Published in final edited form as:

Gene. 2010 November 15; 468(1-2): 8–19. doi:10.1016/j.gene.2010.07.012.

Taking phage integration to the next level as a genetic tool for mycobacteria

Jason Huff¹, Agata Czyz², Robert Landick², and Michael Niederweis^{1,*}

¹Department of Microbiology, University of Alabama at Birmingham, 609 Bevill Biomedical Research Building, 845 19th Street South, Birmingham, AL 35294, USA

²Departments of Biochemistry and Microbiology, University of Wisconsin-Madison, WI 53706, USA

Abstract

Genes must be stably integrated into bacterial chromosomes for complementation of gene deletion mutants in animal infection experiments or to express antigens in vaccine strains. However, with currently available vectors it is cumbersome to create multiple, stable, unmarked chromosomal integrations in mycobacteria. Here, we have constructed a novel integration vector for mycobacteria that enables expression of genes from a cassette protected from transcriptional interference by bi-directional transcriptional terminators proven to be highly efficient in *in vitro* transcription termination assays. Removal of the integrase gene by a site-specific recombinase, easily identifiable by loss of a backbone reporter gene, stabilizes the integration cassette and makes this vector ideally suitable for infection experiments. This integration vector can be easily adapted to different mycobacteriophage attachment sites (*attB*) due to its modular design. Integration of a *gfp* expression cassette at the L5, Giles and Ms6 *attB* sites in the chromosomes of *Mycobacterium smegmatis* and *Mycobacterium tuberculosis* yielded identical *gfp* expression levels, indicating that none of these sites are compromised for gene expression. The copy number of pAL5000-based extrachromosomal plasmids is 23 in *M. smegmatis* as determined by quantitative real-time PCR and accounts for the previously observed drastic reduction of gene expression upon integration of plasmids into the chromosome of mycobacteria. Gfp expression and fluorescence of *M. smegmatis* and *M. tuberculosis* strains with multiple integrations of *gfp* increased concomitantly with the copy number demonstrating that these vectors can be used to generate stronger phenotypes and/or to analyze several genes simultaneously *in vivo*.

Keywords

plasmid; modular design; expression; chromosome; pAL5000; copy number; transcription termination

1. INTRODUCTION

Genetic constructs can be integrated into mycobacterial chromosomes using plasmids carrying a mycobacteriophage integrase gene (*int*) and the corresponding phage attachment

© 2010 Elsevier B.V. All rights reserved.

*For correspondence: Phone: +1-205-996-2711 Fax: +1-205-934-9256 mnieder@uab.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

site (*attP*) (Lee *et al.*, 1991; Ribeiro *et al.*, 1997; Freitas-Vieira *et al.*, 1998; Pham *et al.*, 2007; Morris *et al.*, 2008). Commonly used integration vectors based upon mycobacteriophage L5, henceforth referred to as L5, contain the L5 *int* gene and the L5 *attP* sequence, which is homologous to the bacterial attachment site *attB* residing within the *glyV* tRNA gene (Lee *et al.*, 1991) (Fig. 1). The phage integrase along with the mycobacterial integration host factor mIHF (Pedulla *et al.*, 1996) catalyze site-specific recombination resulting in the integration of the plasmid at the *attB* site. Integration vectors are genetic tools required to maintain stable gene expression, for example during animal infection studies. However, several caveats exist to generating stably integrated expression cassettes. For instance, integrase-mediated excisive recombination destabilizes the integrated plasmid at the L5 attachment site (Springer *et al.*, 2001). Additionally, expression problems have been reported for genes integrated at the L5 *attB* site, often with disproportionately lower expression in comparison to the copy number of pAL5000-based episomal plasmids (Stover *et al.*, 1991) and Mailaender and Niederweis, unpublished results). Further, the number of genes that can be expressed from a single integration vector is limited (Saviola and Bishai, 2004).

Several approaches have been developed to circumvent these obstacles. For example, co-electroporation of an *attP*-containing integration vector and a nonreplicating *int*-containing vector provides integrase activity in trans which is then subsequently lost, thus stabilizing the integrated vector (Pena *et al.*, 1997). Alternatively, the *int* gene can be removed from the chromosome via subsequent expression of site-specific recombinases such as the bacteriophage P1 Cre recombinase (Song and Niederweis, 2007) or the *Saccharomyces cerevisiae* FLP recombinase (Stephan *et al.*, 2004). The presence of an *attB* site on an integration vector reintroduces a new bacterial attachment site after integration into the chromosome and enables serial integrations (Saviola and Bishai, 2004). Another option is to transform separate plasmids containing the integration machinery from different mycobacteriophages, which allows genes to be integrated at different chromosomal sites (Freitas-Vieira *et al.*, 1998). To date, it has not been determined whether integration of a plasmid at other locations in the chromosome leads to higher expression levels in mycobacteria compared to the L5 *attB* site.

The ability to create multiple, stable, unmarked integrations to express several genes simultaneously or to simply increase overall expression by increasing their copy number would be advantageous. In this study, a new integrative plasmid was designed with a multiple cloning site flanked by bi-directional intrinsic transcriptional terminators. A backbone excisable by Cre recombinase enables the removal of the L5 *int* gene, resistance cassette, and remaining plasmid backbone. This can be monitored by loss of the *xylE* backbone reporter gene. As all genetic components are flanked by single restriction sites, this plasmid can be easily adapted to alternative integration machineries from other mycobacteriophages. A series of plasmids for integration at the mycobacteriophages L5, Giles, and Ms6 attachment sites was constructed and expression levels were compared in both fast and slow growing mycobacteria. These plasmids will aid in the examination of mycobacterial genes in mouse infection experiments and/or in vaccine strain development where stable gene expression is required.

2. MATERIALS AND METHODS

2.1. Chemicals, enzymes and DNA

Hygromycin B was purchased from Calbiochem. All other chemicals were purchased from Merck, Roche or Sigma at the highest purity available. Enzymes for DNA restriction and modification were purchased from New England Biolabs. Isolation and modification of

DNA was performed as described (Ausubel *et al.*, 1987). Oligonucleotides were obtained from Integrated DNA Technologies.

2.2. Bacterial strains and growth conditions

Escherichia coli DH5 α , used for cloning experiments, was routinely grown in Luria-Bertani broth at 37°C. *M. smegmatis* strains were grown at 37°C in Middlebrook 7H9 medium (Difco) supplemented with 0.2% glycerol and 0.025% Tyloxapol or on Middlebrook 7H10 agar (Difco) supplemented with 0.2% glycerol. *M. tuberculosis* mc²6230 is an avirulent derivative of *M. tuberculosis* H37Rv (Sambandamurthy *et al.*, 2006) and was grown at 37°C in Middlebrook 7H9 liquid medium supplemented with 0.2% glycerol, 10% oleic acid-albumin-dextrose-catalase enrichment (OADC, Remel), 0.025% Tyloxapol, 0.2% casamino acids, and 24 μ g/mL pantothenate or on Middlebrook 7H10 agar supplemented with 0.2% glycerol, 10% OADC, 0.2% casamino acids, and 24 μ g/mL pantothenate. Antibiotics were used when required at the following concentrations: hygromycin (200 μ g/ml for *E. coli*; 50 μ g/ml for mycobacteria) and kanamycin (50 μ g/ml for *E. coli*; 30 μ g/ml for mycobacteria).

2.3. Synthesis of *xylE_m*, attachment sites, polylinker, loxP sites, and bi-directional terminators

The 1,820 bp fragment between *SspI* and *XmaI* of pML1342 was synthesized by Geneart (Burlingame, CA) and cloned into the plasmid pML1320. The sequence of the *xylE* gene from *Pseudomonas putida* (NCBI accession: V01161.1) was codon-optimized for efficient expression in *Mycobacterium tuberculosis* H37Rv. The *xylE* gene sequence (*xylE_m*) was optimized using codons preferred by mycobacteria (<http://www.kazusa.or.jp.codon/>) as previously described (Song and Niederweis, 2007). The codon usage was further optimized by removing all rare codons that occur in the *M. tuberculosis* genome with a frequency of less than 10 per 1,000 codons. The sequence of *xylE_m* was submitted to Genbank (accession #: GU128960). The *xylE_m* gene sequence was fused to the *p_{wmyc}* promoter (Kaps *et al.*, 2001) and a transcription terminator based on the intrinsic terminator *trpT* downstream of the *E. coli* tryptophan operon (Wu and Platt, 1978) known as *tt_{trpA}*⁺ (AGCCCGCCTAATGGCGGGCTTTTTTTT) was added 15 bp downstream of the *xylE_m* stop codon. Two loxP sites (ATAACTTCGTATAATGTATGCTATACGAAGTTAT) were positioned to flank the L5 *attP* and *attB* attachment sites (Lee *et al.*, 1991), two bi-directional terminator sequences, and a multiple cloning site polylinker. The synthetic bi-directional terminators *tt_{sbiA}* and *tt_{sbiB}* were designed for efficient transcription termination by inclusion of ten leading adenosine residues, seven GC base pairs in the hairpin stem, four unpaired bases in the hairpin loop, and ten thymidine residues downstream of the hairpin (de Hoon *et al.*, 2005).

2.4. Construction of plasmids

The *E. coli* origin of replication was amplified from pML113 (Wolschendorf *et al.*, 2007) by PCR using the primers 99 and 967 (Table S1) and digested with *ClaI* and *SspI*. The resulting 954 bp fragment was cloned into the 2,541 bp fragment of *ClaI/SspI* digested pML113 in order to remove the *bla* gene and to correctly position a required *SspI* restriction site downstream of the *E. coli* replication origin, generating the plasmid pML603. The gene encoding the mycobacteriophage L5 integrase was obtained from pML102 (Wolschendorf *et al.*, 2007) via a one-step combined chain reaction (CCR) to remove the internal *SphI* restriction site using the end primers 1147 and 1376 and the internal *SphI* mutagenesis primer 1377. The L5 *int* gene was ligated as a blunt-end fragment with the 3,421 bp fragment of *HpaI/ClaI* digested and filled-in pML603 to generate the plasmid pML1321. Digestion of pML1321 with *XmaI* and *SspI* resulted in a fragment containing the hygromycin resistance gene, the L5 integrase gene, and the *E. coli* origin of replication. This 3,642 bp fragment was ligated with the 1,812 bp fragment of *XmaI/SspI* digested pML1320

to generate the plasmid pML1330. Because recombination of this plasmid occurred during passage through *E. coli* (Fig. S1), the L5 *attB* site was removed by digestion with *XhoI* and *SpeI*, the overhangs were filled-in, and the resulting 5,404 bp fragment was ligated to generate the new L5 integration vector pML1342.

The *gfp_m²⁺* gene codon optimized for expression in mycobacteria encodes a green fluorescence protein (Gfp) with mutations that enhance fluorescence and protein stability (Steinhauer *et al.*, 2010). In order to clone the *gfp_m²⁺* gene into pML1342, the *gfp_m²⁺* gene was cut out from pMN437 (Steinhauer *et al.*, 2010) via digestion with *ClaI* and *PmeI*, and the resulting 1,122 bp fragment was ligated into the polylinker of similarly digested pML1342 to generate the integrative, Gfp-expressing plasmid pML1335. The same *gfp_m²⁺* digestion fragment was also ligated into *ClaI/HpaI* digested pMV306kan (Stover *et al.*, 1991), a standard mycobacterial integration vector containing L5 *int* and *attP* to create pML1337.

To adapt pML1335 to the mycobacteriophage Giles integration system, the *MluI* restriction site in *gfp_m²⁺* had to first be removed. This was accomplished by CCR with the end primers 145 and 131 and the mutagenesis primer 1598. The 851 bp CCR product was digested with *PacI* and *HindIII* and ligated into similarly digested pML1335, replacing *gfp_m²⁺* with a *gfp_m²⁺* carrying a silent mutation of the *MluI* site, generating pML1347. The mycobacteriophage Giles *int* gene was amplified from pGH1000A (Morris *et al.*, 2008) by PCR using the end primers 1605 and 1606. The resulting 1,344 bp product was digested with *KpnI* and ligated into the 5,043 bp fragment of *KpnI/HpaI* digested pML1347 to generate pML1355. The Giles *attP* site was then amplified from pGH1000A by PCR with the end primers 1603 and 1604. The resulting 338 bp product was digested with *MluI* and *AscI* and ligated into the similarly digested 6,127 bp fragment of pML1355 to generate the mycobacteriophage Giles integration vector pML1357.

To adapt pML1347 to the mycobacteriophage Ms6 integration system, the Ms6 *attP* site was amplified from pNIP40b (kindly provided by Dr. Nathalie Winter, Institut Pasteur, Unité de Génétique Mycobactérienne) by PCR with the end primers 1686 and 1687. The resulting 543 bp product was digested with *MluI* and *AscI* and ligated into the 6,231 bp fragment of similarly digested pML1347 to generate pML1360. The Ms6 *int* gene was then amplified from pNIP40b by PCR using the end primers 1688 and 1689. The resulting 1,278 bp product was digested with *KpnI* and ligated with the 5,323 bp fragment of *HpaI/KpnI* digested pML1360 to generate the mycobacteriophage Ms6 integration vector pML1361. Exchange of the *hyg* gene with the *aph* gene in pML1361 generated pML2300. This was done by amplification of *aph* from pMS2kan (Kaps *et al.*, 2001) by PCR using the primers 1749 and 1750, digestion of the resulting 1,421 bp product with *NruI* and *NsiI*, and ligation of the resulting 1,327 bp fragment into the 5,494 bp fragment of *NruI/NsiI* digested pML1361.

2.5. Construction of transcription termination assay linear templates

The plasmid pRLG3748 contains a -35 consensus promoter and two *E. coli* terminators: *rrnBT1* and *rrnBT2* located 166 bp and 190 bp downstream from the transcription start site, respectively (Gaal *et al.*, 2001). Two restriction sites, *BamHI* and *NcoI*, were introduced in pRLG3748 by site directed mutagenesis located before (*BamHI*) and after (*NcoI*) the *rrnBT1* terminator. These restriction sites were used to replace *rrnBT1* with either the *tt_{sbiA}* or the *tt_{sbiB}* terminators. Oligonucleotides containing *tt_{sbiA}* and *tt_{sbiB}* terminator sequences were purchased from Integrated DNA Technologies. Plasmids were confirmed by restriction analysis and sequencing via a BigDye V3.1 Cycle Sequencing Kit (Applied Biosystems).

Linear templates for *in vitro* transcription were amplified from pRLG3748-derived plasmids containing either the *tt_{sbiA}* or the *tt_{sbiB}* terminators using primers complementary to

sequences upstream of the -35 consensus promoter and downstream of the *rrnBT2* terminator. The resulting 686 bp linear templates, containing the promoter-test terminator-*rrnBT2* terminator, were subsequently gel purified by a QIAquick gel extraction kit (Qiagen) followed by phenol extraction and precipitation. DNA was resuspended in TE buffer and stored at -20°C for further experiments.

2.6. In vitro transcription termination assay

In vitro transcription complexes were generated by addition of 25 nM linear DNA template, 50 nM *E. coli* RNA polymerase (RNAP) or recombinant *M. bovis* RNA polymerase containing σ^A (r*M.b*RNAP), and bovine serum albumin (25 $\mu\text{g/ml}$ final) in transcription buffer (40 mM Tris-HCl, pH 7.9; 10 mM KCl; 10 mM MgCl_2 ; 1 mM DTT; 0.1 mM EDTA). Transcription reactions were initiated by addition of rNTPs (200 μM ATP, 200 μM GTP, 200 μM CTP and 10 μM UTP final) and 2 μM [α - ^{32}P] UTP after a 10 min incubation at 37°C . To ensure a single round of transcription, the inhibitor of re-initiation (heparin), was subsequently added to a final concentration of 25 $\mu\text{g/ml}$. Samples were collected after 30 min of incubation and mixed with an equal volume of a $2\times$ STOP buffer (8 M urea, $1\times$ TBE buffer, 50 μM EDTA, 0.3 mg/ml each bromophenol blue and xylene cyanol). RNA products were resolved in 6% denaturing agarose gels (19:1 acrylamide/bisacrylamide, 8 M urea in $0.5\times$ TBE buffer). Relative amounts of each RNA band were determined via phosphorimaging and the ImageQuant Total Lab software. Efficiency of termination was calculated as a fraction of the particular terminated transcript versus the sum of all terminated and read-through transcripts. Linear template containing the well-studied *E. coli trpA* terminator was included as a control. All experiments were repeated three times.

2.7. Construction of bacterial strains

Integrative vectors were electroporated into *M. smegmatis* mc²155 or *M. tuberculosis* mc²6230 and positive clones were selected for by plating on Middlebrook 7H10 agar containing appropriate antibiotics. Strains constructed from plasmids containing the *gfp_m²⁺* gene were confirmed via observation of Gfp fluorescence. Strains constructed from plasmids containing the *xylE_m* gene were confirmed by dropping 50 μl of 1% catechol (Sigma) on top of the colonies and incubating at room temperature for approximately 5 to 10 minutes until a distinct yellow color appeared.

Strains destined for plasmid backbone removal were transformed with the plasmid pCreSacB and selected for on kanamycin. This allowed for expression of the Cre recombinase and concomitant removal of integrated backbone DNA between flanking *loxP* sites, thus generating an unmarked strain. Plating on 7H10 agar containing either 10% or 2% sucrose for *M. smegmatis* or *M. tuberculosis* strains, respectively, counter-selected against pCreSacB. The absence of the chromosomal *hyg* cassette and the pCreSacB-encoded *aph* cassette was confirmed by growth on 7H10 agar but no growth on 7H10 agar containing either hygromycin or kanamycin. The genomic organization of all strains was examined by colony PCR using PuReTaqTM Ready-To-GoTM PCR beads (GE Healthcare) according to the manufacturer's protocol with appropriate primers.

2.8. Quantification of Gfp fluorescence

Bacterial strains were incubated in Middlebrook 7H9 liquid medium on a shaker at 37°C to an OD₆₀₀ of 0.8-1.0. Triplicate cultures were diluted in 7H9 medium to the same optical density and 200 μl were transferred to wells of a black, 96-well microtiter plate with flat, transparent bottoms (Costar #3631 from Corning). Fluorescence was measured via a Synergy HT plate reader (Bio-Tek) by excitation at 485 nm and emission at 528 nm. The optical density of bacterial cultures at 600 nm was measured, and fluorescence of each sample was normalized to OD₆₀₀.

2.9. Visualization of Xyle and Gfp reporter activity

Bacterial cultures were incubated in 7H9 medium on a shaker at 37°C and then standardized to an OD₆₀₀ of 1.0. Cultures were serially diluted and plated onto 7H10 agar plates. After incubation at 37°C, pictures of individual colonies were taken under standard white light with 1 ms exposure times using an AxioCam MRc camera (Zeiss) mounted on a Stemi 2000-C compound light microscope (Zeiss). To visualize Xyle activity, 50 µl of 1% catechol were dropped on the tops of colonies and incubated at room temperature for approximately 5 to 10 minutes until a distinct yellow color appeared (Curcic *et al.*, 1994). Pictures were taken under the same conditions as white light pictures. To visualize Gfp fluorescence, the white light was turned off and an excitation wavelength of 485 nm was applied while emission was filtered at 528 nm.

2.10. Determination of plasmid copy numbers in mycobacteria by qPCR

Plasmid copy numbers were determined using a previously published protocol (Skulj *et al.*, 2008) with minor modifications. Strains grown to an OD₆₀₀ of 1.0 were harvested and washed three times in PBS containing 0.025% Tyloxapol. Cells were then pelleted, weighed, and 10 µL 1% SDS lysis buffer were added per 1 mg cells. Cells were then frozen on dry ice and immediately heated to 95°C for 15 min. Control experiments showed that high SDS concentrations impeded amplification and that low-power sonication resulted in reduced Ct values (not shown), indicating greater release of cellular DNA. Hence, cell lysates were immediately diluted 20-fold in ultrapure, nuclease-free water (Ambion) and subjected to sonication on ice (Misonix sonicator 3000) with 3 watts for 5 min in 1 sec intervals. Centrifugation of insoluble debris increased Ct values (not shown), indicating that total DNA release from debris was not achieved despite the harsh treatments. Thus, centrifugation was omitted in further experiments. Lysates were stored at -20°C without affecting Ct values.

For qPCR, lysates were thawed and heated to 95°C for 5 min to dissociate debris aggregates, and diluted 10-fold in 0.005% SDS in ultrapure water as cellular debris inhibited amplification efficiency (not shown). From this solution, serial 5-fold dilutions were made in 0.005% SDS. Three to five replicates of three to six dilutions for each sample were analyzed. 20 µL amplification reactions consisted of 10 pmol of each primer (see below), 0.05% dimethyl sulfoxide (DMSO), 5 µL sample, 2 µL ultrapure water, and 10 µL Evagreen supermix (Bio-Rad) and were placed into individual wells of an iCycler iQ PCR plate (Bio-Rad). Amplification and detection were performed in a MyiQ2 real time PCR detection system (Bio-Rad). Amplification efficiency and Ct values were calculated by the iCycler software and used to determine copy number according to the equation: copy number = $(E_c)^{C_{tc}} / (E_p)^{C_{tp}}$ (Skulj *et al.*, 2008), where E_c and E_p refer to the efficiency of amplification from the chromosome and plasmid, respectively, and C_{tc} and C_{tp} refer to the Ct value of the chromosome and plasmid reactions, respectively. Copy numbers were determined for each dilution and were averaged. Only dilutions yielding Ct values lower than 30, Ct values at least five less than no-template controls, and R^2 correlation coefficients >0.99 were used in final calculations. It was observed that the above equation did not produce reliable copy numbers when the efficiencies between the two reactions differed by more than 1.0%. Hence, only reactions with efficiencies differing by less than 1% were compared.

The program Primer3 (<http://frodo.wi.mit.edu/primer3>) was used to identify primers and amplicons optimal for qPCR. Primer3 identified the primers 2085 and 2087 which amplified a 180 bp amplicon in the *E. coli* replication origin present on all plasmids used in this study (Table S1). Primer3 also identified a chromosomal control amplicon located near the *L5 attB* site. This amplicon consisted of the 170 bp fragment positioned between nucleotides 4,729,312 and 4,729,481 of the *M. smegmatis* mc²155 genome

(<http://cmr.jcvi.org/cgi-bin/CMR/GenomePage.cgi?org=gms>) and was amplified with the primers 2038 and 2039.

2.11. Western blot analysis of gfp expression

M. smegmatis strains were grown overnight to an OD₆₀₀ of 1.0 and cells were harvested by centrifugation. Cells were washed in PBS pH 7.4 containing 0.025% Tyloxapol three times and then transferred to pre-weighed 1.5 mL tubes. Cell pellets were resuspended in 10 µL PBS per 1 mg wet weight of cells and immediately frozen at -80°C. After thawing at 50°C, cell suspensions were sonicated for 30 sec, 3 watts, 1 sec on/1 sec off. Protease inhibitor cocktail (Roche) on ice and cell debris was pelleted down. Equal volumes of supernatants were loaded onto a 10% polyacrylamide gel and proteins were separated by gel electrophoresis. After transfer to a nitrocellulose membrane, Gfp was detected by a mouse anti-Gfp monoclonal antibody (Sigma). Quantification by pixel densitometry of Gfp bands was performed using the Labworks 4.6 software (UVP, Inc.). Because *M. smegmatis* ML263 contained the pMV306kan-derived plasmid pML1337, it served as a control for single copy integrations. Thus, expression in this strain was set to 1, to which all other strains were compared.

3. RESULTS AND DISCUSSION

3.1. Improvements to the mycobacteriophage L5 integration system

Our aim was to design a new plasmid for integration into mycobacterial chromosomes that would (i) be easily modifiable due to a modular design, for example for adaptation to multiple integration sites, (ii) prevent transcriptional interference, (iii) allow for stable integration of expression cassettes, and (iv) contain a reporter gene for rapid identification of the backbone. To achieve these goals, a cassette containing two bi-directional intrinsic transcriptional terminators (*tt_{sbiA}* and *tt_{sbiB}*) flanking a polylinker was synthesized (Fig. 2). The L5 *attP* site was positioned adjacent to the protected polylinker cassette. Flanking these components are two *loxP* sites for backbone removal, which can be monitored via the *xylE* gene of *Pseudomonas putida* (Curcic *et al.*, 1994) that was codon-optimized for expression in mycobacteria (*xylE_m*). The entire *xylE_m-loxP-attP-tt_{sbiA}-polylinker-tt_{sbiB}-loxP* cassette was synthesized and cloned into a backbone containing the hygromycin resistance gene, the L5 integrase gene, and the *E. coli* origin of replication, generating the plasmid pML1342 (Table 1 and Fig. 2).

Previous reports have described apparently disproportionately reduced expression when genes were integrated at the L5 attachment site (Stover *et al.*, 1991; Mailaender and Niederweis, unpublished results). Promoter occlusion from RNA polymerases extending from proximal promoters such as the *glyV* promoter may cause such reduced expression of genes integrated at the L5 *attB* site (Adhya and Gottesman, 1982). To protect the expression cassette as well the surrounding genome from transcriptional interference, we designed the synthetic bi-directional intrinsic transcriptional terminators A and B (*tt_{sbiA}* and *tt_{sbiB}*) (Fig. 2). These terminators were optimized for bi-directional function and high termination efficiency by incorporating ten upstream adenine nucleotides, seven G-C base pairs in the stem, four unpaired nucleotides in the hairpin loop, and ten downstream thymine nucleotides (de Hoon *et al.*, 2005).

Though inefficient, it has been shown that continued expression of the L5 *int* gene may cause plasmid instability due to L5 integrase-mediated, Xis-independent excisive recombination (Springer *et al.*, 2001). One solution to this problem is to encode integrase on a second nonreplicating vector and co-transform both plasmids simultaneously. However, reduced transformation efficiencies using this approach were reported (Springer *et al.*,

2001). Therefore, we sought to remove the integrase gene without reducing transformation efficiency by encoding *int* on an excisable backbone. By positioning two *loxP* sites flanking the *hyg-int-colEI-xylE_m* backbone, excision of the backbone upon expression of the Cre recombinase gene leaves behind a stably integrated and terminator-protected expression cassette. Integration of pML1342 and removal of the backbone can be monitored by the backbone reporter gene *xylE_m*. In summary, the modular design of pML1342 contains a polylinker of single restriction endonuclease sites flanked by protective bi-directional terminators, a removable backbone for stable integrations equipped with a codon-optimized reporter gene (*xylE_m*), and single restriction sites flanking all genetic components to enable easy replacement of each module.

3.2. Identification of clones via reporter gene activities

To complement mutants or introduce genes into vaccine strains, it is desirable to create stable integrations in the bacterial chromosome as extrachromosomally replicating plasmids cannot be selected for during infection experiments. To demonstrate that unmarked integrations can be created using pML1342, we used reporter genes to monitor both the integration of an expression cassette and excision of the destabilizing *int*-containing backbone. The transformation efficiency (Table 2) of pML1342 was reduced in comparison to the commonly used integration vector pMV306kan in *M. smegmatis* mc²155 (Snapper *et al.*, 1990) and in *M. tuberculosis* mc²6230 ($\Delta RD1$, $\Delta panCD$) (Sambandamurthy *et al.*, 2006) (Table 1), which may be attributed to the larger size of pML1342 and differences in the antibiotic resistance they confer. To visualize expression of pML1342 and pMV306kan transformants, we inserted at the polylinker of both plasmids *gfp_m*²⁺, a variant of the *gfp* gene encoding green fluorescent protein which was codon-optimized for expression in mycobacteria and carries mutations for enhanced fluorescence and stability (Steinhauer *et al.*, 2010). This generated the pML1342-derived plasmid pML1335 and the pMV306kan-derived plasmid pML1337. *M. smegmatis* transformants with integrated pML1335, referred to as *M. smegmatis* ML261 (Table 1), were easily identified by Gfp fluorescence and XylE activity (Fig. 3). Removal of the pML1335 backbone upon expression of Cre recombinase was visualized by loss of XylE activity in the resulting *M. smegmatis* ML262, a stable strain void of the *int* gene. Clones with pML1337 (*M. smegmatis* ML263) were identifiable by Gfp fluorescence alone. The *int*-containing backbone of pML1337 was not excisable as this pMV306-derived vector does not contain *loxP* sites. *M. smegmatis* mc²155 carrying the *gfp_m*²⁺ replicative plasmid pMN437 (Table 1) displayed much higher Gfp fluorescence in colonies (Fig. 3), reflecting the higher copy number of episomal pAL5000 plasmids (Ranes *et al.*, 1990; Gavigan *et al.*, 1997). Transformation with the empty pML1342 vector generated *M. smegmatis* ML264 and served as a negative control for Gfp fluorescence, but clones were still identifiable via catechol reactivity resulting from *xylE_m* in the plasmid backbone.

3.3. Expression of *gfp_m*²⁺ from the L5 attB site in mycobacteria

Episomal plasmids containing the pAL5000 origin of replication were estimated to have a copy number of three (Ranes *et al.*, 1990) to eight (Gavigan *et al.*, 1997) in mycobacteria. Expression of genes from the L5 *attB* site has been reported to be disproportionately less than expected based on their reduced copy number compared to pAL5000 episomal vectors (Stover *et al.*, 1991). It is possible that transcriptional interference from nearby promoters may inhibit expression of genes at the L5 *attB* site, and the positioning of bi-directional intrinsic transcriptional terminators flanking the expression cassette may prevent this. To quantitatively determine expression levels from unprotected versus terminator-protected *gfp_m*²⁺ expression cassettes integrated at the L5 *attB* site, the Gfp fluorescence of *M. smegmatis* ML261 (protected cassette present in pML1335) was compared to that of *M. smegmatis* ML263 (unprotected cassette present in pMV306-derived pML1337). No

difference in fluorescence was observed between these two strains (Fig. 4). Excision of the pML1335 backbone in ML261 (generating ML262) had also no significant effect on fluorescence. The pAL5000-based plasmid pMN437 provided a 14-fold increased fluorescence in *M. smegmatis* compared to the control strain ML263, consistent with previous estimates of pAL5000 copy numbers (Gavigan *et al.*, 1997).

Similar results were obtained for *M. tuberculosis* using the same vectors (Fig. 5). Fluorescence of *M. tuberculosis* carrying the episomal gfp_m^{2+} vector pMN437 was 9-fold higher than that of the strain integrated with the standard, unprotected gfp_m^{2+} cassette (*M. tuberculosis* ML441). Both the strain with integrated pML1335 containing the protected gfp_m^{2+} cassette (ML440) and the strain with the pML1335-backbone removed (ML442) did not show significantly increased fluorescence levels. These results show that pML1342-based vectors can functionally integrate and express genes in both *M. smegmatis* and *M. tuberculosis* and that episomal expression of genes from pAL5000-based plasmids results in higher expression levels. However, the transcriptional terminators flanking the expression cassette did not significantly increase expression compared to other commonly used integration vectors.

3.4. Efficiencies of the synthetic transcriptional terminators

Previous reports indicated that transcriptional terminators in mycobacteria might be different from those in other bacteria (Unniraman *et al.*, 2001) and are difficult to predict (de Hoon *et al.*, 2005). It is conceivable that the synthetic transcriptional terminators in our integration vector may not work efficiently in mycobacteria. Therefore, we sought to determine the functionality of the terminators *tt_{sbi}A* and *tt_{sbi}B* in an *in vitro* transcription termination assay. In this assay, transcription initiated from a consensus promoter may be terminated by an intrinsic terminator encoded downstream referred to as the test terminator. Alternatively, the RNA polymerase (RNAP) may proceed through the test terminator and be terminated by the *rrnBT2* terminator or read-through both terminators to the end of the template (Fig. 6A). As a positive control for the test terminator, the well-characterized *trpA* terminator from *E. coli* was included. Transcription elongation by *E. coli* RNAP was terminated with 64% efficiency by *trpA*. Conversely, *tt_{sbi}A* and *tt_{sbi}B* terminated transcription with 71% and 88% efficiency, respectively (Fig. 6B). While efficiency of *trpA* was the same when recombinant RNAP from *M. bovis* was used, *tt_{sbi}A* and *tt_{sbi}B* terminated the mycobacterial RNAP with efficiencies near 100% (Fig. 6C). Thus, the bi-directional terminators *tt_{sbi}A* and *tt_{sbi}B* are indeed functional and highly efficient in terminating transcription by RNA polymerases from different bacteria. These results agree with previous findings that canonical Rho-independent transcriptional terminators were efficient in mycobacteria (Timm *et al.*, 1994) and the fact that all bacterial RNA polymerases are highly homologous and are terminated by similar mechanisms (Artsimovitch *et al.*, 2000).

Taken together, these results indicate that gene expression levels are not strongly affected by site-specific effects at the L5 *attB* site in *M. smegmatis* or *M. tuberculosis*. It may be that the putative intrinsic terminator within the L5 *attP* site upstream of the core already prevents read-through transcription from the *glyV* gene (Lee *et al.*, 1991). Addition of flanking bi-directional terminators, however, may prevent read-through transcription from the integrated cassette into the surrounding genome and thereby avoid artificial phenotypes resulting from integration.

3.5. Copy number of pAL5000 plasmids

As transcriptional interference does not play a role in the apparently disproportionately reduced expression at the L5 *attB* site, an alternative explanation may be that the copy number of pAL5000 plasmids is higher than previously reported. The relative increase in

fluorescence intensity when gfp_m^{2+} was expressed from the episomal pMN437 vector compared to an integrative vector was 14-fold and 9-fold in *M. smegmatis* and *M. tuberculosis*, respectively. In the absence of transcriptional interference, these results suggested that the copy number of pAL5000-based plasmids may indeed be higher than previously reported. A copy number of three was determined by Southern blot analysis (Ranes *et al.*, 1990) while a single-cell antibiotic resistance assay placed the copy number closer to eight (Gavigan *et al.*, 1997). However, both methods are prone to error. DNA isolation required for Southern blots is often incomplete and difficult to reproduce quantitatively. Single-cell antibiotic resistance assumes the relationship between copy number of a plasmid and antibiotic resistance is linear without experimental verification. Further, the resistance gene is expressed in different genetic environments and cfu determinations require an arbitrary growth endpoint whose choice can significantly alter the results. An alternative is quantitative real-time PCR (qPCR) of whole cell lysates which provides absolute copy numbers of specific target DNA sequences (Fu *et al.*, 2009). Thus, the copy number of the pAL5000-containing plasmid pMN437 in *M. smegmatis* was determined by qPCR using a published protocol (Skulj *et al.*, 2008) with slight modifications. As a control, the copy number of the plasmid pML1337 integrated in *M. smegmatis* ML263 was determined. In this experiment, the threshold cycle (Ct) and efficiency of PCR amplification are compared between two amplicons: one located in the genome as a control and one positioned in the plasmid. An amplicon in the gfp_m^{2+} gene could not be reproducibly amplified, very likely due to the high and repetitive G+C content in this gene. However, an amplicon in the A+T rich *E. coli* replication origin, present in all plasmids used in this study, could be amplified with high efficiency and reproducibility. Because multiple replication forks may initiate prior to cell division during log phase growth (Rocha, 2004), it was important to select a genomic amplicon located near the L5 *attB* site. When the PCR amplification efficiency and Ct values for the amplicon in the integrated plasmid of ML263 were compared to that of the control genomic amplicon, the copy number was found to be 1.0 ± 0.04 , validating the protocol for use in mycobacteria. When the same two amplicons were compared in cell lysates of *M. smegmatis* containing pMN437, the copy number of this pAL5000 plasmid was 23.5 ± 1.5 . Results from both strains are representative of at least three independent experiments. Thus, the large reduction in expression levels observed for genes on integrated plasmids is due to a large drop in copy number rather than adverse effects on gene expression at the L5 site in mycobacteria.

3.6. Adaptation of pML1335 to alternative integration sites

Due to its large size, the bacterial chromosome needs to be tightly compacted in order to fit within the constraints of the cell. The nonuniform bacterial nucleoid contains different topological domains of varying compaction which may influence transcriptional activity (Luijsterburg *et al.*, 2006). It is conceivable that the L5 *attB* site resides within a tightly compacted region in mycobacterial chromosomes and accessibility to RNA polymerases is restricted, thus limiting gene expression. Alternative integration sites in both *M. smegmatis* and *M. tuberculosis* exist in the *proT* and *alaV* tRNA genes for the mycobacteriophages Giles and Ms6 (Fig. 7). The modular design of pML1335 enabled us to easily replace the L5 *int* gene and *attP* region, generating the integrative plasmids pML1357 and pML1361 (Table 1) specific for the Giles and Ms6 *attB* sites, respectively.

To compare gene expression from different sites in mycobacterial chromosomes, pML1357 and pML1361 were transformed in *M. smegmatis* mc²155, generating the strains *M. smegmatis* ML444 (Giles) and ML446 (Ms6), respectively (Table 1). These two plasmids transformed into *M. smegmatis* mc²155 with nearly ten-fold higher efficiency than the L5 *attB*-specific pML1335 (Table 2). This could be due to higher expression levels or greater efficiencies of their respective integrases. These strains were identifiable in large sized

colonies by their XylE activity (not shown). However, the XylE activity of ML444 (Giles) and ML446 (Ms6) was much weaker than that of ML261 (L5), which readily reacted with catechol and produced a very prominent yellow color in culture smears (Fig. S2). These results indicated that genes in the unprotected backbone are not equally expressed at all integration sites in the *M. smegmatis* chromosome. In contrast, expression of *gfp_m²⁺* was identical at the L5, Giles and Ms6 *attB* sites (Fig. 4). This suggests that transcriptional interference may be present at the Giles and Ms6 attachment sites, which is blocked by bi-directional terminators flanking the *gfp_m²⁺* expression cassette. At all three integration sites, the presence or absence of the plasmid backbone did not affect expression of *gfp_m²⁺* from the protected cassette (Fig. 4). These results show that similar expression can be achieved *in vitro* at the L5, Giles, and Ms6 integration sites by using expression cassettes protected by intrinsic transcriptional terminators.

To determine whether the integration vectors pML1357 (Giles) and pML1361 (Ms6) function similarly in slowly growing mycobacteria, they were transformed into *M. tuberculosis* mc²6230, generating the strains ML448 and ML450, respectively (Table 1). Similarly to *M. smegmatis*, transformation was more efficient for these two plasmids compared to the L5 *attB*-specific pML1335 (Table 2). In contrast to *M. smegmatis*, the backbone reporter *xylE_m* provided similar catechol reactivity when expressed from all three integration sites (not shown). Figure 5 shows that expression of the protected *gfp_m²⁺* cassette from the L5, Giles, and Ms6 integration sites of the *M. tuberculosis* chromosome was also identical at all three sites and is independent of the presence or absence of the plasmid backbones. These results indicate that the L5 *attB* site is as suitable as the other chromosomal sites tested in *M. smegmatis* and *M. tuberculosis* for the expression of a gene of interest. This study demonstrates the versatility of our basic integration vector pML1342 in being adaptable to multiple phage integration sites as well as efficient gene expression from the protected cassette at all chromosomal sites in both fast and slow growing mycobacteria. To our knowledge, this is the first study comparing gene expression levels at different chromosomal sites in mycobacteria.

3.7. Simultaneous integration of expression cassettes at different chromosomal sites

For some studies, high expression levels are required to generate a detectable phenotype. This may be difficult to achieve as episomal plasmids cannot be selected for during infection of animals. An alternative approach to increasing expression from chromosomal constructs is to integrate multiple copies of the gene of interest. After integration, the L5 *attB* sequence is reconstituted at the resulting *attL* site (Lee *et al.*, 1991) (Fig. 1), but is not a functional reconstitution, perhaps due to steric blockade by tight integrase binding (Dr. Graham Hatfull, personal communication). To circumvent this, Saviola and Bishai integrated a vector carrying a new L5 *attB* site into the chromosome of *M. smegmatis*. The plasmid-born *attB* site could then serve as a second site for subsequent integration (Saviola and Bishai, 2004). Therefore, we positioned an L5 *attB* site within the protected polylinker of pML1342 to enable serial integrations. However, a small fraction of the *attP/attB*-carrying pML1330 recombined during passage through *E. coli* (Fig. S1), which made cloning in this vector difficult. Recombination may not have been previously observed as the *attB* sequence was mutated rather than tolerated (Saviola and Bishai, 2004), in contrast to our observations. Since the entire *glyV* gene was used in their construct, the authors proposed that these mutations could result from toxicity of GlyV to *E. coli*. Likewise, an unmutated *attB* site may support recombination with *attP* in *E. coli*. Interestingly, concentrated *E. coli* IHF promoted some L5 *attB/attP* recombination *in vitro*, despite the inability of *E. coli* crude extracts or purified but not concentrated *E. coli* IHF to do the same (Lee and Hatfull, 1993). An alternative to reiterative, adjacent positioning of identical genes at the same attachment site is to integrate plasmids at different locations in the chromosome (Freitas-Vieira *et al.*,

1998). To this end, the hygromycin resistance cassette of the Ms6 integration vector pML1361 was replaced by a kanamycin resistance cassette, generating the plasmid pML2300. This plasmid along with the Giles integration vector pML1357 carrying the *hyg* gene were co-transformed in *M. smegmatis* ML261 and *M. tuberculosis* ML442 (both strains containing L5 *attB*::*gfp_m²⁺*, unmarked), generating the *gfp_m²⁺* triple integration strains *M. smegmatis* ML453 and *M. tuberculosis* ML458, respectively (Table 1). The transformation efficiency (Table 2) of both the Giles (pML1357) and the Ms6 (pML2300) plasmids together in *M. smegmatis* ML261 was nearly ten-fold lower than pML1357 alone (generating the L5/Giles double *gfp_m²⁺* integrated ML451) and pML2300 alone (generating the L5/Ms6 double *gfp_m²⁺* integrated ML452). This was also the case when the same two plasmids were co-transformed into *M. tuberculosis* ML442 in comparison to pML1357 alone (generating the L5/Giles double *gfp_m²⁺* integrated ML454) and pML2300 alone (generating the L5/Ms6 double *gfp_m²⁺* integrated ML455).

The triple integrated strains *M. smegmatis* ML453 and *M. tuberculosis* ML458 were identified by colony PCR (Fig. 8). Figures 4 and 5 clearly show that for both *M. smegmatis* and *M. tuberculosis*, the triple integrated strains provide higher fluorescence than the double integrated strains, which in turn provide higher fluorescence levels than the single integrated strains. It was observed that fluorescence intensity did not increase linearly with the chromosomal copy number. This may be due to self aggregation of Gfp (Cramer *et al.*, 1996) or by inner filter effects (Eftink, 1997). In conclusion, these results show that an increase in gene expression can be achieved by superintegration at multiple sites in the chromosome.

3.8. Gfp protein expression

To examine whether protein levels corroborated with plasmid copy number, we sought to quantify Gfp expression in a Western blot. Cultures of equal optical density were lysed by sonication, centrifuged, and the supernatants were separated on a 10% polyacrylamide gel. After membrane transfer, Gfp was detected with a monoclonal antibody. As the pMV306kan-derived vector pML1337 (in ML263) served as the standard for commonly used integration vectors, the amount of protein quantified by pixel densitometry for this strain was set to 1, to which all other strains were normalized. Indeed, the amount of Gfp protein in cell lysates of ML453 was 3.6 fold more than ML263 (Fig. 9), in agreement with three integrated plasmids containing the *gfp_m²⁺* gene in this strain. Cell lysates of *M. smegmatis* carrying pMN437 contained approximately 26-fold more Gfp than ML263. These results are consistent with the copy number of pAL5000 plasmids determined by qPCR.

4. CONCLUSIONS

This series of integration vectors was designed to stably insert an expression cassette protected from transcriptional interference by bi-directional terminators into different sites of mycobacterial chromosomes. Additionally, the modular design of these plasmids makes them easily adaptable for a new generation of episomal expression vectors for mycobacteria. These vectors may serve as valuable genetic tools where stable expression or multiple integrations are required, for example for complementation in infection experiments or for development of multi-component vaccine strains.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Charles Turnbough for excellent advice on the design of transcriptional terminators. The strain *M. tuberculosis* mc²6230 was a gift from Dr. William Jacobs. The plasmid pML603 was constructed by Dr. Olga Danilchanka. We thank Dr. Adrie Steyn for use of the MyiQ2 real time PCR detection system and for providing the pCreSacB1 plasmid. Valuable advice regarding qPCR was provided by Dr. Ashwani Kumar, Dr. Aisha Farhana, and Loni Guidry. We also thank Ying Wang for outstanding technical assistance. This work was supported by grants from the USDA (WIS-01081) and the NIH (GM38660) to RL and by NIH grants AI063432 and AI083632 to MN.

Abbreviations

bp	base pair(s)
kbp	1,000 bp
hyg	gene encoding hygromycin phosphotransferase
aph	gene encoding aminoglycoside phosphotransferase
ColE1	<i>E. coli</i> origin of DNA replication
int	phage integrase gene
xylE_m	codon-adapted <i>xylE</i> gene for expression in <i>M. tuberculosis</i>
<i>gfp_m</i>²⁺	mutated <i>gfp</i> gene for increased fluorescence and stability and codon-adapted for expression in <i>M. tuberculosis</i>
attB	bacterial attachment site
attP	phage attachment site
tt_{sbi}A/B	synthetic bi-directional intrinsic terminator
loxP	Cre-recombinase recognition site
PCR	polymerase chain reaction
CCR	combined chain reaction
qPCR	quantitative real-time polymerase chain reaction
Ct	threshold cycle
wt	wild-type
kan	kanamycin
hyg	hygromycin

REFERENCES

- Adhya S, Gottesman M. Promoter occlusion: transcription through a promoter may inhibit its activity. *Cell*. 1982; 29:939–44. [PubMed: 6217898]
- Artsimovitch I, Svetlov V, Anthony L, Burgess RR, Landick R. RNA polymerases from *Bacillus subtilis* and *Escherichia coli* differ in recognition of regulatory signals in vitro. *J Bacteriol*. 2000; 182:6027–35. [PubMed: 11029421]
- Ausubel, FM.; Brent, R.; Kingston, RE.; Moore, DD.; Seidmann, JG.; Smith, JA.; Struhl, K. *Current Protocols in Molecular Biology*. John Wiley & Sons; New York: 1987.
- Crameri A, Whitehorn EA, Tate E, Stemmer WP. Improved green fluorescent protein by molecular evolution using DNA shuffling. *Nat. Biotechnol*. 1996; 14:315–319. [PubMed: 9630892]
- Curcio R, Dhandayuthapani S, Deretic V. Gene expression in mycobacteria: transcriptional fusions based on *xylE* and analysis of the promoter region of the response regulator *mtrA* from *Mycobacterium tuberculosis*. *Mol Microbiol*. 1994; 13:1057–64. [PubMed: 7854120]

- de Hoon MJ, Makita Y, Nakai K, Miyano S. Prediction of transcriptional terminators in *Bacillus subtilis* and related species. *PLoS Comput Biol*. 2005; 1:e25. [PubMed: 16110342]
- Eftink MR. Fluorescence methods for studying equilibrium macromolecule-ligand interactions. *Methods Enzymol*. 1997; 278:221–57. [PubMed: 9170316]
- Freitas-Vieira A, Anes E, Moniz-Pereira J. The site-specific recombination locus of mycobacteriophage Ms6 determines DNA integration at the tRNA(Ala) gene of *Mycobacterium* spp. *Microbiology*. 1998; 144(Pt 12):3397–406. [PubMed: 9884232]
- Fu J, Li D, Xia S, Song H, Dong Z, Chen F, Sun X, Tang Z. Absolute quantification of plasmid DNA by real-time PCR with genomic DNA as external standard and its application to a biodistribution study of an HIV DNA vaccine. *Anal Sci*. 2009; 25:675–80. [PubMed: 19430152]
- Gaal T, Ross W, Estrem ST, Nguyen LH, Burgess RR, Gourse RL. Promoter recognition and discrimination by EsigmaS RNA polymerase. *Mol Microbiol*. 2001; 42:939–54. [PubMed: 11737638]
- Gavigan JA, Ainsa JA, Perez E, Otal I, Martin C. Isolation by genetic labeling of a new mycobacterial plasmid, pJAZ38, from *Mycobacterium fortuitum*. *J Bacteriol*. 1997; 179:4115–22. [PubMed: 9209023]
- Kaps I, Ehrt S, Seeber S, Schnappinger D, Martin C, Riley LW, Niederweis M. Energy transfer between fluorescent proteins using a co-expression system in *Mycobacterium smegmatis*. *Gene*. 2001; 278:115–24. [PubMed: 11707328]
- Lee MH, Hatfull GF. Mycobacteriophage L5 integrase-mediated site-specific integration in vitro. *J. Bacteriol*. 1993; 175:6836–41. [PubMed: 8226625]
- Lee MH, Pascopella L, Jacobs WR Jr, Hatfull GF. Site-specific integration of mycobacteriophage L5: integration-proficient vectors for *Mycobacterium smegmatis*, *Mycobacterium tuberculosis*, and bacille Calmette-Guerin. *Proc Natl Acad Sci USA*. 1991; 88:3111–5. [PubMed: 1901654]
- Luijsterburg MS, Noom MC, Wuite GJ, Dame RT. The architectural role of nucleoid-associated proteins in the organization of bacterial chromatin: a molecular perspective. *J Struct Biol*. 2006; 156:262–72. [PubMed: 16879983]
- Morris P, Marinelli LJ, Jacobs-Sera D, Hendrix RW, Hatfull GF. Genomic characterization of mycobacteriophage giles: evidence for phage acquisition of host DNA by illegitimate recombination. *J Bacteriol*. 2008; 190:2172–82. [PubMed: 18178732]
- Pedulla ML, Lee MH, Lever DC, Hatfull GF. A novel host factor for integration of mycobacteriophage L5. *Proc Natl Acad Sci USA*. 1996; 93:15411–6. [PubMed: 8986825]
- Pena CE, Kahlenberg JM, Hatfull GF. Protein-DNA complexes in mycobacteriophage L5 integrative recombination. *J Bacteriol*. 1999; 181:454–61. [PubMed: 9882658]
- Pena CE, Lee MH, Pedulla ML, Hatfull GF. Characterization of the mycobacteriophage L5 attachment site, attP. *J Mol Biol*. 1997; 266:76–92. [PubMed: 9054972]
- Pena CE, Stoner JE, Hatfull GF. Positions of strand exchange in mycobacteriophage L5 integration and characterization of the attB site. *J Bacteriol*. 1996; 178:5533–6. [PubMed: 8808947]
- Pham TT, Jacobs-Sera D, Pedulla ML, Hendrix RW, Hatfull GF. Comparative genomic analysis of mycobacteriophage Tweety: evolutionary insights and construction of compatible site-specific integration vectors for mycobacteria. *Microbiology*. 2007; 153:2711–23. [PubMed: 17660435]
- Ranes MG, Rauzier J, Lagranderie M, Gheorghiu M, Gicquel B. Functional analysis of pAL5000, a plasmid from *Mycobacterium fortuitum*: construction of a “mini” *mycobacterium-Escherichia coli* shuttle vector. *J Bacteriol*. 1990; 172:2793–7. [PubMed: 2158981]
- Ribeiro G, Viveiros M, David HL, Costa JV. Mycobacteriophage D29 contains an integration system similar to that of the temperate mycobacteriophage L5. *Microbiology*. 1997; 143(Pt 8):2701–8. [PubMed: 9274023]
- Rocha EP. The replication-related organization of bacterial genomes. *Microbiology*. 2004; 150:1609–27. [PubMed: 15184548]
- Sambandamurthy VK, Derrick SC, Hsu T, Chen B, Larsen MH, Jalapathy KV, Chen M, Kim J, Porcelli SA, Chan J, Morris SL, Jacobs WR Jr. *Mycobacterium tuberculosis* delta*RD1* delta*panCD*: a safe and limited replicating mutant strain that protects immunocompetent and immunocompromised mice against experimental tuberculosis. *Vaccine*. 2006; 24:6309–20. [PubMed: 16860907]

- Saviola B, Bishai WR. Method to integrate multiple plasmids into the mycobacterial chromosome. *Nucleic Acids Res.* 2004; 32:e11. [PubMed: 14718555]
- Skulj M, Okrsjar V, Jalen S, Jevsevar S, Slanc P, Strukelj B, Menart V. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. *Microb Cell Fact.* 2008; 7:6. [PubMed: 18328094]
- Snapper SB, Melton RE, Mustafa S, Kieser T, Jacobs WR Jr. Isolation and characterization of efficient plasmid transformation mutants of *Mycobacterium smegmatis*. *Mol Microbiol.* 1990; 4:1911–9. [PubMed: 2082148]
- Song H, Niederweis M. Functional expression of the FLP recombinase in *Mycobacterium bovis* BCG. *Gene.* 2007; 399:112–9. [PubMed: 17590537]
- Springer B, Sander P, Sedlacek L, Ellrott K, Böttger EC. Instability and site-specific excision of integration-proficient mycobacteriophage L5 plasmids: development of stably maintained integrative vectors. *Int J Med Microbiol.* 2001; 290:669–75. [PubMed: 11310445]
- Steinhauer K, Eschenbacher I, Radischat N, Detsch C, Niederweis M, Goroncy-Bermes P. Rapid evaluation of the mycobactericidal efficacy of disinfectants in the quantitative carrier test EN 14563 by using fluorescent *Mycobacterium terrae*. *Appl Environ Microbiol.* 2010; 76:546–54. [PubMed: 19948860]
- Stephan J, Stemmer V, Niederweis M. Consecutive gene deletions in *Mycobacterium smegmatis* using the yeast FLP recombinase. *Gene.* 2004; 343:181–90. [PubMed: 15563844]
- Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH, Hatfull GF, et al. New use of BCG for recombinant vaccines. *Nature.* 1991; 351:456–60. [PubMed: 1904554]
- Timm J, Lim EM, Gicquel B. *Escherichia coli*-mycobacteria shuttle vectors for operon and gene fusions to *lacZ*: the pJEM series. *J. Bacteriol.* 1994; 176:6749–53. [PubMed: 7961429]
- Unniraman S, Prakash R, Nagaraja V. Alternate paradigm for intrinsic transcription termination in eubacteria. *J Biol Chem.* 2001; 276:41850–5. [PubMed: 11551936]
- Wolschendorf F, Mahfoud M, Niederweis M. Porins are required for uptake of phosphates by *Mycobacterium smegmatis*. *J Bacteriol.* 2007; 189:2435–2442. [PubMed: 17209034]
- Wu AM, Platt T. Transcription termination: nucleotide sequence at 3' end of tryptophan operon in *Escherichia coli*. *Proc Natl Acad Sci U S A.* 1978; 75:5442–6. [PubMed: 364481]

Fig. 1. Genomic organization of the mycobacteriophage L5 bacterial attachment site and crossover with the phage attachment site

A) A 5.6 kbp fragment of the *M. tuberculosis* genome encompassing the mycobacteriophage L5 attachment site *attB*. The *attB* site overlaps the *glyV* tRNA gene and is composed of a nearly identical to the phage attachment site *attP* as well as an essential but non-homologous 10 bp upstream sequence. Numbers flanking the genomic map correspond to coordinates in the *M. tuberculosis* H37Rv genome (<http://genolist.pasteur.fr/TubercuList/>). Annotations of surrounding genes are listed above. B) Position of overlap between *attB* and *glyV*. The 72 bp fragment encompassing *glyV* and *attB* shows that 52 of the 53 bp in *attB* overlap with *glyV* between coordinates 2765560-2765611. The non-homologous region of *attB* refers to an essential sequence of *attB* that is not present in the core region of mycobacteriophage L5 *attP*. C) Sequence and position of strand exchange. The DNA sequence of the *attB* site is overlaid with the *attP* sequence, which is cloned into a generic integration vector. Positions of strand exchange between the bacterial chromosome and the integrating plasmid (Pena *et al.*, 1996) are noted by the jagged lines in the sequences. The base pairs in bold and underlined are the only mismatching base pairs between *attB* of *M. tuberculosis* and *attP*. Reconstituted *glyV* after integration in *M. tuberculosis* therefore contains a single point mutation. *int*, L5 integrase gene; *aph*, kanamycin resistance gene; *colEI*, *E. coli* origin of replication. D) Chromosomal arrangement of genes resulting from a generic plasmid integration. Site-specific recombination of *attB* (hatched bar, non-homologous region; black bar, homologous region) and *attP* (grey bar) creates the flanking *attL* and *attR* sites. The color scheme shows the hybrid portions of *attL* and *attR* originating from *attB* and *attP* (note: figure is not to scale).

Fig. 2. Plasmid map of integration vector pML1342

The integration vector pML1342 carries the mycobacteriophage L5 integrase gene and the L5 phage attachment site *attP*. The synthesized *xylE_m* gene was adapted to codons preferentially used by *M. tuberculosis*. *LoxP* sites flank an excisable backbone, leaving behind a stably integrated polylinker flanked by two bi-directional intrinsic terminators named *tt_{sb}A* and *tt_{sb}B*. The protected polylinker includes the restriction sites *PacI*, *SmaI* and *PmeI* which occur rarely or not at all in the *M. tuberculosis* genome. Five additional restriction sites are also included in the polylinker for cloning convenience. The assumed secondary structures of the bi-directional terminators are shown to the sides. Each component of the plasmid is flanked by single-cutting restriction endonuclease sites. Flanking P sites in the *attP* region are required for intasome formation, a complex comprising L5 *attP*, L5 Int, and mIHF (Pena *et al.*, 1999).

Fig. 3. Identification of positive clones by visualization of plasmid-encoded reporter genes

Confocal micrographs of single colonies were taken at $20 \times$ magnification. For comparison, micrographs of single colonies were taken under white light. Gfp fluorescence was visualized using an excitation wavelength of 490 nm and an emission filter at 512 nm. Yellow colony color indicative of Xyle activity was visualized under white light after 10 min exposure to 1% catechol. The *M. smegmatis* strain for each colony is notated in bold and underlined. Presence of the genes gfp_m^{2+} and/or $xyleM$ in each strain is listed beneath the strain name. The name and description of the plasmid carried by each strain is also included. enhanced int.: enhanced integration plasmid with gfp_m^{2+} (pML1335) or without gfp_m^{2+} (pML1342). pMV306-reference: standard integration vector pMV306kan with gfp_m^{2+} named pML1337.

Fig. 4. Expression of gfp_m^{2+} from episomal and integrative vectors in *M. smegmatis*

The fluorescence of cultures grown to an $OD_{600} \sim 0.8$ was measured and normalized to optical density. Results are shown as averages of at least three independent experiments with their standard deviations. Strain names are notated below the graph and descriptions of the plasmids used to generate each strain are notated above each bar. The plasmid names associated with each strain and description are: episomal: pMN437; new, L5: pML1335; standard, L5: pML1337; new, empty: pML1342; new, Giles: pML1357; new, Ms6: pML1361; L5, Ms6: pML2300 in ML262; L5, Giles: pML1357 in ML262; L5, Ms6, Giles: pML2300 co-transformed with pML1357 in ML262.

Fig. 5. Expression of gfp_m^{2+} from episomal and integrative vectors in *M. tuberculosis*
 The fluorescence of avirulent *M. tuberculosis* mc²6230 cultures grown to an OD₆₀₀ ~ 0.8 was measured and normalized to optical density. Results are shown as averages of at least three independent experiments with their standard deviations. Strain names are notated below the graph and descriptions of the plasmids used to generate each strain are notated above each bar. The plasmid names associated with each strain and description are: episomal: pMN437; new, L5: pML1335; standard, L5: pML1337; new, empty: pML1342; new, Giles: pML1357; new, Ms6: pML1361; L5, Ms6: pML2300 in ML442; L5, Giles: pML1357 in ML442; L5, Ms6, Giles: pML2300 co-transformed with pML1357 in ML442.

Fig. 6. Transcription termination efficiency of bi-directional terminators

A) Diagram of linear templates. Polymerases extending from an upstream promoter may be terminated at the test terminator (either *tt_{sbiA}*, *tt_{sbiB}*, or the control *trpA*), read-through and be terminated at *rrnBT2*, or read-through both terminators. B) Termination efficiencies of test terminators using the *E. coli* RNAP. C). Termination efficiencies of test terminators using the recombinant *M. bovis* RNAP. Numbers below each lane indicate the termination efficiency (%T) and the standard deviation (SD). To the sides of the gel pictures, T represents RNA products arising from termination at the test terminator and RT represents RNA products arising from termination at the read-through terminator *rrnBT2* and read-through to the end of the template.

Fig. 7. Mycobacteriophage attachment sites in *M. smegmatis* and *M. tuberculosis* chromosomes
 The mycobacteriophage *attB* sites present in both species of mycobacteria are noted along with the tRNA genes in which they overlap. Coordinates are listed for each *attB* site. The origin is defined as the start of the *dnaA* gene and determines the first coordinate base pair.

Fig. 8. Co-transformation and superintegration in *M. tuberculosis*

M. tuberculosis ML442 (L5 *attB*::*gfp_m*²⁺; backbone removed) was co-transformed with pML2300 and pML1357, generating the strain ML458. A) Genomic map of *M. tuberculosis* ML458 with *gfp_m*²⁺ integrated at three positions in the chromosome. Integration of pML2300 and pML1357 in ML442 at the Ms6 and Giles *attB* sites, respectively, triples the number of *gfp_m*²⁺ genes in the newly constructed strain ML448. Approximate positions of primers used for confirmation of ML458 are drawn. After PCR amplification using the following primer pairs, the expected bands on a DNA gel are: P1-P2: 1,991 bp; P3-P4: 497 bp; P5-P6: 868 bp; P7-P8: 622 bp; P9-P10: 883 bp. Refer to Table S1 for primer sequences. Figure is not to scale. B) Gel of PCR-amplified DNAs from ML458 using primer combinations listed in A.

Fig. 9. Copy number-dependent expression of Gfp

M. smegmatis strains were grown to an OD₆₀₀ of 1.0, lysed, centrifuged, and supernatants separated on a gel and transferred to a nitrocellulose membrane for detection with an anti-Gfp mAb in a Western blot. Strain names are noted above the lanes; pMN437 refers to *M. smegmatis* mc²155 carrying the pMN437 plasmid. Gfp bands were quantified by pixel densitometry and normalized to the control ML263, which was set to 1 to represent its single copy number. Normalized to ML263, the following fold-amount of Gfp was detected in each strain: *M. smegmatis* mc²155/pMN437: 26.0; ML261: 1.3; ML264: 0.0; ML453: 3.6.

Table 1
Strains and plasmids used in this study

Ept: improved electroporation due to lipid modification, *pAL5000*: mycobacterial origin of replication, *colE1*: *E. coli* origin of replication, *hyg*: hygromycin resistance gene, *p_{smyc}*: strong mycobacterial promoter, *gfp_m²⁺*: *gfp* gene mutated for enhanced fluorescence and stability and codon optimized green fluorescence protein gene, *aph*: kanamycin resistance gene, *sacB*: sucrose resistance operon, *int*: integrase gene, *bla*: ampicillin resistance gene, *FRT*: Flp recombinase recognition site, *attP*: phage attachment site, *xylE_m*: codon optimized catechol 2,3 dioxygenase gene, *loxP*: Cre recombinase recognition site, *tt_{sb}A/B*: synthetic bi-directional terminators, *p_{wmyc}*: weak mycobacterial promoter, *p_{groEL}*: groEL promoter.

Strains	Genotype	Relevant Description	Reference
<i>M. smegmatis</i> mc ² 155	Ept	competent wild-type <i>M. smegmatis</i>	(Snapper <i>et al.</i> , 1990)
<i>M. smegmatis</i> ML261	mc ² 155 derivative; L5 <i>attB</i> ::pML1335	new integration vector; <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML262	ML261 derivative; L5 <i>attB</i> :: <i>gfp_m²⁺</i>	pML1335 backbone removed; <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML263	mc ² 155 derivative; L5 <i>attB</i> ::pML1337	standard integration vector; <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML264	mc ² 155 derivative; L5 <i>attB</i> ::pML1342	new integration vector; empty	this study
<i>M. smegmatis</i> ML444	mc ² 155 derivative; Giles <i>attB</i> ::pML1357	new integration vector; Giles <i>int/attP</i> , <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML446	mc ² 155 derivative; Ms6 <i>attB</i> ::pML1361	new integration vector, Ms6 <i>int/attP</i> , <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML451	ML262 derivative; Ms6 <i>attB</i> ::pML2300	<i>gfp_m²⁺</i> integrated at both L5 and Ms6 <i>attBs</i>	this study
<i>M. smegmatis</i> ML452	ML262 derivative; Giles <i>attB</i> ::pML1357	<i>gfp_m²⁺</i> integrated at both L5 and Giles <i>attBs</i>	this study
<i>M. smegmatis</i> ML453	ML262 derivative; Giles <i>attB</i> ::pML1357, Ms6 <i>attB</i> ::pML2300	<i>gfp_m²⁺</i> integrated at L5, Ms6 and Giles <i>attBs</i>	this study
<i>M. smegmatis</i> ML454	ML444 derivative; Giles <i>attB</i> :: <i>gfp_m²⁺</i>	pML1357 backbone removed; <i>gfp_m²⁺</i>	this study
<i>M. smegmatis</i> ML455	ML446 derivative; Ms6 <i>attB</i> :: <i>gfp_m²⁺</i>	pML1361 backbone removed; <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> mc ² 6230	<i>M. tuberculosis</i> H37Rv derivative; $\Delta RD1$, $\Delta panCD$	avirulent <i>M. tuberculosis</i> ; “wild-type”	(Sambandamurthy <i>et al.</i> , 2006)
<i>M. tuberculosis</i> ML440	mc ² 6230 derivative; L5 <i>attB</i> ::pML1335	new integration vector; <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> ML442	ML440 derivative; L5 <i>attB</i> :: <i>gfp_m²⁺</i>	pML1335 backbone removed; <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> ML441	mc ² 6230 derivative; L5 <i>attB</i> ::pML1337	standard integration vector; <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> ML268	mc ² 6230 derivative; L5 <i>attB</i> ::pML1342	new integration vector; empty	this study
<i>M. tuberculosis</i> ML448	mc ² 6230 derivative; Giles <i>attB</i> ::pML1357	new integration vector; Giles <i>int/attP</i> , <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> ML450	mc ² 6230 derivative; Ms6 <i>attB</i> ::pML1361	new integration vector, Ms6 <i>int/attP</i> , <i>gfp_m²⁺</i>	this study
<i>M. tuberculosis</i> ML456	ML442 derivative; Ms6 <i>attB</i> ::pML2300	<i>gfp_m²⁺</i> integrated at both L5 and Ms6 <i>attBs</i>	this study
<i>M. tuberculosis</i> ML457	ML442 derivative; Giles <i>attB</i> ::pML1357	<i>gfp_m²⁺</i> integrated at both L5 and Giles <i>attBs</i>	this study
<i>M. tuberculosis</i> ML458	ML442 derivative; Giles <i>attB</i> ::pML1357, Ms6 <i>attB</i> ::pML2300	<i>gfp_m²⁺</i> integrated at L5, Ms6 and Giles <i>attBs</i>	this study
<i>M. tuberculosis</i> ML459	ML448 derivative; Giles <i>attB</i> :: <i>gfp_m²⁺</i>	pML1357 backbone removed; <i>gfp_m²⁺</i>	this study

Strains	Genotype	Relevant Description	Reference
<i>M. tuberculosis</i> ML460	ML450 derivative; <i>Ms6 attB::gfp_m²⁺</i>	pML1361 backbone removed; <i>gfp_m²⁺</i>	this study
Plasmids			
pMN437	<i>pAL5000, colE1, hyg, p_{smyc}-gfp_m²⁺</i>	episomal <i>gfp_m²⁺</i> plasmid; 6,236 bp	(Steinhauer <i>et al.</i> , in press)
pML102	<i>pAL5000, colE1, aph, sacB, L5 int</i>	L5- <i>int</i> containing plasmid; 8,647 bp	(Wolschendorf <i>et al.</i> , 2007)
pML113	<i>colE1, bla, FRT-hyg-FRT, L5 attP</i>	<i>colE1</i> containing plasmid; 4,365 bp	(Wolschendorf <i>et al.</i> , 2007)
pML603	<i>colE1, hyg</i>	<i>colE1</i> and <i>hyg</i> containing backbone; 3,495 bp	this study
pML1321	<i>colE1, hyg, L5 int</i>	pML603 with L5 <i>int</i> ; 4,867 bp	this study
pML1320	<i>colE1, p_{wmyc}-xylE_m loxP, L5 attP, tt_{shf}A, L5 attB, tt_{shf}B, loxP, aph</i>	synthetic <i>xylE_m</i> , protected <i>attB</i> /polylinker flanked by <i>loxP</i> ; 4,841 bp	this study
pML1330	<i>colE1, hyg, L5 int, xylE_m loxP, L5 attP, tt_{shf}A, L5 attB, tt_{shf}B, loxP</i>	pML603 backbone with synthetic construct; 5,459 bp	this study
pML1342	<i>colE1, L5 int, L5 attP, hyg, xylE_m loxP, L5 attP, tt_{shf}A, tt_{shf}B, loxP</i>	new L5 integration vector (no <i>attB</i>); 5,404 bp	this study
pML1335	<i>colE1, L5 int, L5 attP, xylE_m hyg, p_{smyc}-gfp_m²⁺</i>	pML1342 backbone, <i>p_{smyc}-gfp_m²⁺</i> ; 6,484 bp	this study
pMV306kan	<i>colE1, L5 int, L5 attP, aph</i>	standard L5 integration vector; 3,995 bp	(Stover <i>et al.</i> , 1991)
pML1337	<i>colE1, L5 int, L5 attP, aph, p_{smyc}-gfp_m²⁺</i>	pMV306kan backbone, <i>p_{smyc}-gfp_m²⁺</i> ; 5,101 bp	this study
pGH1000A	<i>colE1, Giles int, Giles attP, hyg</i>	standard Giles integration vector; 5,226 bp	(Morris <i>et al.</i> , 2008)
pML1357	<i>colE1, xylE_m gfp_m²⁺, hyg, Giles int, Giles attP</i>	pML 1335 backbone, Giles <i>int/attP</i> ; 6,454 bp	this study
pNIP40b	<i>colE1, Ms6 int, Ms6 attP, hyg</i>	standard Ms6 integration vector; 7,290 bp	Dr. N. Winter
pML1361	<i>colE1, xylE_m gfp_m²⁺, hyg, Ms6 int, Ms6 attP</i>	pML1335 backbone, Ms6 <i>int/attP</i> ; 6,594 bp	this study
pML2300	<i>colE1, xylE_m gfp_m²⁺, aph, Ms6 int, Ms6 attP</i>	kanamycin resistant derivative of pML1361; 6,821 bp	this study
pCreSacBI	<i>colE1, p_{groEL}-cre, p_{groEL}-sacRB, aph</i>	counter-selectable Cre-recombinase vector; 7,891 bp	Dr. A. Steyn

Table 2
Transformation efficiencies

500 ng of each plasmid were transformed into *M. smegmatis* or *M. tuberculosis* strains by electroporation and selected for on Middlebrook 7H10 agar with appropriate antibiotics. Colony-forming units (cfu) were counted and used to calculate transformation efficiencies in cfu per μg DNA. The standard pMV306kan-derived vectors are noted as “standard” while all pML1342-derived vectors are noted as “new.” The site of plasmid integration is noted as L5, Giles, or Ms6. The presence of the gfp_m^{2+} gene is also listed where appropriate, while “empty” refers to those plasmids without the gfp_m^{2+} gene. Plasmids encoding resistance genes to hygromycin and kanamycin are notated as “hyg” and “kan,” respectively.

Organism	Strain	Plasmid	Relevant Plasmid Description	Efficiency
<i>M. smegmatis</i>	mc ² 155	pML1342	new, L5, empty, hyg	8.1×10^3
<i>M. smegmatis</i>	mc ² 155	pML1335	new, L5, gfp_m^{2+} , hyg	4.9×10^3
<i>M. smegmatis</i>	mc ² 155	pMV306kan	standard, L5, empty, kan	2.1×10^4
<i>M. smegmatis</i>	mc ² 155	pML1337	standard, L5, gfp_m^{2+} , kan	3.2×10^4
<i>M. smegmatis</i>	mc ² 155	pML1357	new, Giles, gfp_m^{2+} , hyg	2.8×10^4
<i>M. smegmatis</i>	mc ² 155	pML1361	new, Ms6, gfp_m^{2+} , hyg	4.0×10^4
<i>M. smegmatis</i>	ML262	pML1357	new, Giles, gfp_m^{2+} , hyg	1.7×10^4
<i>M. smegmatis</i>	ML262	pML2300	new, Ms6, gfp_m^{2+} , kan	1.4×10^4
<i>M. smegmatis</i>	ML262	pML1357 + pML2300	new, Giles, gfp_m^{2+} , hyg; new, Ms6, gfp_m^{2+} , kan	2.2×10^3
<i>M. tuberculosis</i>	mc ² 6230	pML1342	new, L5, empty, hyg	2.4×10^3
<i>M. tuberculosis</i>	mc ² 6230	pML1335	new, L5, gfp_m^{2+} , hyg	2.7×10^3
<i>M. tuberculosis</i>	mc ² 6230	pMV306kan	standard, L5, empty, kan	9.6×10^4
<i>M. tuberculosis</i>	mc ² 6230	pML1337	standard, L5, gfp_m^{2+} , kan	9.2×10^4
<i>M. tuberculosis</i>	mc ² 6230	pML1357	new, Giles, gfp_m^{2+} , hyg	4.3×10^4
<i>M. tuberculosis</i>	mc ² 6230	pML1361	new, Ms6, gfp_m^{2+} , hyg	1.3×10^4
<i>M. tuberculosis</i>	ML442	pML1357	new, Giles, gfp_m^{2+} , hyg	1.5×10^4
<i>M. tuberculosis</i>	ML442	pML2300	new, Ms6, gfp_m^{2+} , kan	5.2×10^4
<i>M. tuberculosis</i>	ML442	pML1357 + pML2300	new, Giles, gfp_m^{2+} , hyg; new, Ms6, gfp_m^{2+} , kan	3.4×10^3