Gagneux, S., Long, C.D., Small, P.M., Van, T., Schoolnik, G.K., and Bohannan, B.J.M. (2006). The competitive cost of antibiotic resistance in. Science *312*, 1944–1946. https://doi.org/10.1126/science.1124410.

49. 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., Sassetti, C.M., et al. (2017). Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. Nat. Microbiol. *2*, 16274. ARTN 16274. https://doi.org/10.1038/nmicrobiol.2016.

50. Yan, M.Y., Li, S.S., Ding, X.Y., Guo, X.P., Jin, Q., and Sun, Y.C. (2020). A CRISPR-assisted nonhomologous end-joining strategy for efficient genome editing in Mycobacterium tuberculosis. mBio *11*. ARTN e02364-19. https://doi.org/10.1128/mBio.02364-19.

51. Ghosh, S., Goldgur, Y., and Shuman, S. (2020). Mycobacterial DNA polymerase I: activities and crystal structures of the POL domain as apoenzyme and in complex with a DNA primer-template and of the full-length FEN/EXO POL enzyme. Nucleic Acids Res. *48*, 3165–3180. https://doi.org/10.1093/nar/gkaa075.

52. van Kessel, J.C., and Hatfull, G.F. (2007). Recombineering in Mycobacterium tuberculosis. Nat. Methods *4*, 147–152. https://doi.org/10.1038/nmeth996.

53. Pashley, C.A., and Parish, T. (2003). Efficient switching of mycobacteriophage L5-based integrating plasmids in Mycobacterium tuberculosis. FEMS Microbiol. Lett. *229*, 211–215. https://doi.org/10.1016/S0378-1097(03)00823-1.