Construction of an in-frame deletion of eccD3 in M. smegmatis by efficient allelic exchange method

2.1 Introduction

Mycobacterium sp. represents one of the important and very ancient groups of pathogens. The evolving trends leading to pathogenicity of this organism is very interesting as several unanswered questions relating to *M. tuberculosis*, the most important human pathogen of the genus, may be explored with an objective of identifying potential drug targets.

M. smegmatis, which belongs to the Non Tuberculous Mycobacteria (NTM) is a soil dwelling bacteria and shares about 1080 clusters referred to as core orthologs with *M. tuberculosis* [1]. The core orthologs involve the essential biological processes required for the mycobacterial growth such as amino acid and purine/pyrimidine biosynthetic pathways, processes related to DNA transcription, translation and repair, biosynthetic pathway of cell wall synthesis *etc.* [2]. These orthologs present in the pathogenic strain shares about 70% identity with *M. smegmatis*, the non-pathogenic strain [1]. A notable example of genomic conservation is the unique secretion system known as the Type VII secretion system (T7SS) or the ESAT6 secretion system. This T7SS is conserved among all the species of *Mycobacterium* [3]. The T7SS consists of five gene clusters designated as ESX1, ESX2, ESX3, ESX4 and, ESX5 [4]. Out of the five ESX loci present in *M. tuberculosis*, ESX1, ESX3and, ESX5 play a pivotal role in the virulence of the organism [5-7]. The ESX1 secretion system plays an important role in the host-pathogen interaction. A fragment of DNA (~ 9.5kb) known as the 'Region of Difference1' (RD1) includes the ESX1 locus. It was found later that the only attenuated vaccine, BCG has a deletion of this RD1 region which comprised the ESX1 locus [8, 9]. The two effectors ESAT6 and CFP10 secreted by ESX1 help the bacteria to evade the host immune system by escaping and inhibiting the phagolysosome fusion [10-12]. The ESX3 secretion system is associated with the regulated uptake of iron in *M. tuberculosis* and plays an important role in the virulence of the bacteria [13, 14]. The ESX5 secretion system which is unique to the pathogenic slow growing mycobacteria secretes the ESXA and ESXB paralogs ESXN and ESXM along with some PE/PPE proteins. Reports suggest that ESX5 is essential for the mycobacterial growth and full virulence of the tubercle bacilli. *M. tuberculosis* mutants deficient in the secretion of the PE/PPE proteins were highly attenuated and could not elicit the T-cell response in the host [15, 16]. Although the five ESX clusters are present in *M. tuberculosis* many *Mycobacterium sp.* harbors few of the ESX loci

[4]. The non-pathogenic strain *M. smegmatis* comprises of three out of the five ESX clusters *viz*., ESX1, ESX3, and ESX4. The non-pathogenic strain possesses a highly homologous esxA and esxB locus of the ESX1 secretion system and is essential for the transfer of DNA [17, 18]. The ESX3 secretion system is necessary for the natural growth of *M. smegmatis* under iron deprivation [19]. Each of the ESX clusters is a protein complex of ~1.5MD that shares a conserved inner membrane and a cytosolic apparatus. It consists of the ESX conserved components (Ecc), ESX secretion associated proteins (Esp), ESX and PE/PPE proteins. The core channel is composed of EccA, EccB, EccC, EccD and, EccE. The other components of the ESX systems include membrane-bound mycobacteria-specific subtilisin-like serine proteases named mycosins [20]. Although the function of ESX3 T7SS is known to be absolutely critical for establishing infection by *M. tuberculosis*, its conserved nature in all the pathogenic and nonpathogenic mycobacterial species is intriguing and extends the scope for further exploration for the additional functional roles in *Mycobacterium sp*. This would enable identification of potent targets for drugs in *M. tuberculosis*.

With the development of modern molecular tools, investigation of the functional role of a particular gene or a gene cluster in the organisms has become possible. Construction of mycobacterial mutants is a fundamental approach to study the functional role of a gene. These mutations can be achieved by targeted mutagenesis or random mutagenesis. For a successful genetic manipulation in the mycobacterial DNA efficient allelic exchange or transposon mutagenesis has to occur. This is possible with the help of efficient delivery system harboring the disrupted gene or the transposon. When this recombinant delivery vector is transformed into the Mycobacterial genome, homologous recombination or transposition takes place resulting in the mutation of the desired gene. In this regard a wide range of delivery vectors are used including the suicide plasmid vectors, non-replicating plasmids, counter-selectable markers, long linear DNA fragments, and mycobacteriophages mediated transduction have been developed to obtain mycobacterial mutants [21-25]. The use of the allelic exchange method employed to successfully disrupt the genes has certain advantages. In a targeted disruption method the essentiality of a gene might not be known. This may even lead to polar effect in which deletion of a particular gene influences the expression of the gene present downstream of the gene of interest. However, in mutations achieved through in frame deletion the

disadvantages of the targeted disruption can be avoided. The in frame deleted mutants are antibiotic sensitive and cannot revert back and the polar effect is also absent [25-27].

In light of the earlier studies, the present study relates to the generation of an in-frame deletion of the ESX3 associated *eccD3* locus in *M. smegmatis*, which forms the central channel of the ESX3 secretion system in the cell membrane. Here, using the efficient allelic exchange method the complete copy of the *eccD3* gene was deleted from the *M. smegmatis* genome.

2.2 Materials and methods

2.2.1 Sequence alignment of msm MSMEG_0623 and mtb Rv0290

The amino acid sequence of EccD3, the transmembrane protein from *M. tuberculosis* and *M. smegmatis* with the gene annotation Rv0290 and MSMEG_0623 were retrieved from http://svitsrv8.epfl.ch/tuberculist/ and http://svitsrv8.epfl.ch/mycobrowser/smegmalist.html databases respectively. The amino acid sequences were then aligned and analyzed using the CLC sequence viewer.

2.2.2 Bacterial strains and growth conditions

All the cloning procedures explained in this work were conducted in the *E. coli* strain DH5α.This strain was grown in Luria Bertani (LB, HiMedia) broth or LB agar (2%) solid medium. The parent wild-type strain of *M. smegmatis* used in this study was mc^2155. *M. smegmatis* was grown in Middlebrook 7H9 (MB-7H9) broth (HiMedia) (per liter: $(NH_4)_2SO_4$-0.5g, Na_2HPO_4-2.5g, KH_2PO_4-1.0g, $MgSO_4$-0.05g, $CaCl_2$-0.0005g, $ZnSO_4$-0.001g, $CuSO_4$-0.001g, L-Glutamic acid-0.5g, SodiumCitrate-0.1g, Pyridoxine-0.001g, Biotin-0.0005g, Ferric ammonium citrate-0.04g, supplemented with 0.05%Tween 80 and 2% glucose or Middlebrook 7H9 (MB-7H9, HiMedia) agar (1.5%) solid medium supplemented with 2% glucose. Ampicillin (100µg/ml for *E. coli*), Kanamycin (25µg/ml for *E. coli* and 25µg/ml for *M. smegmatis*) and Gentamicin (10µg/ml for *E. coli* and 10µg/ml for *M. smegmatis*) were used as appropriate. Growth conditions for *E. coli* was fixed at 37^0C for 12 hours, whereas, growth conditions for *M. smegmatis* was modified as per requirement.

2.2.3 Plasmids and primers used in the study

The plasmids and the primers used in the study are listed in Table 2.1 and Table 2.2 respectively.

Table 2.1: The table summarizes all the plasmid vectors used as a part of the study.

Plasmids	Details
pGEM7Zf(+)	2.9kb *E. coli* cloning vector, *lacZ*, *amp*^{r}, f1 origin.
pUC4K	3.9kb *E. coli* vector, source of *kanr* gene.
pPR27	*E. coli – Mycobacterial* shuttle vector, *ts oriM, sacB, gm*^{r}.
pD3UP*.	1 kb upstream region of *eccD3* cloned into *XbaI-EcoRI* sites of pGEM-7Zf(+) vector
pD3UPDN*	1.2 kb downstream region of *eccD3* cloned into *EcoRI-HindIII* sites of pD3UP.
pD3UPDNkan*	*EcoRI* insert of pUC4K containing the *kan*^{r} gene cloned into *EcoRI* site of pD3UPDN.
pPRD3KO*	3.4 kb construct containing upstream and downstream region of *eccD3* with *kan*^{r} gene, cloned into pPR27 between *XbaI* sites.

Table 2.2: This table summarizes all the primers used for Polymerase Chain Reactions (PCR) in the present study.

Primers	Sequences	Restriction sites
Ms0622For	GCC**TCTAGA**CAGAACACCATGGCGATGAG	*XbaI*
Ms0622Rev	GCC**GAATTC**GCATCACAGTGTTCTCGGAC	*EcoRI*
MS0624For	GCC**GAATTC**TGATCCACAAGAGTCTGGGC	*EcoRI*
Ms0624Rev	GCC**AAGCTT**CATTGCTCACCGAGACGATG	*HindIII*
pD3UPDNKan For	GCC**TCTAGA**CAGAACACCATGGCGATGAG	*XbaI*
pD3UPDNKan Rev	GCC**TCTAGA**CATTGCTCACCGAGACGATG	*XbaI*

Set1For	ACGCGGCACTACGCCGACAGGGTTG	-
Set1Rev	GCTGGCCGTCGCGGCCATCGGGGACC	-
Set2 For	GCCTGAGCGAGACGAAATACGCGATCG	-
Set2 Rev	GGCGGTTAGAAGCTTAGACC	-
Set3 For	CAGAAAAGCTAGCGTTGGAC	-
Set3 Rev	GGCGGTTAGAAGCTTAGACC	-

2.2.4 Construction of a suicide delivery vector for the replacement of eccD3 with kan^r using pPR27 vector.

A suicide delivery vector bearing the recombination cassette was constructed (Figure 2.1) using the principle of allelic exchange mutagenesis in *M. smegmatis* [25]. The vector pPR27, a mycobacterial shuttle vector was used, which consisted of an *E. coli* origin of replication, *oriE* and a thermo-sensitive mycobacterial origin of replication, *ts-oriM*. Thus, the vector is able to reside and replicate in mycobacteria at 30℃ or below, but it gets disintegrated and is unable to replicate at temperatures 39℃ or above. The vector backbone also carries the different selectable markers such as the gene for gentamycin resistance, *gent^r* and a *sacB* gene which renders sucrose sensitivity to the host harboring it. The *sacB* gene encodes for the secreted enzyme levansucrase which catalyzes the hydrolysis of sugar and synthesis of high molecular weight fructose polymers, levan. Mycobacterial cell transformed with the vector harboring the *sacB* gene gets expressed in presence of 10% sucrose and thereby, results in lethality of the mycobacterial cells. Thus, *sacB* and *ts-oriM* provide efficient counter selective measures for identification of the mutants.

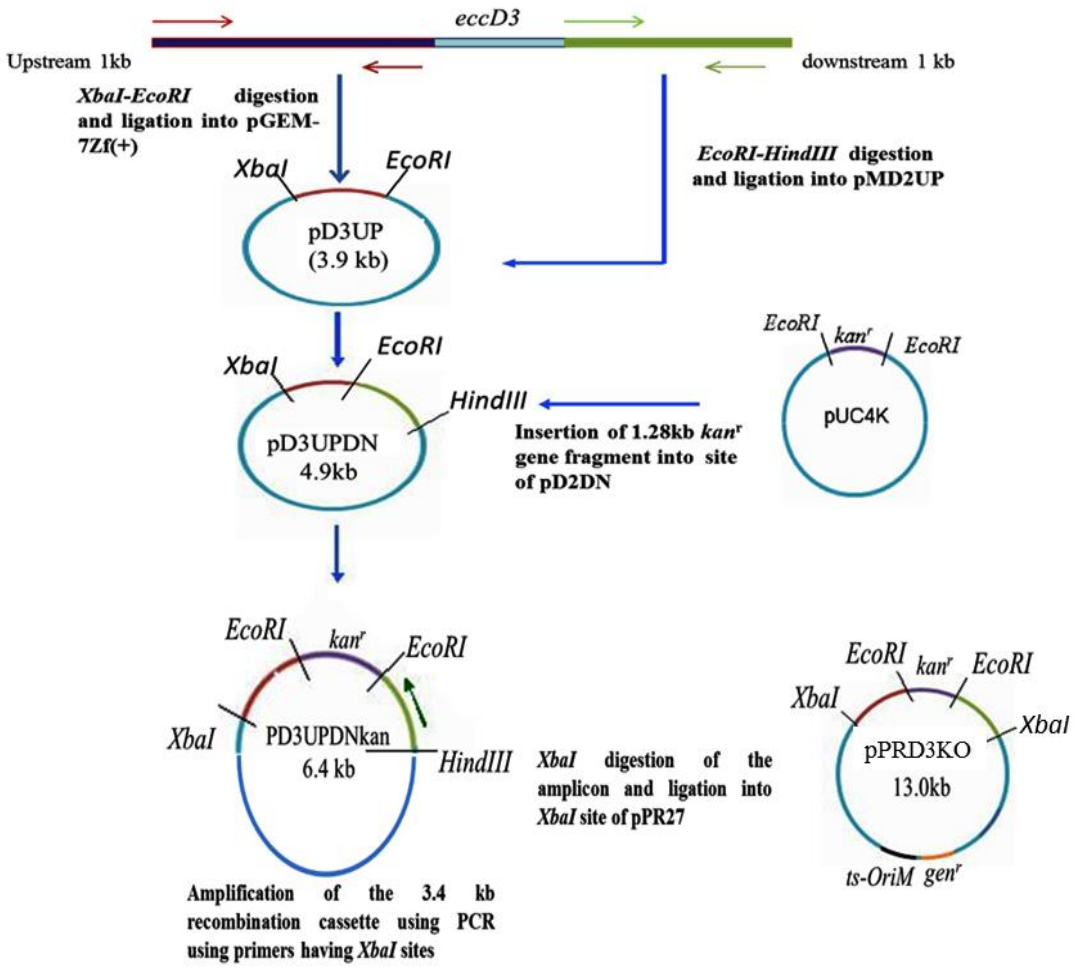


Figure 2.1: A schematic representation of all the steps involved in the construction of the recombination cassette for the disruption of *eccD3* in *M. smegmatis* using *kan*^{r} (*aph*) gene. The sizes of the vectors are approximate estimations of their size.

2.2.4.1 Cloning of the upstream 1kb region of eccD3 into pGEM-7Zf(+) vector

A DNA fragment upstream of the *eccD3* gene, approximately 975bp in size, was amplified by Polymerase Chain Reaction (PCR) using the primers; forward primer- Ms0622For and reverse primer- Ms0622Rev (Table 2.2). The concentration of the components used in the PCR reaction and the conditions for the amplification of the target genes have been summarized in the Tables 2.3 and 2.4 respectively. The amplified upstream fragment PCR product was digested with *XbaI* and *EcoRI* (Fermentus) and cloned into the vector pGEM-7Zf(+) digested with the same enzymes to yield pD3UP. The recombinant plasmid pGEM-7Zf(+) named as pD3UP was used to transform the competent *E.coli* DH5α cells. The clones were screened for gain in ampicillin resistance rendered by the parent vector on LB-agar plates

containing Ampicillin (100µg/ml). The clones were confirmed using restriction digestions to release the insert.

Table 2.3: The table summarizes the components of the PCR reaction mixtures.

Components	Concentration
Primers	0.2µM each
dNTPs	0.2mM each
Buffer	1X
Template DNA	50ng – 100ng
DMSO	5%
Nuclease free H_2O	As per requirement
Total reaction volume	25µl

The DNA dependent DNA polymerase used was DyNAzyme EXT Polymerase (Fermentus) at 1 U/50µl reaction. DyNAzyme EXT polymerase, being a proofreading polymerase with 3′-5′proof reading capacity was used to minimize the error rate. This enzyme is also capable of long range amplification on a GC-rich template, which makes it suitable for amplifications from the high GC-containing mycobacterial genome. The thermal cycling parameters are summarized in Table 2.4. The parameters were kept constant for most of the PCR amplifications.

Table 2.4: The table summarizes the parameters for thermal cycling reactions in the PCR.

Steps	Temperature (^{0}C)	Time (mins)
1. Initial denaturation	95^{0}C	5 mins
2. Denaturation	95^{0}C	1 min (29 cycles)
3. Annealing	56.2^{0}C	1 min (29 cycles)
4. Extension	72^{0}C	4 mins (29 cycles)
5. Final extension	72^{0}C	7 mins
6. Hold at	4^{0}C	α

2.2.4.2 Cloning of the 1.2 kb region downstream of eccD3 into pD3UP

A DNA segment approximately 1kb downstream of the *eccD3* gene was amplified using the primers MS0624For and Ms0624Rev (Table 2.2). The PCR conditions were the same as described earlier (Table 2.3 and Table 2.4). The amplified product was digested with *EcoRI* and *HindIII* and ligated into *EcoRI*-

HindIII digested pD3UP to generate pD3UPDN. The clones harboring the pD3UPDN plasmid were screened for gain in ampicillin resistance as previously described.

2.2.4.3 Cloning of kan^r gene into pD3UPDN

The vector pUC4K having the *kan^r* gene encoding for kanamycin resistance was digested with *EcoRI* to release a 1282bp long *kan^r* gene insert. The insert with the sticky ends was purified and ligated into the unique *EcoRI* site present in the pD3UPDN plasmid and named as pD3UPDNKan. The transformation of the DH5α *E.coli* pD3UPDNKan plasmid resulted in the gain of resistance to kanamycin by the ampicillin resistant parent vector. The transformed clones were checked for the gain in resistance to kanamycin (25µg/ml) and ampicillin along with restriction digestions that would result in the release of the insert from the vector.

2.2.4.4 Cloning of the recombination cassette into pPR27

The pPR27 suicide vector has a few restriction sites in its multiple cloning sites (MCS) and hence limits the option for selecting the restriction sites for cloning into it. The absence of the desired restriction site in pPR27 prompted to amplify the entire recombination cassette pD3UPDNKan using the primers pD3UPDNKaFor and pD3UPDNKaRev each carrying *XbaI* site. The resulting amplified product having the *XbaI* site was digested with *XbaI* and ligated into pPR27 vector digested with the same restriction digestion enzyme. The high GC-content present in the mycobacterial genome often interferes in the amplification of long DNA fragment as it have a tendency to assume strong secondary structures. Therefore, Phusion DNA polymerase (Fermentus) was used to amplify the recombination cassette, pD3UPDNKan. The ~3.28 kb PCR amplified product was digested with *XbaI* (Fermentus) and inserted into the unique *XbaI* site of pPR27 to generate the suicide delivery vector pPRD3KO. The *E. coli* cells successfully transformed with pPRD3KO exhibited resistance to kanamycin, ampicillin and gentamycin. The clones were further checked for release of insert from the vector backbone.

2.2.5 Transformation of M. smegmatis with rpPR27

The electroporation method was used for the transformation of *M. smegmatis*. 1-5µg of the recombinant plasmid pPRD3KO was added to eletrocompetent *M. smegmatis* cells and incubated on ice for 5 mins. Then the cells were transferred to cuvettes (Bio-Rad, 2mm) and subjected to electroporation (Bio-Rad electroporator) at

1.5kV/mm. After the pulse was given, 5 ml liquid MB-7H9 containing 2% glucose and 0.05% Tween 80 was added to the electroporated cells and allowed to grow till 6 hours at 30^0C for recovery. After 6 hours, 1ml of the culture was pelleted and plated onto MB-7H9 agar plates with gentamycin (10µg/ml) and 2% glucose at different dilution. The plates were incubated at 30^0C for 5 days.

2.2.6 Screening for clones with targeted disruption of eccD3 with kan^r in M. smegmatis

The colonies that appeared after transformation were restreaked on a MB-7H9 containing gentamycin (10µg/ml) and 2% glucose and kept at 30^0C to obtain patches. A loop of bacteria colonies from one of the patches were inoculated into liquid MB-7H9 medium containing kanamycin (25µg/ml), 2% glucose and 0.05% Tween-80 and grown in a shaker incubator at 30^0C for 36 hours. This step was critical for allowing the plasmid to increase its copy number in the cells and thus creating an environment for allelic exchange through homologous recombination to take place.

After the incubation period the cells were pelleted and plated onto MB-7H9 agar plates containing 10% sucrose and kanamycin (25µg/ml) at different dilutions and incubated at 39^0C to screen for double recombinants. The colonies obtained on these plates were again plated on MB-7H9 agar plates with glucose, sucrose and gentamycin (10µg/ml) to check for gentamycin sensitivity. Effectively, the clones with deleted copy of the *eccD3* gene would exhibit resistance to kanamycin and sucrose and sensitivity towards gentamycin. The single crossovers clone would retain the vector backbone and therefore exhibit sucrose sensitivity and gentamycin resistance.

2.2.7 Genomic DNA extraction from M. smegmatis

Mycobacterial genomic DNA was isolated as follows. *M. smegmatis* mc^{2}155 cells were harvested from 5ml of saturated cell culture (OD_{600} ~ 0.8-1) by centrifugation at 3,000g for 10mins and resuspended in 400µl of solution I (50mM Tris-HCl [pH 8.0], 50mg/ml lysozyme, 0.25 mg/ml RNase). After a 2 hours incubation step at 37^0C, 750µl of solution II (150mM Tris-HCl [pH 8.0], 100mM EDTA, 1% [wt/vol] SDS, 2mg/ml proteinase K) was added to the reaction mixture. The mixture was then incubated at 45^0C for 16 hours. The mycobacterial DNA was extracted with 5ml phenol-chloroform-isoamyl alcohol (25:24:1, vol/vol/vol). After centrifugation at 3,000g for 15 mins, the upper phase was washed twice with 5ml

chloroform-isoamyl alcohol (24:1, vol/vol). The mycobacterial DNA was precipitated by 0.7 volume of isopropanol in the presence of 0.1 volume of sodium acetate (3M), resuspended in distilled water, and stored at -20^{0}C.

2.2.8 Sequence analysis

In order to determine if the different vectors (pGEMT7Zf+, pPR27) used for the in frame deletion of the *eccD3,* contained the correct copy of the upstream and downstream gene of *eccD3,* the amplified copy of the respective gene was PCR purified with a PCR purification column kit (Qiagen) and subjected for sequencing analysis. The retrieved sequences were then viewed in the Chromas software and analyzed in the NCBI nucleotide BLAST (nBLAST).

2.2.9 PCR based confirmation of the mutant

2.2.9.1 Amplification using Set-1primers: Kan^{r}For and Kan^r Rev:

The *kan^r* gene specific primers (Forward primer-Kan^{r}For and Reverse primer-Kan^{r}Rev) were used for initial screening. Since the deletion event would result in the incorporation of the *kan^r* gene in place of the *eccD3* gene, the amplification of the gDNA of the putative knock out clone was done using specific primers for the *kan^r* gene. In the case of a successful targeted deletion of *eccD3,* the amplification would result in a product size of 1450bp comprising of the 1282 bp long of *kan^r*, 60 bp from the upstream region and 99 bp from the downstream region of *eccD3.* In the case of non-recombinants, the amplification of the *eccD3* gene would yield a product size of 975 bp. In case of single crossover events or mixed populations of cells, both the amplicons corresponding to 1450 bp and 975 bp would be obtained. The PCR was carried out with Dynazyme (Fermentus). The PCR reaction mixture was same as in Table 2.3. The thermal cycling parameters were kept same as Table 2.4 except that the annealing temperature was kept at 59^{0}C for 1 min.

2.2.9.2 Amplification using Set-2 primers: Ms0621For and Kan^{r}Rev

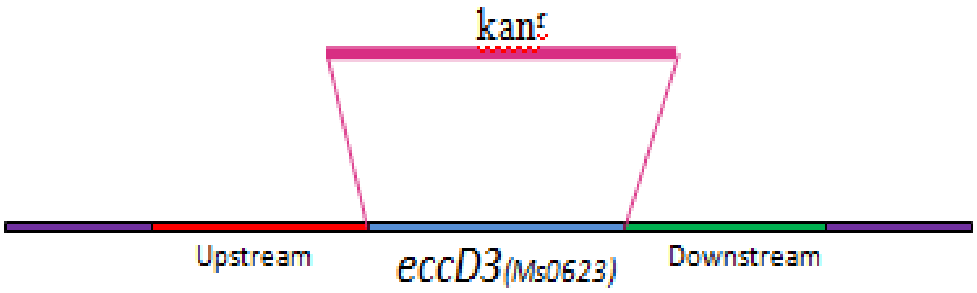
The amplification of the gDNA of the putative knockout clones with the above set of primers would confirm the incorporation of the *kan^r* gene and deletion of the *eccD3* gene but it would not help in identifying if the recombination event occurred at the correct position and the recombination cassette was incorporated in the correct sequence of genes. The amplification with the Kan^{r}For and Kan^{r}Rev as forward and reverse primers respectively would also take place if the recombination occurs at any

non-specific site. So, a set of primers were used which would result in the amplification of a fragment of DNA comprising of the upstream gene, the upstream fragment of the recombination cassette and the *kan*^{r} gene. The forward primer, Ms0621For was used such that it would anneal to the upstream region outside the construct, whereas, Kan^{r}Rev was used as the reverse primer is specific to the kanamycin resistance gene. This amplification would only take place if the recombination takes place at the desired site. This PCR amplification therefore confirms both the double crossover event and the position of the exchange. The expected amplicon size is 1730bp in the deletion mutant. The PCR was carried out with high fidelity Pfu DNA Polymerase using the same parameters as mentioned in Table 2.3 and Table 2.4. The annealing temperature was set at 59^{0}C for 1 min.

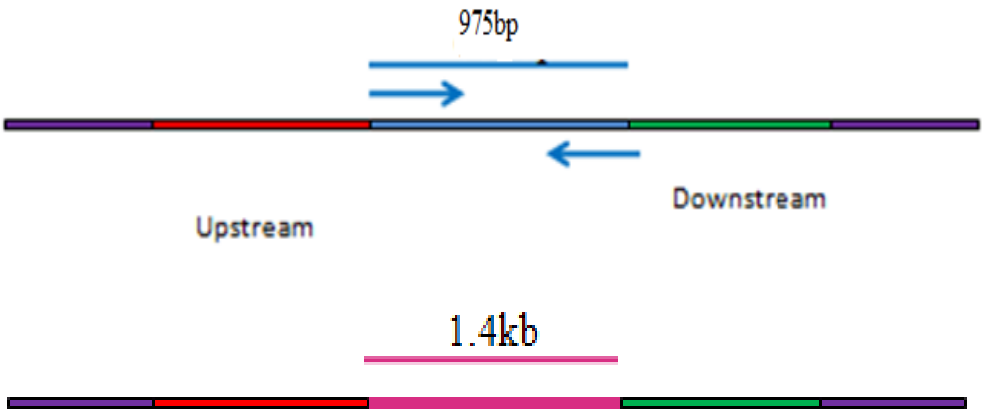
2.2.9.3 Amplification using Set-3 primers: Kan^{r}For and Ms0625Rev

This combination of primers would result in the amplification of a 2300 bp long amplicon in case of the double recombinant mutant whereas, in the wild type or single crossovers, it will produce a product of 1.7kb. Since this reaction will work for both double recombinants and wild type, so it is easy to standardize and also troubleshoot. Moreover, the result of this reaction will also be a clear indication of the site of recombination as *Kan*^{r}For is specific to the construct whereas, Ms0625Rev binds to a genomic region outside the construct. The reaction was performed with high fidelity Pfu DNA Polymerase using its specific buffer in the presence of 5% DMSO. The annealing temperature was kept at 59^{0}C for 1 min. The remaining conditions were kept the same as in Tables 2.3 and 2.4.

A.



B.



C.

D.

Figure 2.2: A-D: A schematic representation of the PCRs used for genotypic verification of the *eccD3* deleted clone in *M. smegmatis***.**

2.3 Results

2.3.1 Homology between the M. tubercuosis and M. smegmatis eccD3 gene

The amino acid sequence alingment of the msm MSMEG_0622 and Rv0290 corresponding to the EccD3 protein of *M. smegmatis* and *M. tuberculosis* (Figure 2.3) respectively showed that there was 61% identity between the two amino acid sequences that covered 98% of the query (Figure 2.4).

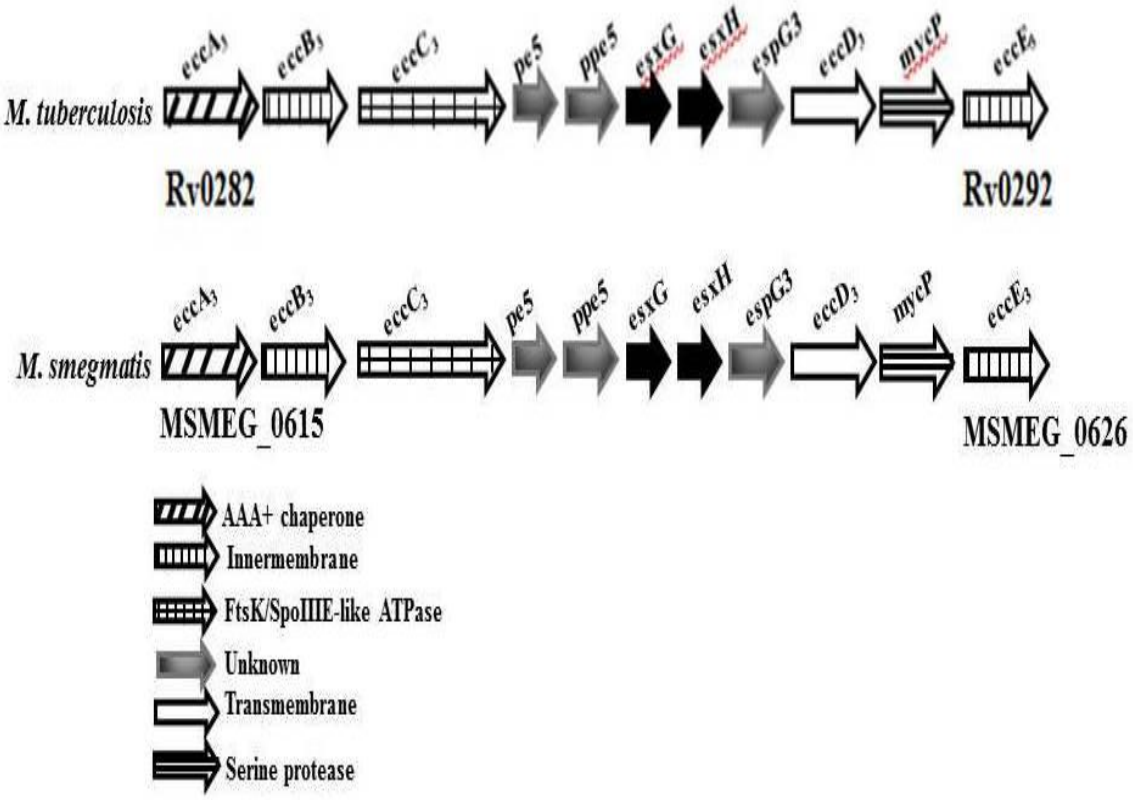


Figure 2.3: Schematic representation of the esx-3 loci from *M. tuberculosis* and *M. smegmatis*

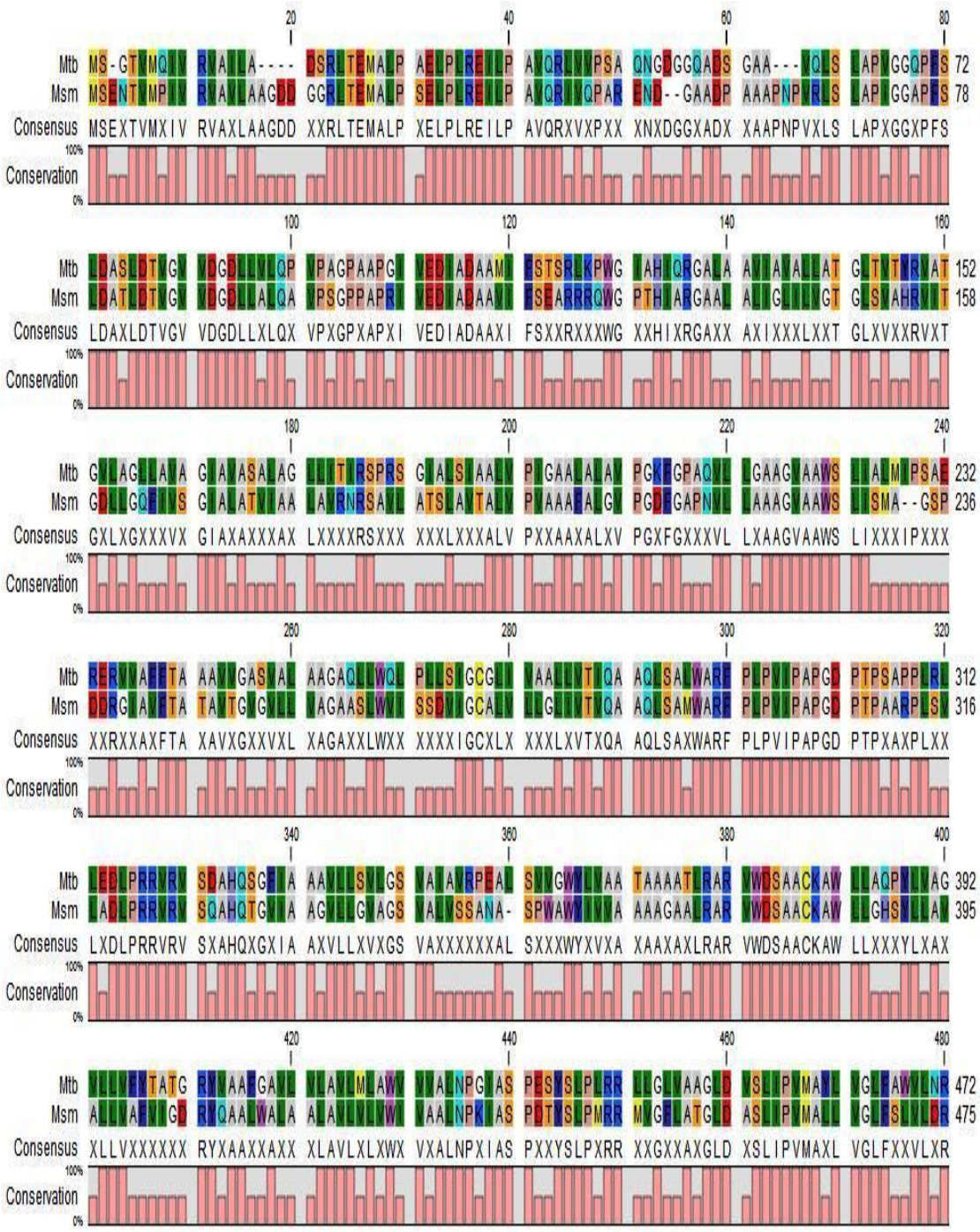


Figure 2.4: Sequence alignment of MSMEG_0622 and Rv0290.

2.3.2 Cloning of the upstream 1kb region of eccD3 into pGEM-7Zf(+) vector

A fragment of approximately 975bp DNA corresponding to the upstream region of the *eccD3* gene was amplified successfully. The DNA polymerase enzyme, DyNAzyme EXT polymerase (Fermentus) was used in this reaction. The DyNAzyme EXT polymerase has an advantage of amplifying GC rich region. In addition, the proofreading capacity of this enzyme allows minimal incorporation of wrong bases in the amplicon. The PCR conditions were standardized using a gradient thermal cycler

and the annealing temperature was fixed at 59°C (Figure 2.5[A]). The amplicon was then successfully digested with *XbaI* and *EcoRI* and ligated into *XbaI-EcoRI* digested pGEM- 7Zf(+). The clones showing resistance to ampicillin were screened and further verified. The plasmids isolated from the putative clones showed a distinct shift in size on 0.8% agarose gels (Figure 2.5[B]). The clones were further verified by restriction digestion and observed for the release of approximately1kb insert (Figure 2.6). Clone 1 was chosen for further study and renamed as pD3UP.

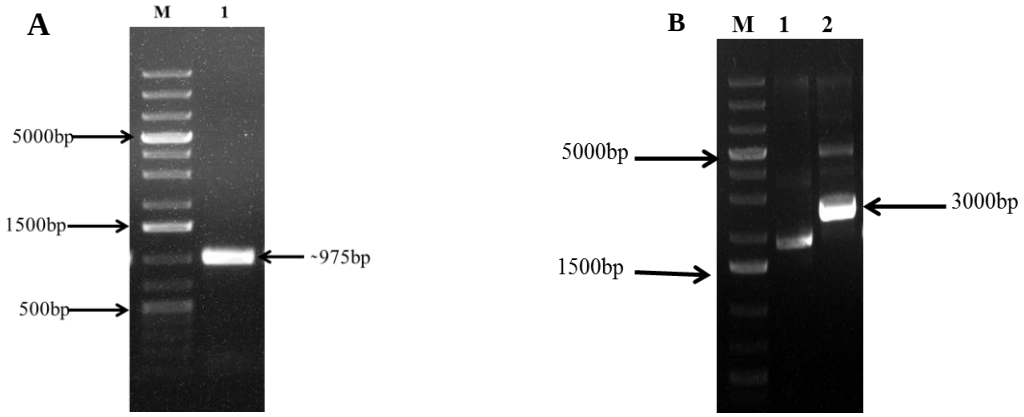


Figure 2.5: [A] 0.8% agarose gel showing the results of PCR amplification of 1 kb upstream region of *eccD3*. The Lane 1 in the gel display the products formed at 59˚C annealing temperature. [B] The gel showing the shift in mobility of plasmids isolated from the clones. Lane1: pGEM7Zf(+) plasmid, Lane2: pD3UP plasmid. Lane M: 1kbPlus DNA Ladder (Fermentus).

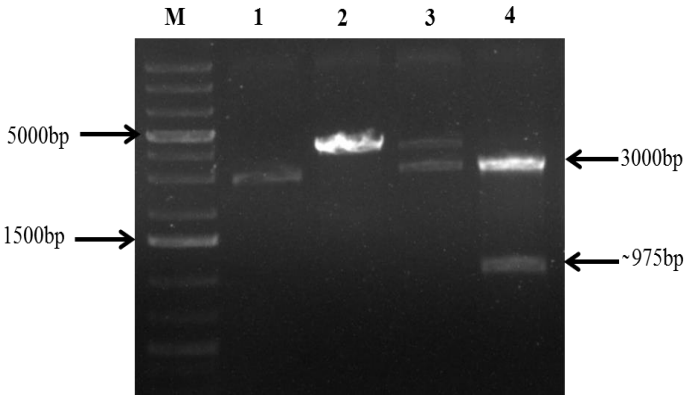


Figure 2.6: 0.8% agarose gel showing the results of restriction digestion of the plasmids isolated from the putative clones. *XbaI-EcoRI* digestion of the plasmids isolated from clone 1 shows an insert release of size 975 bp and therefore is confirmed clone of pD3UP. Lane M: 1kb plus DNA ladder (Fermentus) Lane1: undigested pGEM7Zf(+) vector; Lane2**:** XbaI digested pD3UP vector; Lane4:putative clone 1of pD3UP showing an insert release of ~1kb on *XbaI-EcoRI* digestion.

2.3.3 Sequencing of the upstream fragment cloned in pGEMT7Zf(+)

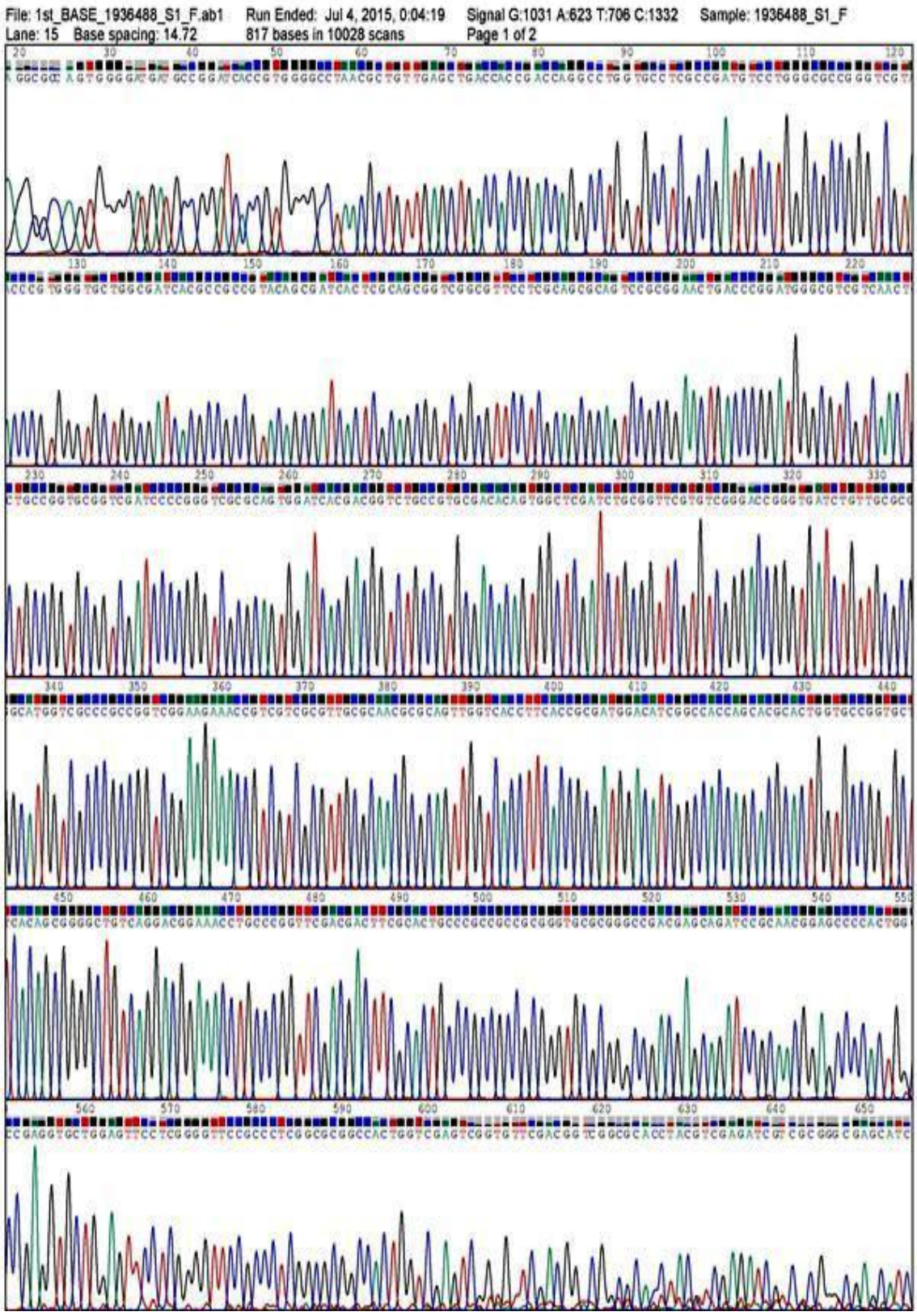


Figure 2.7: Chromatogram of the upstream fragment of *eccD3* gene viewed in the Chromas software.

Score	Expect	Identities	Gaps	Strand
1242 bits(672)	0.0	692/701(99%)	4/701(0%)	Plus/Plus

```
Query  1    GTGGGGCCTAACGCTGTTGAGCTGACCACCGACCAGGCCTGGTGCCTCGCCGATGTCCTG  60
Sbjct  49   GTGGGGCCTAACGCTGTTGAGCTGACCACCGACCAGGCCTGGTGCCTCGCCGATGTCCTG  108

Query  61   GGCGCCGGGTCGTACCCGTGGGTGCTGGCGATCACGCCGCCGTACAGCGATCACTCGCAG  120
Sbjct  109  GGCGCCGGGTCGTACCCGTGGGTGCTGGCGATCACGCCGCCGTACAGCGATCACTCGCAG  168

Query  121  CGGTCGGCGTTCCTCGCAGCGCAGTCCGCGGAACTGACCCGGATGGGCGTCGTCAACTCT  180
Sbjct  169  CGGTCGGCGTTCCTCGCAGCGCAGTCCGCGGAACTGACCCGGATGGGCGTCGTCAACTCT  228

Query  181  GCCGGTGCGGTCGATCCCCGGGTCGCGCAGTGGATCACGACGGTCTGCCGTGCGACACAG  240
Sbjct  229  GCCGGTGCGGTCGATCCCCGGGTCGCGCAGTGGATCACGACGGTCTGCCGTGCGACACAG  288

Query  241  TGGCTCGATCTGCGGTTCGTGTCGGGACCGGGTGATCTGTTGCGCGGCATGGTCGCCCGC  300
Sbjct  289  TGGCTCGATCTGCGGTTCGTGTCGGGACCGGGTGATCTGTTGCGCGGCATGGTCGCCCGC  348

Query  301  CGGTCGGAAGAAACCGTCGTCGCGTTGCGCAACGCGCAGTTGGTCACCTTCACCGCGATG  360
Sbjct  349  CGGTCGGAAGAAACCGTCGTCGCGTTGCGCAACGCGCAGTTGGTCACCTTCACCGCGATG  408

Query  361  GACATCGGCCACCAGCACGCACTGGTGCCGGTGCTCACAGCGGGGCTGTCAGGACGGAAA  420
Sbjct  409  GACATCGGCCACCAGCACGCACTGGTGCCGGTGCTCACAGCGGGGCTGTCAGGACGGAAA  468

Query  421  CCTGCCCGGTTCGACGACTTCGCACTGCCCGCCGCCGCGGGTGCGCGGGCCGACGAGCAG  480
Sbjct  469  CCTGCCCGGTTCGACGACTTCGCACTGCCCGCCGCCGCGGGTGCGCGGGCCGACGAGCAG  528

Query  481  ATCCGCAACGGAGCCCCACTGGCCGAGGTGCTGGAGTTCCTCGGGGTTCCGCCCTCGGCG  540
Sbjct  529  ATCCGCAACGGAGCCCCACTGGCCGAGGTGCTGGAGTTCCTCGGGGTTCCGCCCTCGGCG  588

Query  541  CGGCCACTGGTCGAGTCGGTGTTCGACGGTCGGCGCACCTACGTCGAGATCGTCGCGGGC  600
Sbjct  589  CGGCCACTGGTCGAGTCGGTGTTCGACGGTCGGCGCACCTACGTCGAGATCGTCGCGGGC  648

Query  601  GAGCATCGCGACGGCCACCGCGTCACCACCGAGGTGGGGGTCAGCATCATCGACACCCCA  660
Sbjct  649  GAGCATCGCGACGGCCACCGCGTCACCACCGATGTGGGGGTCAGCATCATCGACACCCCA  708

Query  661  CACGGCCGG-ATACTGG-TTCACCC-GACGAAA-GCCTTCG  697
Sbjct  709  CACGGCCCGTATACGGGGTTCACCCTGACGTAATGTCTTCG  749
```

Figure 2.8: Sequence alignment of the upstream gene cloned in pGEMT7Zf(+) and *M. smegmatis* MSMEG_0622. The upstream gene sequence cloned in the pGEM7Zf(+) showed 99% sequence similarity over a query cover of 79% with MSMEG_0622, the gene present upstream of the *eccD3* gene in *M. smegmatis*.

2.3.4 Cloning of the 1.2 kb region downstream of eccD3 into pD3UP

The amplification of the genomic region downstream of *eccD3* gene resulted in an amplicon size of approximately 1kb evident from the appearance of the band in a 0.8% agarose gel (Figure 2.9[A]). The 1kb amplicon was then digested with *EcoRI* and *HindIII* and cloned into *EcoRI-HindIII* digested pD3UP. The putative clones were screened for gain in ampicillin resistance and verified by relative mobility shift on 1% agarose gel (Figure 2.9[B]), and insert release of the cloned fragment (Figure 2.10). Further studies were conducted with clone 1, renamed as pD3UPDN.

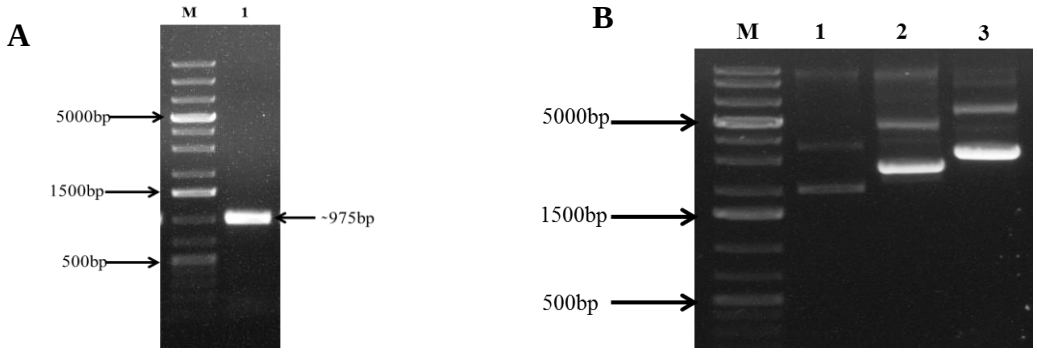


Figure 2.9: [A] 0.8% agarose gel showing the results of PCR amplification of~1kb region downstream of *eccD3* gene. Lanes 1: Amplified 1kb products at 59^0C annealing temperature. [B] 0.8% agarose gel showing the shift in mobility of plasmids isolated from the clones. Lane 1: pGEM7Zf(+) plasmid, Lane 2: pD3UP plasmid, Lane 3 pD3UPDN, Lane M: 1kbPlus DNA Ladder (Fermentus).

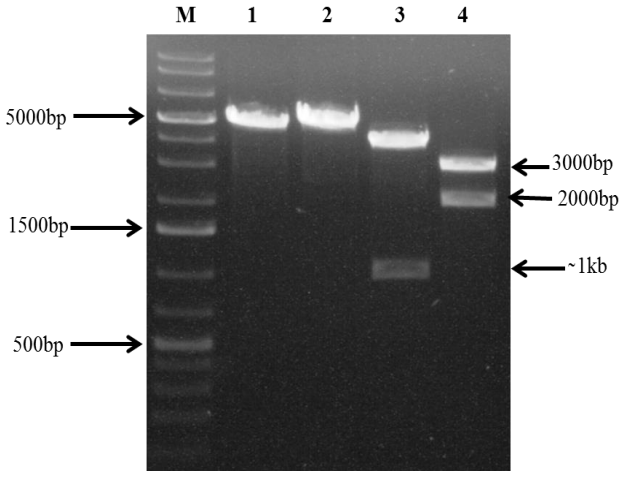


Figure 2.10: 1% agarose gel showing the results of insert specific digestions of pD3UPDN. Lane 1: *XbaI* digested pD3UPDN vector, Lane 2: *HindIII* digested pD3UPDN vector, Lane 3: putative clone 1of pD3UPDN showing an insert release of 1kb on *HindIII-EcoRI* digestion, Lane 4: putative clone 1of pD3UPDN showing an insert release of 2.2kb and the pEMT7Zf vector (3.3kb) on *XbaI-HindIII* digestion, Lane M: 1kb plus DNA ladder (Fermentus).

2.3.5 Sequencing of the downstream fragment cloned in pD3UP

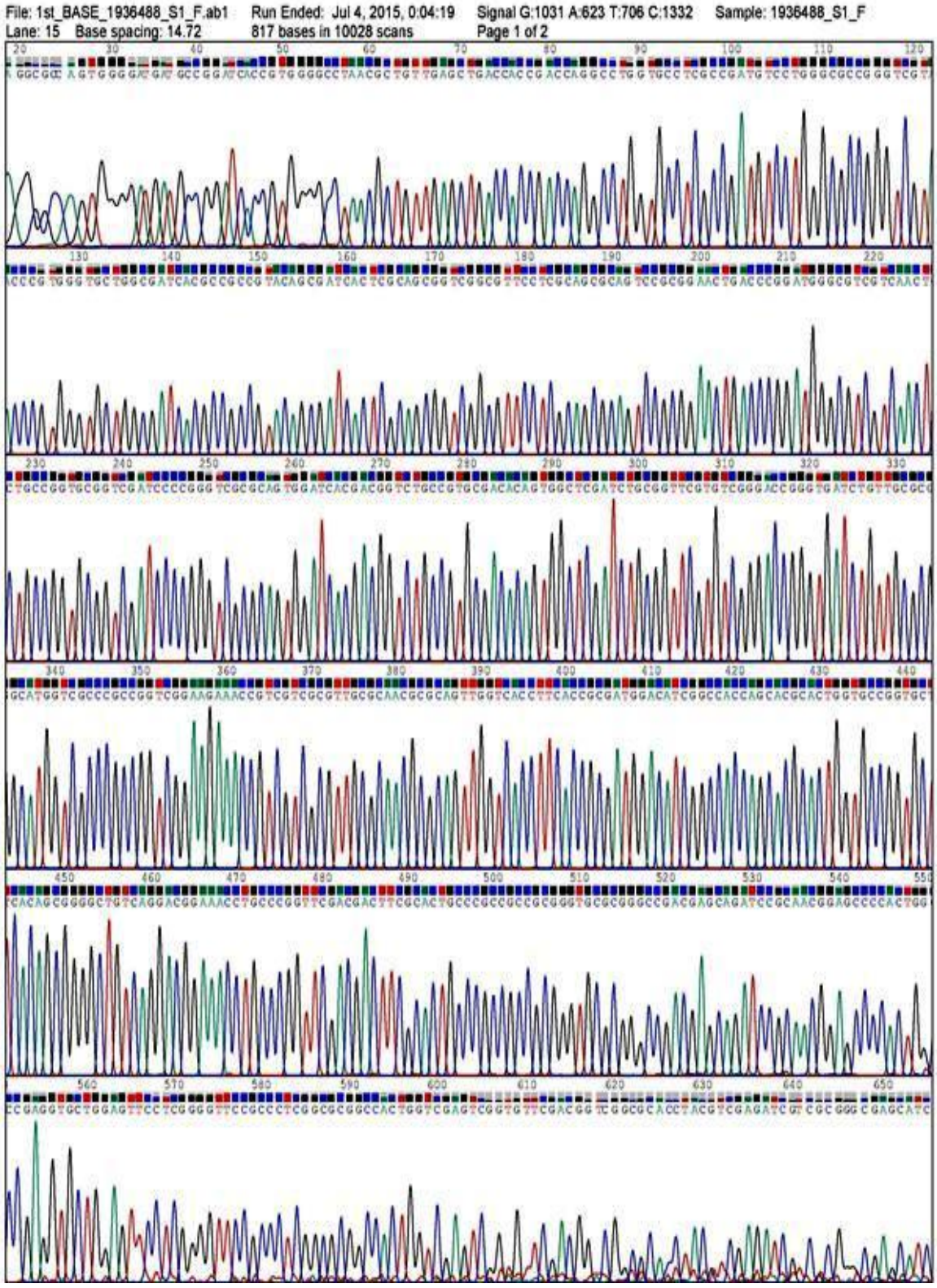


Figure 2.11: Chromatogram of the downstream fragment (MSMEG_0624) of *eccD3* gene cloned in pD3UP as viewed in the Chromas software.

Score	Expect	Identities	Gaps	Strand
675 bits(365)	0.0	390/403(97%)	1/403(0%)	Plus/Minus

```
Query  518  CAACCAGGGCTGCAAGCGATGCCGAGACCCTGGCGCGTGCCGTGGTGCGCGCGGCCGACA  577
            ||||||||| | ||| |||| |||| |||||||||| |||| |||||| ||||||||||
Sbjct  415  CAACCAGGACGTCAACCGACGCCCAGACCCTGGCGTGTGCGGTGGTGAGCGCGGCCGACG  356

Query  578  TGGGCGCGCGCGTCATCAACATCTCACTGGTGACGTGTCTGCCCGCCGACCGGACGATCG  637
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  355  TGGGCGCGCGCGTCATCAACATCTCACTGGTGACGTGTCTGCCCGCCGACCGGACGATCG  296

Query  638  ACCAGTCCGTACTCGGTGCGGCACTGCGCTATGCGGCGCTCGAGAAGGACGCCGTGATCG  697
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  295  ACCAGTCCGTACTCGGTGCGGCACTGCGCTATGCGGCGCTCGAGAAGGACGCCGTGATCG  236

Query  698  TCGCGGCCGCGGGCAACAACCGCGGCGGCGTCTCCACCGGCGCAGCGTGCGAATCGAATC  757
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  235  TCGCGGCCGCGGGCAACAACCGCGGCGGCGTCTCCACCGGCGCAGCGTGCGAATCGAATC  176

Query  758  CGCTGCCTTCGGGCACACCGGGCGATCCGCGCAACTGGAACGGCGTCACGTCGGTGTCCA  817
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  175  CGCTGCCTTCGGGCACACCGGGCGATCCGCGCAACTGGAACGGCGTCACGTCGGTGTCCA  116

Query  818  TCCCGTCGTGGTGGCAGCCCTACGTGCTGTCGGTGGGTGCGGTCGACTCCACCGGGCAGC  877
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  115  TCCCGTCGTGGTGGCAGCCCTACGTGCTGTCGGTGGGTGCGGTCGACTCCACCGGGCAGC  56

Query  878  CGTCGAGTTTCACGATGGCAGGTCCGTGGGTCGGGATCGCCGC  920
            |||||||||||||||||||||||||||||||  |||||||
Sbjct  55   CGTCGAGTTTCACGATGGCAGGTCCGTGGGTCGNN-TCGCCGC  14
```

Figure 2.12: Sequence alignment of the downstream gene cloned in pD3UP and *M. smegmatis* MSMEG_0624. The downstream gene sequence cloned in the pD3UP showed 97% sequence similarity with MSMEG_0624, the gene present downstream of the *eccD3* gene in *M. smegmatis*.

2.3.6 Cloning of kan^r gene into pD3UPDN

The fragment of DNA encoding kanamycin resistance, *kan*^{r} or *aph* was excised out from pUC4K plasmid by digesting with *EcoRI*. The excised 1282 bp insert was then successfully cloned into the unique *EcoRI* site of pD3UPDN to generate the clone pD3UPDNkan. The clones were screened for gain in resistance against kanamycin (25µg/ml) by the ampicillin resistant parent vector. Plasmids of the clones chosen from the above screening were isolated and the mobility shift of the plasmids was observed (Figure 2.13[A]). The clones were further verified by insert release of the cloned fragment (Figure 2.13[B]).

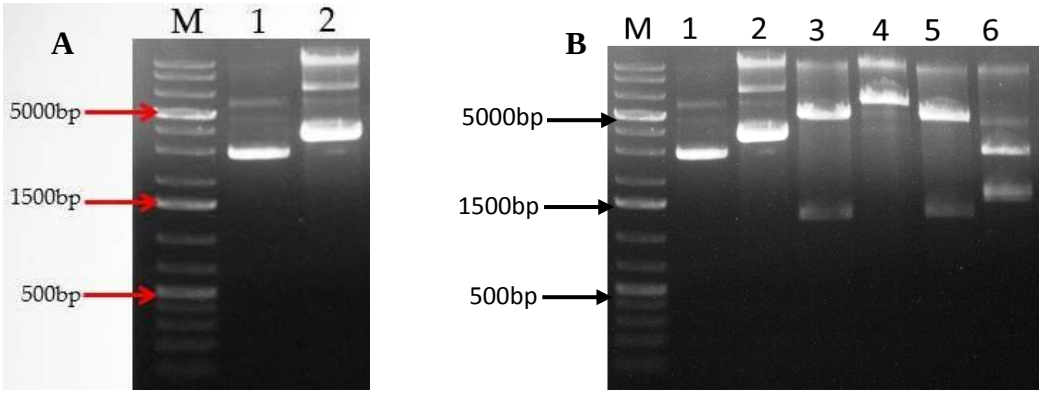


Figure 2.13: [A] 0.8 % agarose gel showing the band shift of the pUPDNKan plasmid. Lane 1: pD3UPDN plasmid, Lane 2: pD3UPDNKan plasmid, Lane M: 1kb plus DNA ladder (Fermentus). [B] 0.8% agarose gel showing the results of restriction digestion of the plasmids isolated from the putative clone. Lane 1: undigested pD3UPDN vector, Lane 2: undigested pD3UPDNKan vector, Lanes 3, 5: putative clone of pD3UPDNKan showing an insert release of 1kb on *EcoRI* digestion, Lane 4: XbaI digested pD3UPDNKan, Lane 6: putative clone showing a release of 3.28kb and the pEMT7Zf vector (3kb) on *XbaI-HindIII* digestion, Lane M: 1kb plus DNA ladder (Fermentus)

2.3.7 Cloning of the recombination cassette into pPR27

Although the use of the pPR27 shuttle vector presents several advantages, one disadvantage of using pPR27 (Figure 2.14) is that it provides a very limited choice of restriction sites in its MCS (multiple cloning sites). Because of the absence of the desired restriction site in the pPR27 vector, the entire construct from pD3UPDNkan *i.e.,* the recombination cassette was amplified with a set of primers having the *XbaI* restriction site. The appearance of a band corresponding to 3.28kb confirmed the successful amplification of the recombination cassette comprising of the upstream fragment, *kan*^{r} gene and the downstream fragment (Figure 2.15[A]). However, the PCR amplification could not be achieved by using the DyNAzyme II DNA Polymerase because of the high GC-content, strong secondary structure formation and the long size of the amplicon. Furthermore, the proofreading capacity of the DyNAzyme II DNA Polymerase limits the rate of amplification and processivity. Therefore, Phusion High- Fidelity DNA polymerase (Fermentus) having an enhanced processivity was used for this reaction. Its enhanced processivity domain ensures higher processivity in long range PCRs even on a difficult template with high GC content and strong secondary structures. The amplified product was digested with *XbaI* and ligated into the *XbaI* site of pPR27 to generate pPRD3KO. The plasmid isolated from the clones showed a distinct gel based mobility shift compared to the empty vector on a 0.8% agarose gel (Figure 2.15[B]). All the insert specific restriction

digestions worked and gave products of expected sizes. An insert release of 3.28 kb was obtained when digested with *XbaI* (Figure 2.16).

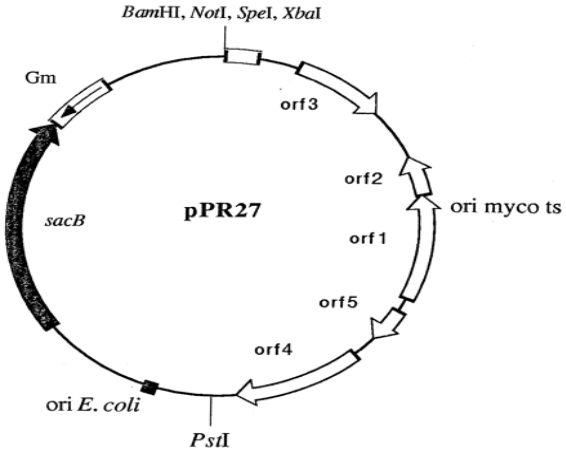


Figure 2.14: Vector map of pPR27, suicide vector.

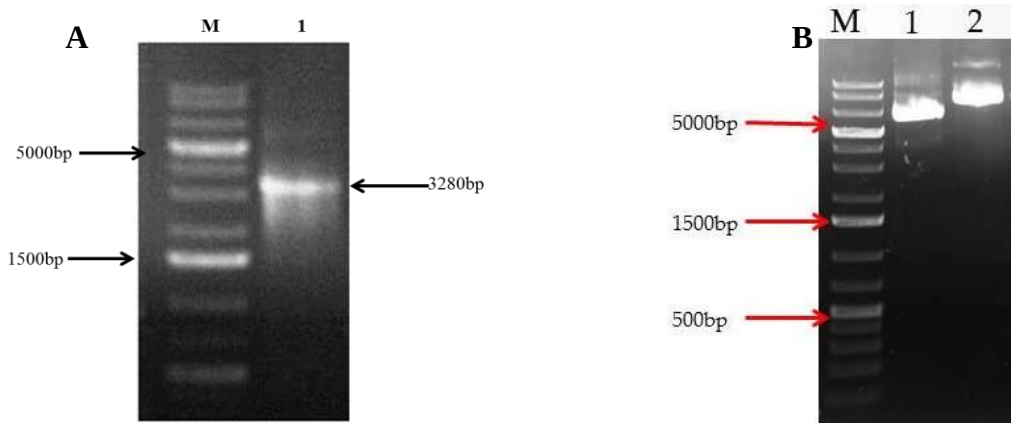


Figure 2.15: [A] 1% agarose gel showing the results of PCR amplification of 3.3 kb region from pD3UPDNkan corresponding to the recombination cassette. Lane 1: amplicon having XbaI as restriction site. [B] 0.8 % agarose gel showing the band shift of the pD3UPDNkan plasmid. Lane 1: pPR27plasmid, Lane 2: pPRD3KO plasmid, Lane M: 1kb Plus DNA Ladder.

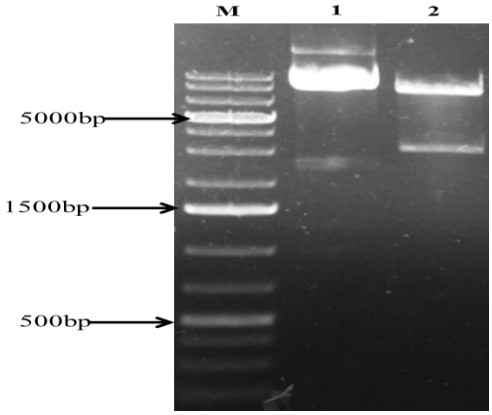


Figure 2.16: 1% agarose gel showing the results of restriction digestion of putative clones of pD3UPDNkan to confirm the clone. Lanes1: undigested pPRD3KO vector, Lane2: XbaI digested pPRD3KO vector showing insert release of 3.28kb, Lane M: 1kb plus DNA ladder (Fermentus).

2.3.8 Screening for clones with targeted deletion of eccD3 with kan^r in M. smegmatis

The pPR27 vector presents the advantage of screening the putative clones, as the vector comprises of selectable markers such as the *sacB* and *gent^r* gene and *ts-oriM* in the vector backbone. The putative clones transformed with pPRD3KO were screened for gain in resistance for both kanamycin and gentamycin (Figure 2.17). The clones were then checked for their ability to survive in 7H9 media containing 10% sucrose. The integration of the recombinant plasmid, pPRD3KO in the wild type *M. smegmatis* genome would result in the death of the cells because of the expression of the levansucrase enzyme from the *sacB* gene. However, in a situation of double cross over event, the *sacB* gene and the *gent^r* gene would be lost from the putative clones. Moreover, the growth of the putative clones at 39^0C ensured the incorporation of the *kan^r* gene in the genome and disintegration and loss of the suicide delivery vector (Figure 2.17). The table 2.5 summarizes the expected phenotypes of the wild-type carrying the plasmid, single crossovers and double crossovers.

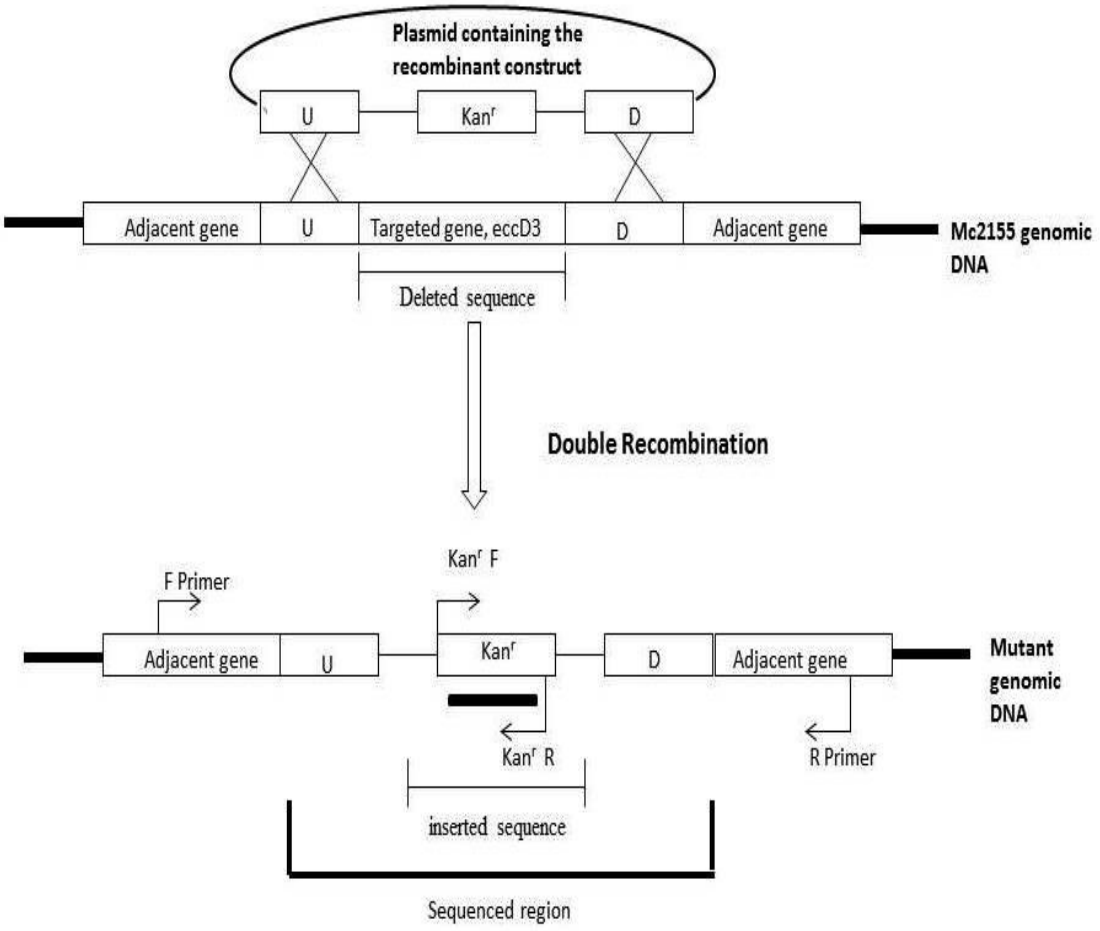


Figure 2.17: Strategies used to generate the *M. smegmatis eccD3* deletion mutant

Table 2.5: The table summarizes the expected phenotypes of the different mutants and the wild type.

Conditions	Single crossover	Double crossover	Nonspecific/WT
Kanamycin resistance (25µg/ml)	yes	yes	yes
Gentamycin resistance (10µg/ml)	yes	No	yes
Growth at 39^0C	yes	yes	No
Sucrose resistance	No	yes	-

Thus, the clones growing in sucrose containing media and exhibiting resistance to kanamycin and sensitivity to gentamycin are the putative *eccD3* deleted mutants. As can be seen in the Figure 2.18, certain colonies sensitive to gentamycin (Figure 2.19) were identified to be the putative *eccD3* deleted strains and subjected to further verification of the deletion event.

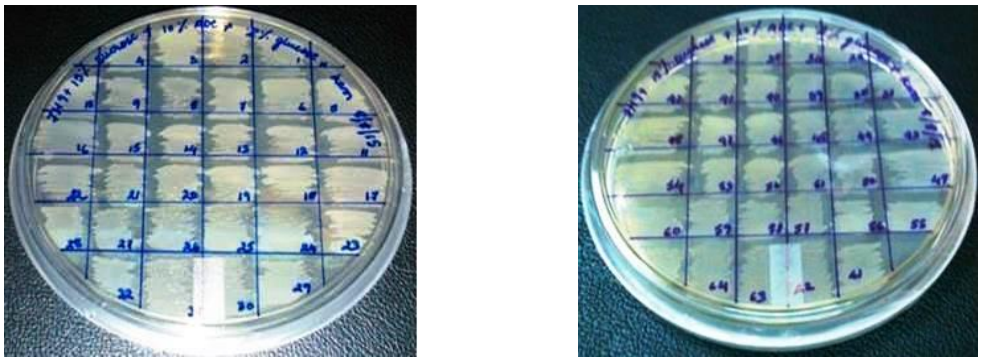


Figure 2.18: Putative colonies being checked for kanamycin resistance at 39^0C in 7H9 media supplemented with 10% ADC, 15% sucrose, 0.2% glucose and kanamycin.

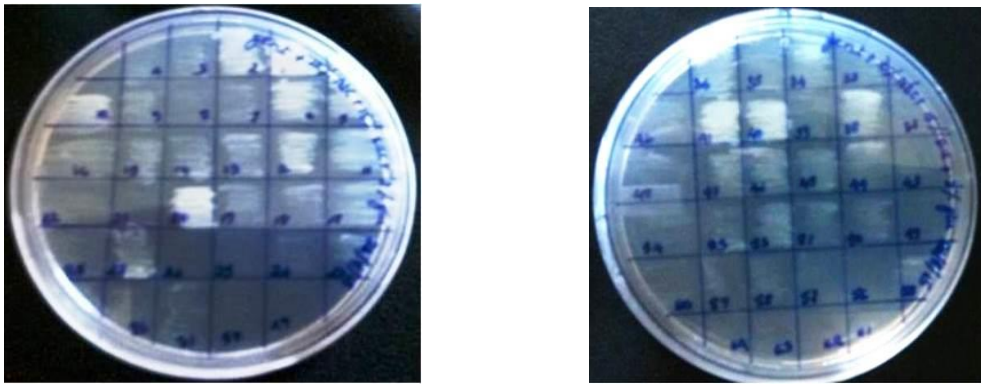


Figure 2.19: Checking for gentamycin sensitivity of the same clones as in Figure 2.18 The putative clones were then subjected to PCR reactions to verify the mutants.

2.3.9 PCR based identification of the eccD3 disruption mutant in M. smegmatis

The PCR reactions were standardized with the wild type genomic DNA and pPRD3KO plasmid. The results for the standardization are shown below:

2.3.9.1 PCR amplification using Set-1: kan^{r}For and kan^{r}Rev primers

The amplification of the genomic DNA of the putative clones with the targeted disruption of the *eccD3* gene with the primer set consisting of the *kan^r*For as the forward primer and *kan^r*Rev as the reverse primer was done using Polymerase Chain Reaction. The appearance of the specific band at 1.4kb demonstrated the successful insertion of the kanamycin resistant gene in place of the *eccD3* gene (Figure 2.20). The absence of the band corresponding to 975bp demonstrated the absence of the *eccD3* gene in the putative colony.

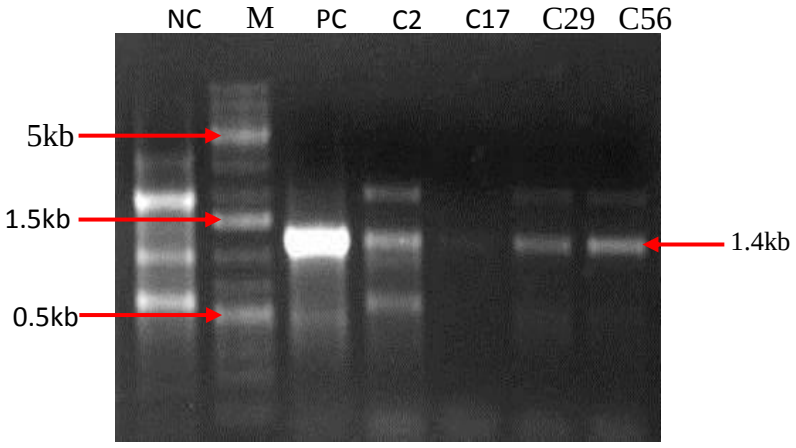


Figure 2.20: 0.8% agarose gel showing the products of PCR amplification with set 1primers. Lane NC- negative control, genomic DNA (gDNA) of mc^{2}155, Lane PC- positive control, pPRD3KO, Lane C2- gDNA of colony 2, Lane C17- gDNA of colony 17, LaneC29- gDNA of colony 29, Lane C56- gDNA of colony 56, Lane M- 1kb Plus DNA Ladder.

2.3.9.2 PCR amplification using Set 2: Ms0621For and Kan^{r}Rev

The amplification of the genomic DNA of the putative clones with deleted *eccD3* using Ms0621For as the forward primer and Kan^{r}Rev as the reverse primer resulted in the appearance of the expected band at 1.7kb (Figure 2.21). The band at 1.7kb demonstrated that the double recombination event was not non-specific and has taken place in the specific site between the upstream and downstream gene of *eccD3*.

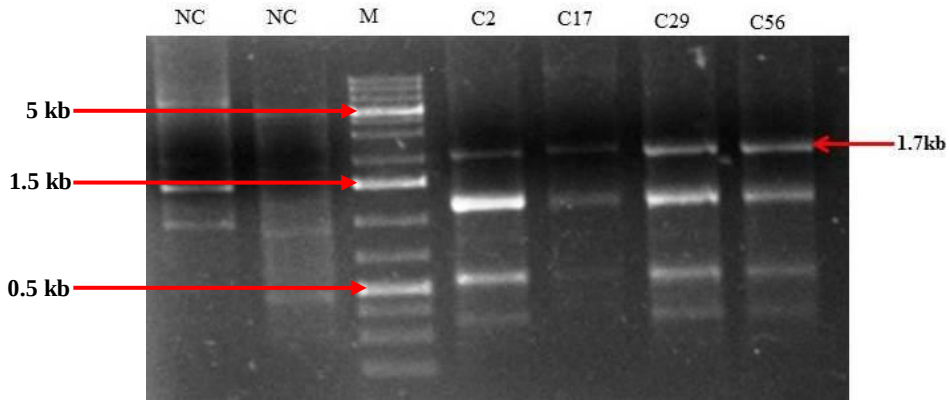


Figure 2.21: 0.8% agarose gel showing the products of PCR amplification with set 2 primers. Lane NC- negative control, gDNA of mc^{2}155, Lane NC- negative control pPRD3KO plasmid, Lane C2- gDNA of colony 2, Lane C17- gDNA of colony 17, LaneC29- gDNA of colony 29, Lane C56- gDNA of colony 56, Lane M- marker. Product size: 1730bp.

2.3.9.3 PCR amplification using Set 3: Kan^{r}For and Ms0625Rev

Out of the four putative colonies, colony 56 was selected as the mutant strain with the deleted copy of *eccD3*. The amplification of the genomic DNA of colony 56 with the primer combination, Kan^{r}For and Ms0625Rev as the forward and reverse respectively yielded a specific band at 2.3 kb which confirmed the site specific deletion of *eccD3* (Figure 2.22). The absence of any band at 1.7kb demonstrated the absence of a single cross over event. The final clone which fulfilled all the criteria of a successful double recombination was assigned the name, Ms0622.

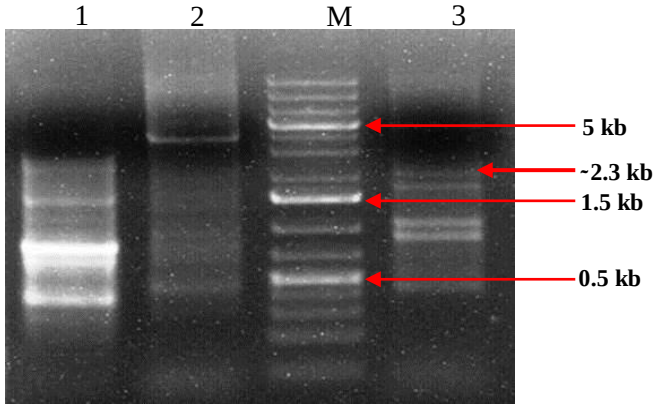


Figure 2.22: 0.8% agarose gel showing the products of PCR amplification with set 3 primers. Lane 1- negative control 1, gDNAof mc^{2}155, Lane2- Negative control2, pPRD3KO plasmid, Lane 3: gDNA of colony 56, Lane M- 1kb plus DNA Ladder.

2.4 Discussion

Genetic modifications such as horizontal gene transfer, deletion and insertion have led to the evolution of successful obligate pathogen such as *M. tuberculosis* and *M. leprae*. Members of the Mycobacteriaceae family can adapt to different

environmental niches and comprises of varied sizes of genome. It has been reported that the pathogenic Mycobacteria contains few of the homologous virulence causing gene and most of these virulence causing genes are present as orthologs in many non-pathogenic strain [28]. Insights into the dynamics of genomic evolution and difference between the pathogenic and non-pathogenic strain could shed light on pathogenicity causing genes thereby providing it as potent targets for drugs.

In the present study, an *eccD3 M. smegmatis* mutant has been generated. This was achieved by in-frame deletion of the gene using the efficient allelic exchange method. The use of the counter selectable marker helps in the appropriate screening of the recombinant clones. Several studies on the knockout of genes are done using different selectable markers such as *rpsL*, *pyrF*, and *sacB* [27, 29-31]. Out of all the selectable markers used in the study of *Mycobacterium*, *sacB* selectable marker is the most promising selectable marker. This selectable marker has been used in several studies including targeted disruptions of *ureC* gene in *M. bovis* and *M. tuberculosis* and the *erp* gene of *M. bovis* BCG and *M. tuberculosis* [25, 32, 33]. The *eccD3* gene in *M. smegmatis* was replaced with the *aph* gene conferring kanamycin resistance (*kan*^{r}) to the mutant strain with the help of the pPR27 suicide vector. We used pPR27 suicide vector for availing the advantage of efficient screening offered by the *sacB* gene and *ts-oriM* present in the vector backbone [34]. The wild-type cells transformed with the recombinant pPR27 (pPRD3KO) are capable of growing at temperatures above 39 ^{0}C and are resistant to both kanamycin and gentamycin. Moreover, the *sacB* gene from the plasmid backbone renders these cells sensitive to sucrose. In *M. smegmatis* allelic exchange occurs relatively easy in comparison to the slow growing mycobacteria. The present work describes the deletion of the gene of interest *i.e.*, *eccD3* through homologous recombination. The successful integration of the recombinant suicide vector into the *M. smegmatis* genome is essential for a homologous recombination event to take place. With the advent of the single crossover event, the recombinant pPRD3KO would integrate into the chromosome ensured by its ability to grow at 39^{0}C and death in the sucrose-containing medium. Whereas, a double recombination event ensured the loss of *gent*^{r} and *sacB* gene from the vector backbone. Thus, the colonies growing in sucrose containing media and exhibiting resistance to kanamycin and sensitive to gentamycin are the putative *eccD3* deleted mutants thereby, satisfying all conditions of a successful targeted deletion event. Out of the 64 putative colonies, clone 56 successfully surpassed all the

screening steps. The putative colony 56 subjected to PCR reactions to verify the deletion exhibited the desired band patterns. Thus, from the present study it can be concluded that *eccD3* gene had been deleted in the colony 56 of *M. smegmatis*.

Reference List

[1] Malhotra, S., Vedithi, S. C., & Blundell, T. L. Decoding the similarities and differences among mycobacterial species. *PLoS neglected tropical diseases*, 11(8): e0005883, 2017.

[2] Parish, T. and Stoker, N. G. The common aromatic amino acid biosynthesis pathway is essential in Mycobacterium tuberculosis. *Microbiology*, 148(Pt 10):3069-3077, 2002.

[3] Stanley, S. A., Raghavan, S., Hwang, W. W., and Cox, J. S. Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system. *Proceedings of the National Academy of Sciences*, 100(22):13001-13006, 2003.

[4] Gey Van Pittius, N. C., Gamieldien, J., Hide, W., Brown, G. D., Siezen, R. J., and Beyers, A. D. The ESAT-6 gene cluster of Mycobacterium tuberculosis and other high G+C Gram-positive bacteria. *Genome biology*, 2(10):research0044-1, 2001.

[5] Houben, E. N., Bestebroer, J., Ummels, R., Wilson, L., Piersma, S. R., Jimenez, C. R., Ottenhoff, T. H., Luirink, J., and Bitter, W. Composition of the type VII secretion system membrane complex. *Molecular microbiology*, 86(2):472-484, 2012.

[6] Bottai, D., Di Luca, M., Majlessi, L., Frigui, W., Simeone, R., Sayes, F., Bitter, W., Brennan, M. J., Leclerc, C., Batoni, G. Disruption of the ESX-5 system of Mycobacterium tuberculosis causes loss of PPE protein secretion, reduction of cell wall integrity and strong attenuation. *Molecular microbiology*, 83(6):1195-1209, 2012.

[7] Bitter, W., Houben, E. N., Bottai, D., Brodin, P., Brown, E. J., Cox, J. S., Derbyshire, K., Fortune, S. M., Gao, L. Y., Liu, J. Systematic genetic nomenclature for type VII secretion systems. *PLoS pathogens*, 5(10):e1000507, 2009.

[8] Mahairas, G. G., Sabo, P. J., Hickey, M. J., Singh, D. C., and Stover, C. K. Molecular analysis of genetic differences between Mycobacterium bovis BCG and virulent M. bovis. *Journal of bacteriology*, 178(5):1274-1282, 1996.

[9] Harboe, M., Oettinger, T., Wiker, H. G., Rosenkrands, I., and Andersen, P. Evidence for occurrence of the ESAT-6 protein in Mycobacterium tuberculosis and virulent Mycobacterium bovis and for its absence in Mycobacterium bovis BCG. *Infection and immunity*, 64(1):16-22, 1996.

[10] Simeone, R., Bottai, D., and Brosch, R. ESX/type VII secretion systems and their role in host-pathogen interaction. *Current opinion in microbiology*, 12(1):4-10, 2009.

[11] van der, W. N., Hava, D., Houben, D., Fluitsma, D.,van Zon, M., Pierson, J., Brenner, M., and Peters, P. J. M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells. *Cell*, 129(7):1287-1298, 2007.

[12] Simeone, R., Bobard, A., Lippmann, J., Bitter, W., Majlessi, L., Brosch, R., and Enninga, J. Phagosomal rupture by Mycobacterium tuberculosis results in toxicity and host cell death. *PLoS pathogens*, 8(2):e1002507, 2012.

[13] Serafini, A., Boldrin, F., Palu, G., and Manganelli, R. Characterization of a Mycobacterium tuberculosis ESX-3 conditional mutant: essentiality and rescue by iron and zinc. *Journal of bacteriology*, 191(20):6340-6344, 2009.

[14] Siegrist, M. S., Unnikrishnan, M., McConnell, M. J., Borowsky, M., Cheng, T. Y., Siddiqi, N., Fortune, S. M., Moody, D. B., and Rubin, E. J. Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proceedings of the National Academy of Sciences*, 106(44):18792-18797, 2009.

[15] Bottai, D., Di Luca, M., Majlessi, L., Frigui, W., Simeone, R., Sayes, F., Bitter, W., Brennan, M. J., Leclerc, C., Batoni, G. Disruption of the ESX-5 system of Mycobacterium tuberculosis causes loss of PPE protein secretion, reduction of cell wall integrity and strong attenuation. *Molecular microbiology*, 83(6):1195-1209, 2012.

[16] Demangel, C., Brodin, P., Cockle, P. J., Brosch, R., Majlessi, L., Leclerc, C., and Cole, S. T. Cell envelope protein PPE68 contributes to Mycobacterium tuberculosis RD1 immunogenicity independently of a 10-kilodalton culture filtrate protein and ESAT-6. *Infection and immunity*, 72(4):2170-2176, 2004.

[17] Coros, A., Callahan, B., Battaglioli, E., and Derbyshire, K. M. The specialized secretory apparatus ESX-1 is essential for DNA transfer in Mycobacterium smegmatis. *Molecular microbiology*, 69(4):794-808, 2008.

[18] Flint, J. L., Kowalski, J. C., Karnati, P. K., and Derbyshire, K. M. The RD1 virulence locus of Mycobacterium tuberculosis regulates DNA transfer in Mycobacterium smegmatis. *Proceedings of the National Academy of Sciences*, 101(34):12598-12603, 2004.

[19] Siegrist, M. S., Unnikrishnan, M., McConnell, M. J., Borowsky, M., Cheng, T. Y., Siddiqi, N., Fortune, S. M., Moody, D. B., and Rubin, E. J. Mycobacterial Esx-3 is required for mycobactin-mediated iron acquisition. *Proceedings of the National Academy of Sciences*, 106(44):18792-18797, 2009.

[20] Simeone, R., Bottai, D., Frigui, W., Majlessi, L., and Brosch, R. ESX/type VII secretion systems of mycobacteria: Insights into evolution, pathogenicity and protection. *Tuberculosis*, 95 Suppl 1:S150-S154, 2015.

[21] Husson, R. N., James, B. E., and Young, R. A. Gene replacement and expression of foreign DNA in mycobacteria. *Journal of Bacteriology*, 172(2):519-524, 1990.

[22] Balasubramanian, V., Pavelka, M. S., Jr., Bardarov, S. S., Martin, J., Weisbrod, T. R., McAdam, R. A., Bloom, B. R., and Jacobs, W. R., Jr. Allelic exchange in Mycobacterium tuberculosis with long linear recombination substrates. *Journal of bacteriology*, 178(1):273-279, 1996.

[23] Malaga, W., Perez, E., and Guilhot, C. Production of unmarked mutations in mycobacteria using site-specific recombination. *FEMS microbiology letters*, 219(2):261-268, 2003.

[24] Pashley, C. A., Parish, T., McAdam, R. A., Duncan, K., and Stoker, N. G. Gene replacement in mycobacteria by using incompatible plasmids. *Applied and environmental microbiology*, 69(1):517-523, 2003.

[25] Pelicic, V., Reyrat, J. M., and Gicquel, B. Generation of unmarked directed mutations in mycobacteria, using sucrose counter-selectable suicide vectors. *Molecular microbiology*, 20(5):919-925, 1996.

[26] Pavelka, M. S., Jr. and Jacobs, W. R., Jr. Comparison of the construction of unmarked deletion mutations in Mycobacterium smegmatis, Mycobacterium bovis bacillus Calmette-Guerin, and Mycobacterium tuberculosis H37Rv by allelic exchange. *Journal of bacteriology*, 181(16):4780-4789, 1999.

[27] Knipfer, N., Seth, A., and Shrader, T. E. Unmarked gene integration into the chromosome of Mycobacterium smegmatis via precise replacement of the pyrF gene. *Plasmid*, 37(2):129-140, 1997.

[28] Malhotra, S., Vedithi, S. C., and Blundell, T. L. Decoding the similarities and differences among mycobacterial species. *PLoS neglected tropical diseases*, 11(8):e0005883, 2017.

[29] Otero, J., Jacobs, W. R., Jr., and Glickman, M. S. Efficient allelic exchange and transposon mutagenesis in Mycobacterium avium by specialized transduction. *Applied and environmental microbiology*, 69(9):5039-5044, 2003.

[30] Sander, P., Meier, A., and Bottger, E. C. rpsL+: a dominant selectable marker for gene replacement in mycobacteria. *Molecular microbiology*, 16(5):991-1000, 1995.

[31] Snapper, S. B., Melton, R. E., Mustafa, S., Kieser, T., and Jacobs, W. R., Jr. Isolation and characterization of efficient plasmid transformation mutants of Mycobacterium smegmatis. *Molecular microbiology*, 4(11):1911-1919, 1990.

[32] Berthet, F. X., Lagranderie, M., Gounon, P., Laurent-Winter, C., Ensergueix, D., Chavarot, P., Thouron, F., Maranghi, E., Pelicic, V., Portnoi, D. Attenuation of virulence by disruption of the Mycobacterium tuberculosis erp gene. *Science*, 282(5389):759-762, 1998.

[33] Pelicic, V., Jackson, M., Reyrat, J. M., Jacobs, W. R., Jr., Gicquel, B., and Guilhot, C. Efficient allelic exchange and transposon mutagenesis in Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, 94(20):10955-10960, 1997.

[34] Stover, C. K., de, l. C., V,Fuerst, T. R., Burlein, J. E., Benson, L. A., Bennett, L. T., Bansal, G. P., Young, J. F., Lee, M. H., Hatfull, G. F. New use of BCG for recombinant vaccines. Nature, 351(6326):456-460, 1991.