BioMed Central

Research article

I-TRAP: A method to identify transcriptional regulator activated promoters

Molly M McLendon[1,2,3] and Thomas M Shinnick*[2]

Address: [1]Department of Microbiology and Immunology, Emory University, Atlanta, Georgia 30322, USA, [2]Division of AIDS, STD, and TB Laboratory Research, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333, USA and [3]Present address: Department of Microbiology, University of Iowa, Iowa City, Iowa 52242, USA

Email: Molly M McLendon - molly-mclendon@uiowa.edu; Thomas M Shinnick* - tms1@cdc.gov

* Corresponding author

Abstract

Background: The differential expression of virulence genes is often used by microbial pathogens in adapting to the environment of their host. The differential expression of such sets of genes can be regulated by RNA polymerase sigma factors. Some sigma factors are differentially expressed, which can provide a means to identifying other differentially expressed genes such as those whose expression are controlled by the sigma factor.

Methods: To identify sigma factor-regulated genes, we developed a method, termed I-TRAP, for the identification of transcriptional regulator activated promoters. The I-TRAP method is based on the fact that some genes will be differentially expressed in the presence and absence of a transcriptional regulator. I-TRAP uses a DNA library in a promoter-trap vector that contains two reporter genes, one to allow the selection of active promoters in the presence of the transcriptional regulator and a second to allow screening for promoter activity in the absence of the transcriptional regulator.

Results: To illustrate the development and use of the I-TRAP approach, the construction of the vectors, host strains, and library necessary to identify SigmaE-regulated genes of *Mycobacterium tuberculosis* is described.

Conclusion: The I-TRAP method should be a versatile and useful method for identifying and characterizing promoter activity under a variety of conditions and in response to various regulatory proteins. In our study, we isolated 360 clones that may contain plasmids carrying SigmaE-regulated promoters genes of *M. tuberculosis*.

Background

The ability of *Mycobacterium tuberculosis* bacteria to infect and replicate inside host mononuclear phagocytes is key to its ability to establish an infection and cause tuberculosis. To elucidate the mechanisms used by tubercle bacilli to survive in macrophages, researchers at several laboratories are studying mycobacterial genes specifically expressed during growth in macrophages [1–6]. Among the genes specifically expressed during intracellular growth are genes encoding two extracytoplasmic function sigma factors, SigE (σ^E) and SigH (σ^H) [3–5]. Sigma factors play critical roles in the recognition of promoters by RNA polymerase. In several systems, individual sigma factors are known to regulate the differential expression of

Table 1: Bacterial strains and plasmids. Plasmid pAWS2 was provided by Rosalind Van Landingham, pBluescript-int was provided by Bonnie Plikaytis, and pUS267 was provided by Jeremy Dale

Strains	Relevant Characteristics	Source/Reference
Electromax DH5α	Cloning host	Invitrogen
M. smegmatis LR222	Cloning host	[20]
M. smegmatis::SEint.2	constitutively expresses *M. tuberculosis sigE* gene	This study
M. smegmatisΔsigE	*sigE* knockout mutant	This study
M. tuberculosis H37Rv TMC102	Virulent laboratory strain	[21]
Plasmids		
pAWS2	Source of the *lacZ* gene	R. Van Landingham
pBluescript-int	Source of integrase gene, *int*	B. Plikaytis
pCL5	Backbone for pCL8	[9]
pCL8	Source of the *rpsL* promoter	This study
pCRScript SK (+)	Vector with multiple cloning site	Stratagene
pMV261	Source of the *hsp60* promoter and OriM	[22]
pPICZalphaC	Source of bleomycin resistance gene, *ble*	Invitrogen
pPR27	Source of gentamicin resistance gene, *aacC1*	[23]
pSE.1	pUC19 with promoterless *aph* and *xylE* genes	This study
pSE.2	pSE.1 with the *hyg* gene	This study
pSE.3	pSE.2 with the OriM	This study
pSE.4	promoter-trap vector	This study
pSEint.1	Integrating plasmid containing the *ble* gene	This study
pSEint.2	Integrating plasmid, constitutive *M. tuberculosis sigE* expression	This study
pSEKO.1	5' region of the *M. smegmatis sigE* gene	This study
pSEKO.2	pSEKO.1 with 3' end of *M. smegmatis sigE* gene	This study
pSEKO.3	pSEKO.2 with the *aacC1* gene	This study
pSEKO.4	Suicide plasmid with interrupted *M. smegmatis sigE* gene	This study
pT7Blue	Blunt cloning vector	This study
pTKMX97	Source of the *xylE* gene	[10]
pUC19	Vector with multiple cloning site	Amersham Biosciences
pUC4-KIXX	Source of *aph* gene	Amersham Biosciences
pUS267	Source of hygromycin resistance gene, *hyg*	J. Dale

sets of genes involved in specific processes, such as genes involved in the heat-shock response or those required for sporulation in *Bacillus subtilis* [7,8].

Identifying a differentially expressed sigma factor can provide a means to identifying other differentially expressed genes – genes whose expression are controlled by the sigma factor. To identify sigma factor-regulated genes, we developed a method, termed I-TRAP, for the identification of transcriptional regulator activated promoters. I-TRAP takes advantage of the fact that some genes will be differentially expressed in the presence and absence of a transcription regulatory protein. To identify transcription factor-dependent promoters, a genomic DNA library is generated in a promoter-trap plasmid containing two reporter genes. The library is transformed into strain expressing the transcription factor, and recipients with plasmids containing an active promoter are selected by virtue of expression of one of the reporter genes. The plasmids are recovered and transformed into a strain that does not express the transcription factor. Promoters dependent on the expression of the regulator are differentiated from

other promoters by screening for lack of reporter activity in bacteria not expressing the regulatory protein. To illustrate the development and use of the I-TRAP approach, we describe here the construction of the vectors, host strains, and library necessary to identify σ^E-regulated genes of *M. tuberculosis*.

Methods
Bacterial strains and growth condition
Bacterial strains are listed in Table 1. Escherichia coli strains were grown in Luria Broth (LB) or on LB agar. Mycobacterium smegmatis LR222 strains were grown in Middlebrook 7H9 broth supplemented with 10% albumin-dextrose-catalase (ADC, v/v, Difco Laboratories, Detroit, MI) and 0.05% Tween 80 (v/v, Sigma, St. Louis, MO) at 37°C, with agitation on a rotating platform (100 rpm), or were grown on trypticase soy agar (TSA) plates. M. tuberculosis H37Rv TMC102 cultures were grown in Middlebrook 7H9 broth supplemented with ADC and 0.05% Tween 80 at 37°C in 250 ml nephelometer flasks on a rotating platform (60 rpm) or were grown on Middlebrook 7H10 plates with 10% oleic acid-albumin-dex-

trose-catalase (OADC, Difco). For selection of E. coli bacteria, drugs were incorporated into the medium at the following concentrations: hygromycin (Invitrogen, Carlsbad, CA), 200 µg/m; kanamycin (Sigma), 25 µg/ml; zeocin (Invitrogen), 25 µg/ml; and gentamicin (Invitrogen), 20 µg/ml. For selection of M. smegmatis bacteria, drugs were incorporated into medium at the following concentrations: hygromycin (Invitrogen), 50 µg/ml; kanamycin (Sigma), 50 µg/ml; zeocin (Invitrogen), 50 µg/ml; and gentamicin (Invitrogen), 5 µg/ml.

Plasmids
Plasmids used in this study, along with their key features and sources, are listed in Table 1. Two of the source plasmids, pCL8 and pAWS2, were constructed as part of unpublished studies. pCL8 is a derivative of pCL5 [9] and contains the luciferase (*luc*) gene from pT3/T7LUC (Clontech, Palo Alto, CA), the mycobacteriophage L5 *int* gene, the kanamycin-resistance gene (*aph*) from pUC4-KIXX (Amersham Biosciences, Piscataway, NJ), and the P$_{rpsL}$-*xylE* cassette from pTKmx97 in which the *Bam*HI site between the *M. smegmatis rpsL* promoter and the *xylE* gene in pTKmx97 [10] was changed into a *Kpn*I site by site-directed PCR mutagenesis. pAWS.2 is a promoter-trap plasmid containing a promoterless *lacZ* gene isolated from λgt11 in which the *Eco*RI site in the 3' end of the *lacZ* gene was removed by site-directed polymerase-chain-reaction (PCR) mutagenesis (personal communication R. Van Landingham).

Construction of pSEint.2
1) The *ble* gene was PCR amplified from pPICZalphaC as a 375-bp fragment using primers Ble-N and Ble-C; the amplicon was blunt-ended and ligated into the blunt-cloning site of pT7Blue to generate pT7Blue-bleA. 2) The *ble* gene was recovered as a 382-bp *Bam*HI fragment from pT7Blue-bleA and cloned into *Bam*HI cleaved pMV261 to generate pMV261-ble. 3) A 1077-bp P$_{hsp60}$-*ble* cassette was PCR amplified from pMV261-ble using primers 981816 and 981819 and blunt cloned into pT7Blue to generate pT7Blue-bleB. 4) The P$_{hsp60}$-*ble* cassette was recovered from pT7Blue-bleB by *Xba*I digest and cloned into the *Xba*I site of a pBluescript KS+ based plasmid that had the mycobacteriophage L5 integrase gene cloned into the *Eco*RI site to generate pSEint.1. 5) The *M. tuberculosis sigE* gene was PCR amplified from H37Rv genomic DNA using primer 987917 and primer 987916 (which contains a *Bam*HI site), digested with *Bam*HI, and cloned into pMV261 digested with *Hpa*I and *Bam*HI. 6) The P$_{hsp60}$-*sigE* cassette was PCR amplified using primers 981816 and 981819 and cloned into pT7Blue to generate pT7Blue-P$_{hsp60}$-sigE. 7) The P$_{hsp60}$-*sigE* cassette was recovered from pT7Blue-P$_{hsp60}$-sigE by digestion *Kpn*I and cloned into the *Kpn*I site of pSEint.1 to generate pSEint.2.

Construction of pSEKO.4
1) A 757-bp fragment containing 246 bp of the 3' end of the *M. smegmatis sigE* gene and 511 bp of the 5' end of the adjacent *ORF1* gene was PCR amplified using primer 7208 (which contains a *Hpa*I restriction site) and primer 7209 and cloned into the *Eco*RV blunt-cloning site of pT7Blue to generate pSEKO.1. 2) A 760-bp fragment containing 526 bp of the 5' region of the *M. smegmatis sigE* gene and 234 bp of the region upstream was PCR amplified using primer 7206 and primer 7207 (which contains a *Hpa*I restriction site), digested with *Hpa*I, and cloned into the *Hpa*I site of pSEKO.1 to generate pSEKO.2. These steps resulted in a 2-bp deletion of the *sigE* gene. 3) The gentamicin-resistance gene (*aacC1*) was PCR-amplified from pPR27 as a 1270-bp fragment using primer 7204 and primer 7205 (both of which contain *Hpa*I restriction sites), digested with *Hpa*I, and cloned into the *Hpa*I site located between the two fragments of the *sigE* gene in pSEKO.2 to generate pSEKO3. 4) The *rpsL* promoter was isolated from pCL8 as a *Bam*HI-*Kpn*I fragment and cloned into *Bam*HI-*Kpn*I digested pAWS2 upstream of the *lacZ* gene. 5) The P$_{rpsL}$-*lacZ* cassette was removed from this recombinant by *Bam*HI and *Eco*RI digestion and cloned into *Bam*HI and *Eco*RI digested pBluescript SK+. 6) The P$_{rpsL}$-*lacZ* cassette was isolated from this subclone as a *Hin*dIII fragment, and the ends of the fragment were blunt-ended using the Perfectly Blunt Cloning Kit (Novagen, Madison, WI) and cloned into the *Sca*I site of pSEKO.3 to generate pSEKO.4.

Construction of pSE.4
1) A promoterless TN5 aph gene was PCR amplified from pUC4-KIXX as an 850-bp fragment using primers 1915 and 1916, and a promoterless xylE gene was PCR amplified from pTKMX97 as a 935-bp fragment using primers 1917 and 1918. Primers 1916 and 1917 have 20 nucleotides of complementary sequence so that the two PCR products can anneal. The two PCR products were mixed, allowed to anneal, and amplified using two primers, one (1915) complementary to the 5' end of the aph gene and a second (1918) complementary to the 3' end of xylE to produce a 1785-bp fragment. The amplified annealed fragment was gel purified and cloned into the SmaI site of pUC19 to generate pSE.1. 2) The hygromycin B-resistance gene (hyg) was recovered from pUS267 by HindIII digestion and cloned into the HindIII site of pSE.1 to generate pSE.2. 3) A 1903-bp fragment containing oriM was recovered from pMV261 by digestion with NotI and MluI, gel purified, blunt ended using the Perfectly Blunt Cloning Kit (Novagen), and ligated into pSE.2 digested with SspI to generate pSE.3. 4) An amplicon containing the E. coli rRNA transcriptional terminator was PCR amplified from pCL5 using primers 5077 and 5078 (which contain XbaI sites), digested with XbaI, and ligated into XbaI digested pSE.3 to generate pSE.4. Function of

Table 2: Oligonucleotide primers. Underlined bases are restriction sites relevant to the construction. MS: *Mycobacterium smegmatis*; **MTB:** *Mycobacterium tuberculosis*; **P**$_{hsp60}$: promoter for the *M. tuberculosis hsp60* gene. **P***rpsL*: promoter for the *M. smegmatis rpsL* gene

Primer Designation	Sequence 5' to 3'	Gene Amplified
Ble-N	ataaat ggatccaatggccaagttgaccagt	*ble*
Ble-C	aaataa ggatcctcagtcctgctcctcggc	*ble*
TN9-5R	cagggcaccggacaggtcggtc	*aph*
981816	atgtacgtggcgaactccgttgta	P$_{hsp60}$
981819	gatgatatattttttatcttgtgca	P$_{hsp60}$
987916	gggcccaa ggatccaatggaactcctcggcggaccc	MTB*sigE*
987917	agcgaactgggttgacgtgaactgc	MTB*sigE*
106777	tggaactcctcggcggacc	MTB*sigE*
106778	tgacgtgaactgcgcactcg	MTB*sigE*
7208	tgcga gttaaccgaaccccgagcagatctacgacg	MS*sigE*
7209	gcatctctagaagccgatcgcgtgtcggcgtc	MS *sigE*
7206	gcatcaagctttgctctagatcggagatcgaccgtttctg	MS*sigE*
7207	tcggt gttaaccctcggcgggcacgcggtcg	MS*sigE*
7204	cgtga gttaacctcgcgcagcgtgggtgc	*aacC1*
7205	gaccg gttaacggcgttgtgacaatttaccgaac	*aacC1*
7997	cctcgacatggtgcgccg	MS*sigE*
7998	caggcgcgaatcgtggtag	MS*sigE*
7999	gccgtctgcccaattgcag	MTB*sigE*
1915	gatca ggatccgatcaagagac	*aph*
1916	gcattacacttttgttcatttcgaaccccagagtcccgctc	*aph*
1917	agcgggactctggggttcgaaatgaacaaaggtgtaatgcg	*xylE*
1918	ccgggccggaccatcaggtcagc	*xylE*
5077	cgtgc tctagagcgagagtagggaactgc	rRNA
5078	CGTCGT CTAGAGACGACAGGAAGAGTTTGTAG	rRNA
47	CGGAGGAATCACTTCGCAATG	*hsp60*
39A	GGAGTTGAAGGCGATCTGCTT	*hsp60*
MA1	GTGCCCGCACGCCG GGTACCATGAACTATGAGG	P*rpsL*
MA2	CCTCATAGTTCAT GGTACCCGGCGTGCGGGCAC	P*rpsL*
TK97L	TCGA AAGCTTCGTCGCGGTCGGCCC	P*rpsL*

the terminator was determined by patching hygromycin B-resistant transformants LB medium containing kanamycin.

DNA manipulations and DNA sequencing

Restriction enzyme reactions were performed as recommended by the manufacturer (Invitrogen). Wizard Plus Minipreps DNA Purification System (Promega, Madison, WI) was used to isolate plasmid DNA from E. coli bacteria and from M. smegmatis bacteria as previously described [11]. Phosphatase reactions using Calf Intestinal Alkaline Phosphatase (Invitrogen) and ligation reactions using T4 DNA Ligase (Invitrogen) were carried out according to manufacturer's directions. Sequencing reactions were carried out using the Big Dye Terminator Cycle Sequencing Ready Reaction Kit as recommended by manufacturer (Applied Biosystems). Sequence reactions were analyzed using an ABI 373 DNA sequencer and sequence data were assembled using SeqEd 675 DNA Sequence Editor, version 1.0.3 (Applied Biosystems).

Genomic DNA isolation

M. smegmatis genomic DNA was purified by a modification of the glass-bead lysis method as previously described [12]. M. tuberculosis genomic DNA was isolated as previously described [13]. DNA concentration was estimated using a GeneQuant II apparatus (Amersham Biosciences).

Polymerase chain reaction

The primers used in this study are listed in Table 2. Primers were synthesized on a 381A DNA synthesizer (Applied Biosystems, Foster City, CA) at the Biotechnology Core Facility, National Center for Infectious Diseases, Centers for Disease Control and Prevention. 50 μl amplification reactions contained 100–200 ng DNA template and 45 μl of a reaction mix (200 μM each, deoxynucleotide triphosphates; 1.0 μM each primer; 1.25 U Taq polymerase; 10 mM Tris hydrochloride (pH 8.3); 50 mM KCl; 1.5 mM MgCl$_2$; 0.01% gelatin). Samples were amplified for 30 cycles in a GeneAmp PCR system 2400 thermocycler (Applied Biosystems) with a three-step cycle of

denaturation for 30 s at 94°C, annealing for 30 s at 55°C to 70°C (optimized for each primer pair), and extension for 1 min at 72°C. Amplification products were analyzed by electrophoresis through 0.8 to 2.0% agarose-TBE gel and visualized by ethidium-bromide staining.

Electroporation
Electrocompetent *M. smegmatis* bacteria and *E. coli* bacteria (Biorad) were produced and DNA was electroporated into these bacteria as previously described [14,15].

Selection of active promoters
Electrocompetent *M. smegmatis* LR222::SEint.2 bacteria were electroporated with 5 μg of the library plasmid DNA, and recipients were plated on TSA with kanamycin (50 μg/ml). After incubation at 37°C for 3 days, colonies were pooled and collected by flooding plates with 2 ml Middlebrook 7H9 broth and scraping the colonies off the plate. Plasmid DNA was harvested from the transformants, ethanol precipitated, and resuspended in water.

Screening to identify σ^E-dependent promoters
Electrocompetent *M. smegmatis* LR222Δ*sigE* cells were electroporated with approximately 1 μg plasmid DNA from the pool of kanamycin-resistant transformants, and recipients were plated on TSA with hygromycin B (50 μg/ml). After incubation at 37°C for 3 days, colonies were chilled for 2 h at 4°C. Chilled colonies were sprayed with 0.5 M catechol (Sigma) in 0.05 mM potassium phosphate buffer pH 7.4.

Results
Experimental strategy
The basic I-TRAP approach (Figure 1) uses a genomic DNA library in a plasmid vector containing a promoterless 'operon' with two reporter genes. One reporter gene, the *aph* gene which confers resistance to kanamycin, allows for selection of active promoters and a second reporter gene, the *xylE* gene which encodes catechol 2,3 dioxygenase, allows for screening for loss of promoter activity. In the first step of the I-TRAP approach, the recombinant plasmid library is transformed into bacteria expressing the transcriptional regulator. Transformants containing plasmids with active promoters are isolated by selection on kanamycin-containing media. A subset of the active promoters will be those whose expression depends on the transcriptional regulator. To differentiate between promoters dependent on the regulator and other promoters, plasmid DNA from a pool of the kanamycin-resistant transformants is transformed into bacteria that do not express the transcriptional regulator. The hygromycin B-resistant transformants are screened for XylE expression by spraying colonies with catechol. Yellow colonies arise from bacteria that contain a plasmid with a transcriptionally active promoter. White colonies arise from bacteria

that contain a plasmid with a promoter that is inactive in the absence of the transcriptional regulator.

To develop the I-TRAP method for identifying *M. tuberculosis* promoters dependent on the expression of σ^E, a strain of *M. smegmatis* over-expressing the *M. tuberculosis* σ^E, a σ^E mutant strain of *M. smegmatis*, and a *M. tuberculosis* promoter-trap library were created.

Construction of σ^E-over-expressing strain of **M. smegmatis**
The σ^E-over-expressing strain of *M. smegmatis* was made by transforming *M. smegmatis* LR222 bacteria with pSEint.2 (Figure 2). The pSEint.2 plasmid contains an *oriE* for plasmid replication in *E. coli*, the mycobacteriophage L5 integrase and attP site to allow integration into the *M. smegmatis* chromosome, two selectable genes, *ble* (bleomycin or zeocin resistance) and *amp* (ampicillin resistance), and the *M. tuberculosis* RV1221 ORF cloned downstream of the *M. tuberculosis* hsp60 promoter. In this strain, the *M. tuberculosis* sigE gene is maintained in a single copy in the genome and is constitutively expressed from the *hsp60* promoter. This strain is designated *M. smegmatis* LR222::SEint.2 (Table 1).

Construction of sigE *mutant strain of* **M. smegmatis**
The suicide plasmid pSEKO.4 (Figure 3) was constructed to interrupt the *M. smegmatis* sigE gene by gene replacement [14]. pSEKO.4 contains the *M. smegmatis* sigE gene interrupted with a gentamicin-resistance gene and the *lacZ* gene. The pSEKO.4 DNA (1 μg) was UV-irradiated in a UV Stratalinker 1800 (Stratagene) at 100 mJ cm^-2 and electroporated into *M. smegmatis* cells. Following outgrowth at 37°C for 4 h, transformants were recovered by plating on TSA containing gentamicin and X-Gal. 24 white transformants were analyzed by PCR using primers (7997 and 7998) that bind to the *sigE* gene on either side of the *aacC1* gene (Table 2). Of the 24 transformants, 11 had the *M. smegmatis* sigE gene disrupted by *aacC1*, 9 had only a single recombination event, and 4 only had a wild-type copy of *sigE* (data not shown). One transformant was chosen as the *M. smegmatis* sigE mutant strain and designated LR222Δ *sigE* (Table 1).

Construction of a library in the promoter-trap plasmid
pSE.4 (Figure 4) is a promoter-trap plasmid containing a promoterless *aph* gene and *xylE* gene, an *E. coli* rRNA transcriptional terminator upstream of the reporter genes, an *oriM* for plasmid replication in *M. smegmatis*, an *oriE* for plasmid replication in *E. coli*, and a *hyg* gene for selection in *E. coli* and *M. smegmatis* bacteria. The plasmid contains a unique *Bam*HI site upstream of the promoterless reporter operon for cloning of potential promoters. To generate fragments of the *M. tuberculosis* H37Rv genome, ~3.0 μg of genomic DNA was digested with 1 U *Sau*3AI in
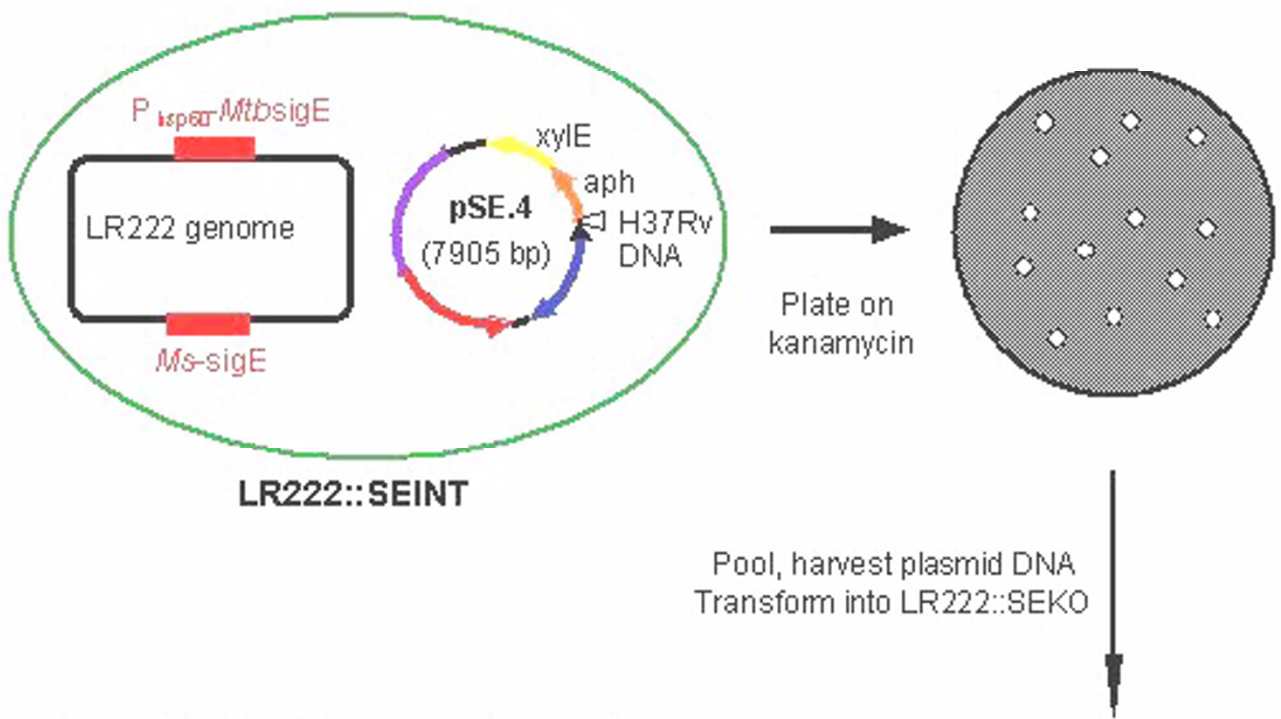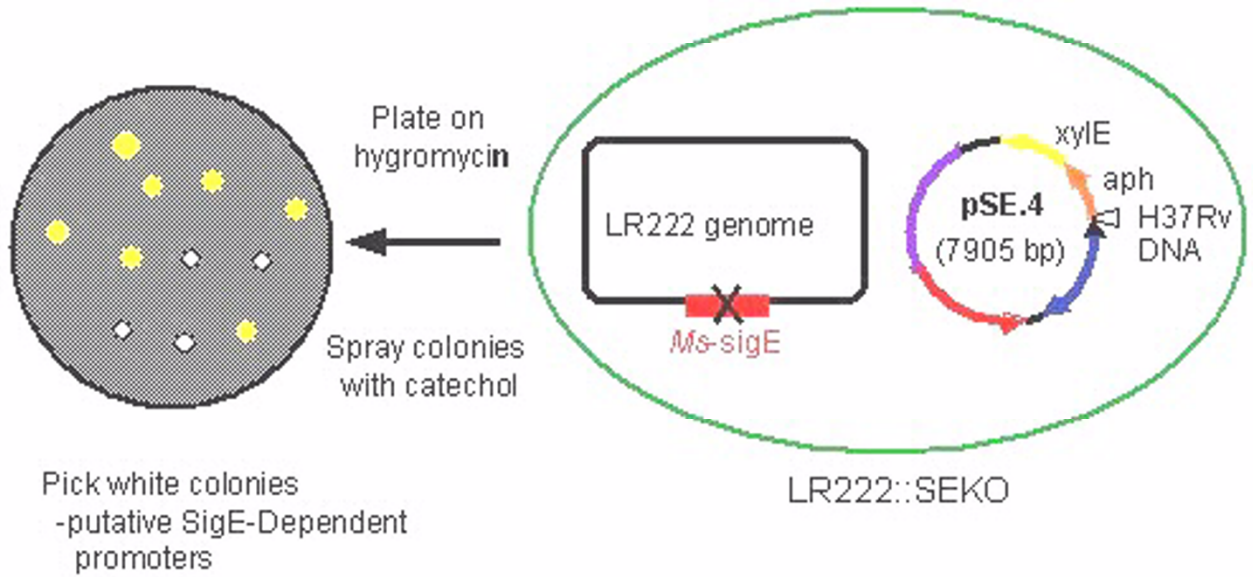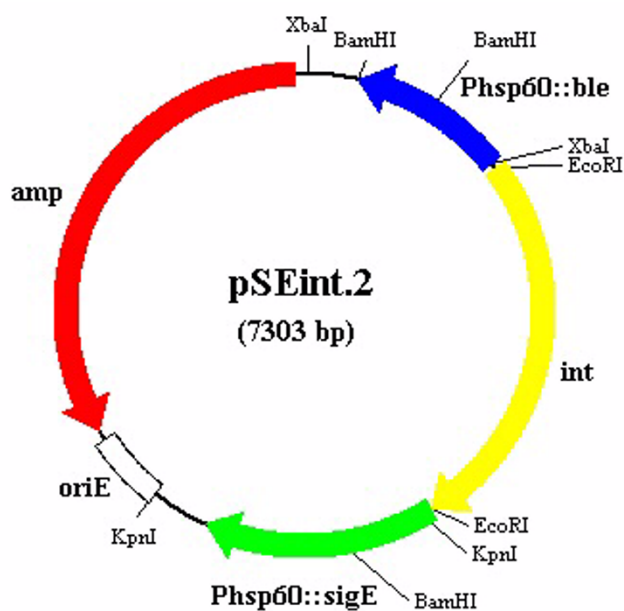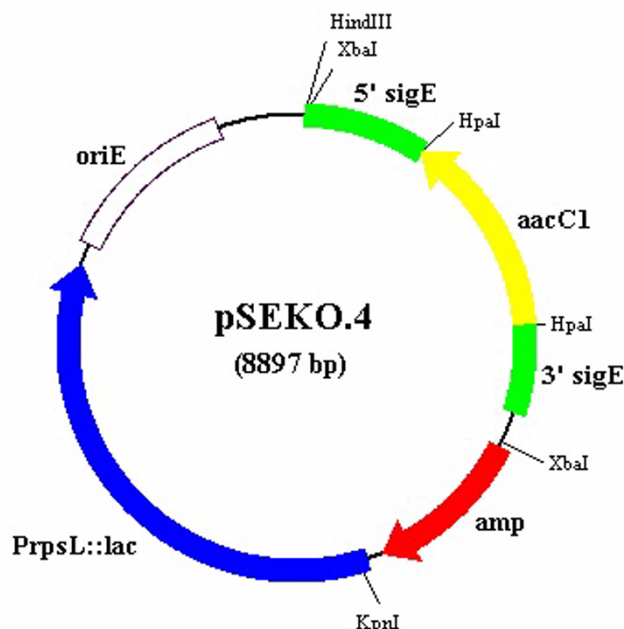
Figure 1
Schematic of I-TRAP approach.

Figure 2
Map of pSEint.2. pSEint.2 contains an *oriE* for plasmid replication in *E. coli*, the mycobacteriophage L5 integrase (*int*) and attP to allow integration into the *M. smegmatis* chromosome, a ble gene (bleomycin resistance) for selection in mycobacteria, an *amp* gene (ampicillin resistance) for selection in *E. coli*, and the *M. tuberculosis sigE* gene under the control of the constitutively expressed *hsp60* promoter. Arrowheads indicate direction of transcription.

Figure 3
Map of pSEKO.4 pSEKO.4 contains the *M. smegmatis sigE* gene interrupted with a gentamicin-resistance gene, *aac1*, the *lacZ* gene for use as a screen against single recombination events, and an oriE for plasmid replication in *E. coli*. Arrowheads indicate direction of transcription. The construction of the mutated *sigE* gene resulted in an insertion of the *aac1* cassette and a 2-bp deletion of the *sigE* gene at the site of insertion.

20 μl total volume for 2 min at 37°C and the reaction stopped by heating at 65°C for 10 min. This generated fragments between 100 and ~6,000 bp (data not shown). *Sau*3AI was removed from the reaction by phenol-chloroform extraction. DNA was ethanol precipitated and resuspended in 20 μl Tris buffer pH 7.4. Approximately 470 ng digested genomic DNA was ligated with ~200 ng pSE.4 DNA that had been digested with *Bam*HI and phosphatased. DNA was ethanol precipitated from the ligation reaction and resuspended in water; 200 ng DNA was electroporated into Electromax DH5α cells. Recipient bacteria were spread on LB agar containing hygromycin B. After overnight incubation at 37°C, 50 colonies were selected at random, plasmid DNA was harvested, and genomic DNA inserts were sequenced using primer TN5-9R (Table 2). All plasmids contained a unique *M. tuberculosis* DNA segment and had an average insert size of approximately 200 bp (data not shown).

To generate a library of recombinant plasmids, 200 ng of the ligated sample were electroporated into Electromax

DH5α cells. The entire mixture was plated on 117 150 mm LB agar plates containing hygromycin B. About 117,000 transformants were collected and pooled by flooding each plate with 2 ml LB and scraping the colonies off the plate. Plasmid DNA was harvested from the pooled *E. coli* transformants, ethanol precipitated, and resuspended in water.

Promoter selection and scoring
To recover plasmids containing active promoters, the plasmid library DNA was electroporated into *M. smegmatis* LR222::SEint.2 bacteria, which constitutively express the *M. tuberculosis sigE* gene. To determine the proportion of active promoters in the library, a portion of the transformation mixture was spread on medium containing hygromycin B (all recipients), and a portion was spread on medium containing kanamycin (recipients with an active promoter). 1.93×10^6 hygromycin B-resistant colony forming units (CFUs) /ml and 3.75×10^4 kanamycin-resistant CFU/ml were recovered; therefore, approximately 1.9% of transformants contained an active pro-
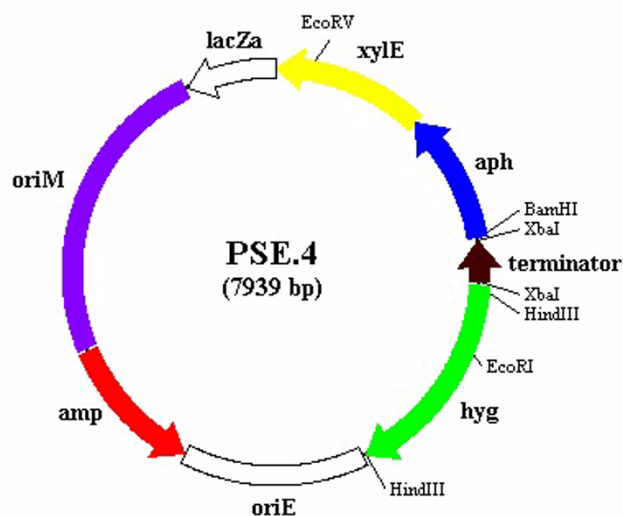
Figure 4
Map of pSE.4. pSE.4 is a promoter-trap plasmid containing a promoterless *aph* gene and *xylE* gene, an *E. coli* rRNA transcriptional terminator cloned upstream of the promoterless *aph* and *xylE* genes, a unique *BamHI* cloning site, an *oriM* for plasmid replication in mycobacteria, an *oriE* for plasmid replication in *E. coli*, and a *hyg* gene for selection in *E. coli* and *M. smegmatis* bacteria. Arrowheads indicate direction of transcription.

moter fragment. When the initial pool of kanamycin-resistant colonies sprayed with catechol, the colonies were light to dark yellow indicating a wide range of XylE activity.

To differentiate between σ^E-dependent and -independent promoters, approximately 40,000 kanamycin-resistant transformants were pooled, and plasmid DNA was isolated and electroporated into *M. smegmatis* LR222::Δ *sigE* bacteria, in which the wild type copy of the *sigE* gene had been disrupted with the gentamicin-resistance gene. Transformants were plated on hygromycin B and assayed for catechol 2,3 dioxygenase activity by spraying colonies with 0.5 M catechol. About 20,000 colonies were sprayed with catechol, and a range of color intensity from white to bright yellow was observed. A total of 360 white and light yellow colonies were recovered. Overall, about 1.8% of hygromycin B-resistant colonies were white or light yellow.

To confirm the phenotypes of plasmids recovered in this two-step procedure, plasmids were recovered from 18 white colonies and individually transformed into *M. smegmatis* LR222::SEint.2 (SigE-expressing) and *M. smegmatis* LR222::Δ *sigE* bacteria. All 18 plasmids generated

only white colonies when transformed into the σ^E-mutant strain. However, although all 18 plasmids generated yellow colonies when transformed into σ^E-expressing bacteria, some plasmids produced a mixture of white and yellow colonies. These white colonies were consistently white on retesting and their plasmids usually contained deletions that removed the cloned *M. tuberculosis* sequence. Replating of the yellow colonies produced a mixture of white and yellow colonies, suggesting that the expression of the *xylE* or *aph* gene in the SigE-expressing strain may be detrimental to the cell and may lead to plasmid instability.

A preliminary analysis of the sequences of the inserts in the 18 clones identified 6 inserts that had matches with the SigE consensus sequence [6]. Interestingly, none of these genes were identified as being SigE-induced genes in microarray studies [6], although the expression of one of the genes, Rv3223c (*sigH*), is thought to be SigE-regulated [5].

Discussion
A key feature of the I-TRAP method is that it allows two degrees of promoter analysis, i.e., with and without a transcriptional regulator. The first step allows selection of all active promoters under one condition, and the second step sorts promoters by activity dependent on a second condition. In the study reported here, we used this method to identify potential σ^E-induced genes. The basic I-TRAP approach might also be used to identify promoters repressed by a particular protein. The transformation steps are reversed so that the promoters are selected in the mutant background first (i.e., no repressor) and then screened for loss of activity in the presence of the repressor protein.

The I-TRAP method might also be used to characterize promoter activity in two environmental conditions in a manner similar to the IVET approach [16]. This is useful because, in some cases, the factor necessary for transcriptional activation during a particular environmental condition is unknown. In this case, after exposing bacteria containing the recombinant library to an environmental condition (e.g., a heat shock or acid shock or growth in medium A), antibiotics could be added to the medium to kill any bacteria that had not expressed the antibiotic resistance reporter gene during the exposure. For mycobacteria, a two-hour treatment with kanamycin would be sufficient to kill bacteria that had not expressed the kanamycin-resistance gene during the stress. The resistant bacteria would then be grown in a second medium or in the absence of a stress, and colonies scored for lack of promoter activity in the second condition by expression of the *xylE* reporter gene. White colonies would contain a plasmid with a promoter that is active only during the first

condition. Using I-TRAP for this purpose may provide a means to identify promoters dependent on any of several regulatory proteins that may be active during the stress and to obtain a global overview of gene expression during the stress.

In the I-TRAP method, the recipient strain can be the native bacterium or a surrogate host. We used *M. smegmatis* as the host because it grows much more rapidly than does *M. tuberculosis* (generation time of 3 hrs vs. 24 hrs), does not require BSL-3 facilities, and is more easily manipulated genetically than *M. tuberculosis*. One possible advantage of using a surrogate host is that studies could be focused on one particular regulator which might avoid complications due to other regulatory proteins in the cell that may be able to recognize promoters in the library. Of course, the surrogate host must not have a regulatory protein that recognizes the same promoters as the regulator being studied. In our studies, the *M. smegmatis* σ^E protein has 92% homology with the *M. tuberculosis* σ^E protein [17]. While this suggests that the *M. tuberculosis* σ^E protein should be able to interact properly with the *M. smegmatis* RNA polymerase, it also suggests that the *M. smegmatis* σ^E homolog might be able to recognize *M. tuberculosis* σ^E-dependent promoters. Because of this, we constructed an *M. smegmatis* host strain that lacked any functional σ^E protein.

An assumption of the I-TRAP approach is that the sequences cloned in the plasmids will be recognized by transcriptional regulatory proteins in the same manner as the intact sequences in the genome. The cloning process might generate false-positive or false-negative results because of the cloned sequences being recognized out of the context of the surrounding genomic sequences.

Another assumption of the I-TRAP approach is that over-expression of a transcriptional regulator will be sufficient to induce gene expression from its regulated promoters. This suggests that proteins, such as PhoP, that require activation to promote transcription may not be approachable with this method unless the over-expressed regulatory protein can also be activated. Also, certain regulatory proteins may not be suitable for study in a surrogate host bacterium [18]. For proteins that require the presence of additional factors, such as σ^N, or interaction with host proteins to be active, such as σ^E with RNA polymerase, use of the native bacterium or closely related surrogate host may be required to ensure the presence of the necessary interacting proteins [19]. Some of these additional factors, however, may only be expressed or active under a particular growth condition and therefore may also need to be conditionally expressed along with the transcription factor. Promoters that require both a positive transcription regulator and the absence of a repressor protein may be missed because expressing the positive transcription factor alone may not be sufficient to induce transcription at such promoters. Another limitation of the I-TRAP method is that some transcription factors may not be stably maintained in the bacterium or may be lethal to the bacterium when over-expressed. To avoid this problem, an inducible promoter could be used to promote transcription of the gene encoding the transcription factor.

Another possible complication of this approach is that expressing a foreign protein may turn on the stress response of the host bacterium. Transcription factors mediating the stress response to the expressed foreign protein may be able to recognize promoters within the library and promote transcription of the reporter genes, thus producing the false positives in the two-step screen.

Because the first step of the I-TRAP method involves selection for active promoters, one could easily begin with a library containing recombinants representing all possible promoters in a bacterial genome. Unfortunately, the recombinant DNA library used in this study contained only ~65% of the genome in the correct orientation upstream of the reporter genes, which will limit the number of different promoters we can recover using this library. For the screening step, we used the *xylE* gene because catechol 2,3 dioxygenase activity is easily measured by spraying colonies with catechol and the amount of color produced is a rough measure of promoter activity [10]. In our screen, white, pale yellow, and bright yellow colonies were observed, indicating a wide range of promoter activity. Pale yellow colonies may represent weak σ^E-independent promoters or promoters that are recognized by another sigma factor in addition to σ^E, providing a low level of constitutive expression. These possibilities could be distinguished by transformation of plasmids from individual clones into the σ^E-over-expressing strain and directly comparing of XylE activity in the presence and absence of σ^E. Unfortunately, it appears that expression of XylE in *M. smegmatis* may be detrimental and may lead to plasmid instability and loss of the inserted DNA. This raises the possibility that some white colonies may arise in the second step of I-TRAP by loss of a promoter fragment from a plasmid recovered in the first step. Nonetheless, the I-TRAP method should greatly enrich for promoters with the desired activities. As is generally true for genetic screens, it is essential to study any identified promoters in their native state to understand the regulation of their expression.

Conclusions

The I-TRAP method is a versatile and useful method for characterizing promoter activity under a variety of conditions and in response to various regulatory proteins. In our study, we isolated 360 clones that may contain plas-

mids carrying SigE-regulated promoters genes of *M. tuberculosis*. The analysis of these clones will be reported elsewhere.

Competing interests

No competing interests for either author.

Authors' contributions

The work is a portion of MA's doctoral thesis. MA designed and conducted the studies, analyzed the results, and drafted the manuscript. TS participated in the design and coordination of the study, analysis of results, and preparation of the manuscript. All authors read and approved the final manuscript.

Acknowledgements

References

1. Barker LP, Brooks DM and Small PL: **The identification of *Mycobacterium marinum* genes differentially expressed in macrophage phagosomes using promoter fusions to green fluorescent protein** *Mol Microbiol* 1998, **29:**1167-1177.
2. Dubnau E, Fontan P, Manganelli R, Soares-Appel S and Smith I: **Mycobacterium tuberculosis Genes Induced during Infection of Human Macrophages** *Infect Immun* 2002, **70:**2787-2795.
3. Graham JE and Clark-Curtiss JE: **Identification of *Mycobacterium tuberculosis* RNAs synthesized in response to phagocytosis by human macrophages by selective capture of transcribed sequences (SCOTS)** *Proc Natl Acad Sci USA* 1999, **96:**11554-11559.
4. Jensen-Cain DM and Quinn FD: **Differential expression of *sigE* by *Mycobacterium tuberculosis* during intracellular growth** *Microb Pathog* 2001, **30:**271-278.
5. Manganelli R, Dubnau E, Tyagi S, Kramer FR and Smith I: **Differential expression of 10 sigma factor genes in *Mycobacterium tuberculosis*** *Mol Microbiol* 1999, **31:**715-724.
6. Manganelli R, Voskuil MI, Schoolnik GK and Smith I: **The *Mycobacterium tuberculosis* ECF sigma factor sigmaE: role in global gene expression and survival in macrophages** *Mol Microbiol* 2001, **41:**423-437.
7. Erickson JW and Gross CA: **Identification of the sigma E subunit of *Escherichia coli* RNA polymerase: a second alternate sigma factor involved in high-temperature gene expression** *Genes Dev* 1989, **3:**1462-1471.
8. Kroos L, Zhang B, Ichikawa H and Yu YT: **Control of sigma factor activity during Bacillus subtilis sporulation** *Mol Microbiol* 1999, **31:**1285-1294.
9. Marston BJ and Shinnick TM: **Differentially expressed genes of *Mycobacterium tuberculosis*** *Ann N Y Acad Sci* 1996, **797:**32-41.
10. Kenney TJ and Churchward G: **Genetic analysis of the *Mycobacterium smegmatis* rpsL promoter** *J Bacteriol* 1996, **178:**3564-3571.
11. Plikaytis BB and King CH: **Nucleic-acid extraction: plasmid/cosmid DNA** In: *Mycobacterial Protocols Volume 101.* Edited by: *Parish T, Stoker NG. Totowa, NJ: Humana Press*; 1998:45-50.
12. Plikaytis BB, Gelber RH and Shinnick TM: **Rapid and sensitive detection of Mycobacterium leprae using a nested-primer gene amplification assay** *J Clin Microbiol* 1990, **28:**1913-1917.
13. van Soolingen D, de Haas PE, Hermans PW and van Embden JD: **RFLP analysis of Mycobacteria** In: *RFLP analysis of Mycobacteria.* Bilthoven, The Netherlands, National Institute of Public Heath and Environmental Protection 1995:5-10.
14. Hinds J, Mahenthiralingam E, Kempsell KE, Duncan K, Stokes RW, Parish T and Stoker NG: **Enhanced gene replacement in mycobacteria** *Microbiol* 1999, **145:**519-527.
15. Jacobs WR Jr, Kalpana GV, Cirillo JD, Pascopella L, Snapper SB, Udani RA, Jones W, Barletta RG and Bloom BR: **Genetic systems for mycobacteria** *Meth Enzymol* 1991, **204:**537-555.
16. Mahan MJ, Slauch JM and Mekalanos JJ: **Selection of bacterial virulence genes that are specifically induced in host tissues** *Science* 1993, **259:**686-688.
17. Wu QL, Kong D, Lam K and Husson RN: **A mycobacterial extracytoplasmic function sigma factor involved in survival following stress** *J Bacteriol* 1997, **179:**2922-2929.
18. Groisman EA: **The pleiotropic two-component regulatory system PhoP-PhoQ** *J Bacteriol* 2001, **183:**1835-1842.
19. Merrick MJ: **In a class of its own – the RNA polymerase sigma factor sigma 54 (sigma N)** *Mol Microbiol* 1993, **10:**903-909.
20. Miller LP, Crawford JT and Shinnick TM: **The *rpoB* gene of *Mycobacterium tuberculosis*** *Antimicrob Agents Chemother* 1994, **38:**805-811.
21. Steenken W Jr, Oatway WH Jr and Petroff SA: **Biological studies of the tubercle bacillus III. Dissociation and pathogenicity of the R and S variants of the human tubercle bacillus (H37)** *J Exp Med* 1934, **60:**515-540.
22. Stover CK, de la Cruz VF, Fuerst TR, Burlein JE, Benson LA, Bennett LT, Bansal GP, Young JF, Lee MH and Hatfull GF: **New use of BCG for recombinant vaccines** *Nature* 1991, **351:**456-460.
23. Pelicic V, Jackson M, Reyrat JM, Jacobs WR Jr, Gicquel B and Guilhot C: **Efficient allelic exchange and transposon mutagenesis in *Mycobacterium tuberculosis*** *Proc Natl Acad Sci USA* 1997, **94:**10955-10960.

Pre-publication history

The pre-publication history for this paper can be accessed here:

http://www.biomedcentral.com/1471-2334/3/15/prepub